Dividing and Discarding: A Procedure for Taking Decisions with Non-transferable Utility*

Vinicius Carrasco† William Fuchs‡

October 29, 2008

Abstract

We consider a setting in which two players must take a single action. The analysis is done within a private values model in which (i) the players' preferences over actions are private information, (ii) utility is non-transferable, (iii) implementation is bayesian and (iv) the welfare criterion is utilitarian. We characterize an optimal monotonic allocation rule. Instead of asking the agents to directly report their types, this allocation can be implemented dynamically. The agents are asked if they are to the left or to the right of the midpoint of the interval of possible types (e.g., $1/2$ for the initial interval $[0,1]$). If both reports agree, the section of the interval which none preferred is discarded and the remaining interval is divided in two parts, and the process continues until one agent chooses left and the other right. In that case, the midpoint of this remaining interval is implemented. This implementation can be carried out by a Principal who lacks commitment, implying this process is an optimal communication protocol.

1 Introduction

Many situations require two agents to take a joint action. Some examples are: managers of two different divisions within a firm, tariff negotiations in a trade block, Monetary Union members deciding on monetary policy, parties in a political coalition, and two members of a household.

Before reaching a decision, it is common for the agents to be involved in long conversations or negotiations that can take several rounds. Typically, a broad set of alternatives is considered at first, and slowly the set of alternatives "on the table" is refined until a decision is reached.

A conflict naturally arises between the agents' incentive to share information so that a better decision is taken and their fear that, if they reveal too much information, the other party might take advantage of it. A way around this problem is to reveal information coarsely at first, and slowly refine it as agents learn their interests are more aligned. In this case, by sharing more information, a better decision for both can

*We have benefited from conversations with Ricardo Alonso, Simon Board, Wouter Dessein, Sergio Firpo, Niko Matouschek, Juan Torres Martinez, Humberto Moreira, Roger Myerson, Alessandro Pavan, Phil Reny, Leo Rezende, Yuliy Sannikov, Andy Skrzypacz and Balazs Szentes, and seminar participants at MIT, Chicago, SED, GTS meetings, Chicago GSB, Fucape, PUC and EPGE-FGV. We are particularly thankful to Ferdinando Monte for careful comments on an early draft. Antonio Sodre provided excellent research assistance.

†Department of Economics, PUC-Rio.

‡Department of Economics, University of Chicago.

be attained. In contrast, once the agents learn their positions conflict there is no more scope for further communication.

In this paper, we study the problem of finding an optimal mechanism for a setting in which two agents have to take a joint action. We consider the case in which the players' preferences over actions are private information and uniformly distributed. Utility is non-transferable, the common action to be chosen belongs to an interval, and implementation is Bayesian. The lack of transfers and the focus on Bayesian implementation makes this a very hard problem to solve since standard mechanism design techniques cannot be readily applied.[1]

Finding an optimal joint action is hard because the scope for the players to misreport their preferences is very large and the instruments available to induce truthfulness are limited. We further restrict attention to the case where the Agent's preferences only depend on their own private information. Hence, unlike other papers that analyze decisions in committees, there is no advantage in sharing information to uncover some underlying truth.[2] This assumption lowers the incentives for agents to be truthful.

When transfers are available, and the players's utility is quasi-linear and satisfy a single crossing condition, one can solve for the optimal allocation using the virtual utility representation of the players´ preferences and applying standard maximization techniques. Once an optimal is found (and upon satisfying some monotonicity requirements), one can back up the necessary transfers to satisfy incentive compatibility.[3] In general, solving for optimal mechanisms with non-transferable utility and Bayesian implementation is a very hard problem.

Given the dificulty of the problem, we restrict the analysis to mechanisms that are monotonic type by type rather than mechanisms that are monotonic in expectation.[4] Also, instead of trying to characterize the optimal allocation directly we do it in two steps. First, we prove that we can weakly improve upon any monotonic mechanism that does not have $\frac{1}{2}$ as the allocation when the preferred actions of the players are on different sides of $\frac{1}{2}$. Second, we show that if an allocation has this property, to characterize the remaining part of the allocation the problem can be separated into two regions; one in which both agents are above $\frac{1}{2}$ and the one in which both are below $\frac{1}{2}$. Furthermore, finding the optimal monotonic allocation in these regions is, subject to rescalling, identical to the original problem. Therefore, the optimal allocation must be self-similar and we can easily characterize it.

The self-similarity of the solution relies heavily on the assumption that types are uniformly distributed, as the uniform distribution will continue being uniform if we condition on any particular region. Once we depart from the uniform case, it is hard to fully characterize the solution but we can show it will still be a step function.

For the uniform case, the optimal monotonic allocation rule is very simple and can be implemented with the following dynamic procedure. A Principal simultaneously asks the players if they are above or below the midpoint of the remaining choice set. If their reports agree we discard the section of the interval which none preferred, and continue dividing the remaining interval in this way until one chooses above and the other

[1] If the decision, choices and valuations were instead binary, a simple voting mechanism would be able to attain the efficient outcome. See, for example, the analysis of enforceable voting by Maggi and Morelli (2006). Alternatively, if transfers were possible and players had quasilinear utility, the problem could be easily solved using the expected externality mechanisms proposed by Arrow (1979), and d'Aspremont and Gerard-Varet (1979).

[2] See, for example, Persico (2004).

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

[4] The chosen allocation must be weakly increasing on the agents' reported types. See Section 3 for a precise definition.

below. In that case, the midpoint of the last remaining interval is implemented. We therefore name this mechanism the Divide and Discard mechanism (DD for short). Fleckinger (2008) has previously studied the DD allocation rule. However, his analysis is limited to quadratic preferences and, more importantly, and in contrast to our paper, rather than proving optimality, he only shows that the DD generates an improvement over the optimal ex-post incentive compatible allocation described in Moulin (1980).[5]

The DD is appealing for a number of reasons. In spite of the complexity brought up by the lack of side payments, the DD is an extremely simple mechanism. Its dynamic implementation resembles many real world situations (such as bilateral trade agreements). There are many rounds of negotiations, and alternatives are successively discarded until an agreement is reached. The principal provides a way to mitigate the conflict that arises between the agents' incentives to share information in order to achieve a better allocation, and their fear that, if they reveal too much information about their preferred actions, the other player may manipulate the allocation to his advantage. This is resolved by having agents report information coarsely at first, and gradually refine their reports as they learn that their interests are partially aligned. The DD hints on why contracting parties that negotiate sequentially may commit not to consider choices that were eliminated in previous rounds, i.e., agree to rule out – Hart and Moore (2007) –: as players move along further rounds of negotiation, they can be confident about the alignment of their interests. Therefore, whenever an agreement is reached, it will necessarily deliver an outcome that cannot be Pareto dominated by those that were ruled out.

The DD can be implemented by a principal who lacks commitment. This is important because in many circumstances, it might be difficult for a mediator or principal to commit to a mechanism. Within a firm, for instance, it is not clear that a CEO with authority will commit to not overrule the divisions' managers. The fact that the optimal mechanism can be implement even without commitment is also surprising. In general, we would expect that allowing the principal to commit would deliver strictly better outcomes.

The implementation of the DD allocation may call for several rounds of communication. Remarkably, for the case in which preferences are quadratic, the same expected value can be attained with just one round of cheap talk. The allocation that attains this value was derived by Alonso, Dessein and Matouschek (2008a) and (2008b) (ADM from now on).[6] In our setting, an advantage of long cheap talk is that the resulting allocation is renegotiation proof. With just one round of cheap talk, both players could report their preferred allocation to be in the same partition, but would not be allowed to divide this partition further into smaller subdivisions in search of a better allocation.

Moulin (1980), Barberà and Jackson (1994), and Barberà (2001) have studied the implementation of social choice functions in dominant strategies in more general settings than ours. In contrast to their work, we just require the allocations to be Bayesian incentive compatible. On the one hand, this makes the task of finding an optimal rule difficult as it is somewhat hard to pin down the set of all allocations that are interim incentive compatible. On the other hand, it allows for a better outcome for the players.

Our work also relates to the cheap talk literature. ADM extend Crawford and Sobel (1982) result that, if communication takes place just once players would communicate their private information coarsely by

[5] Moulin's phantom voter rule essentially places a ficticious player at 1/2 and then takes the median of the 3 reports as the allocation.

[6] Their very interesting papers focus mainly on the issue of centralized vs. decentralized decision making. In both cases, they consider quadratic preferences and only decision making with no commitment and one round of communication. Unlike our paper or ADM (2008b), in ADM (2008a) the authors allow for two actions and include a coordination failure cost from taking different actions. Our model relates to the limiting case in which this cost is prohibitively expensive.

reporting intervals rather than precise types. In the ADM model, as in ours, the existence of a set of states for which there is perfect alignment of incentives for all players makes it possible to have an (countably) infinite number of messages being sent even with just one round of communication.[7] In contrast to ADM and in the spirit of Krishna and Morgan (2004) and Aumann and Hart (2003), who analyze multistage communication, we allow for long cheap talk. By showing that it is an optimal monotonic mechanism, we prove that the communication protocol induced by DD is optimal in our setting. As the ADM allocation generates the same expected value as the DD, talking through partitions with infinite intervals is *also* an optimal communication protocol.

Goltsman et al (2008) solve for optimal communication protocols in the Crawford and Sobel (1982) model. They study three different processes: (i) (possibly long) cheap- talk (negotiations), (ii) non-binding recommendations by a third party (mediation), and (iii) binding recommendations by a third party (arbitration). They show that, if the misalignment of incentives is low, negotiation and mediation lead to the same outcome. Moreover, only two rounds of cheap talk are needed to obtain the mediation outcome. For most parameter values, however, arbitration always dominates the other protocols. We, in turn, show that arbitration and negotiations lead to the same outcomes in our setting.

In the next section we introduce the model.In Section 3 we establish the optimality of the DD allocation. In Section 4 we look at the dynamic implementation of the DD allocation and show it is implementable even if the principal lacks commitment. Section 4 also compares the DD implementation to an implementation with just one round of cheap talk. All proofs are relegated to the Appendix.

2 The Model

We consider a setting in which two ex-ante symmetric players, $i = 1, 2$, have to take a joint action a. Player $i's$ type is determined by his favorite action, θ_i. Types are independent and privately known by the players. We first consider the case in which $\theta_i \sim U[0,1]$.

The players' (Bernoulli) utility function is given by a twice continuously differentiable function $u(a, \theta_i)$, with $u(\theta_i, \theta_i) \geq u(a, \theta_i)$ for all a,

$$\frac{\partial^2 u(a, \theta_i)}{\partial a^2} < 0 < \frac{\partial^2 u(a, \theta_i)}{\partial \theta \partial a}$$

and such that any two equidistant actions from θ_i leads to the same utility, i.e.:

$$||a_1 - \theta_i|| = ||a_2 - \theta_i|| \Rightarrow u_i(a_1, \theta_i) = u_i(a_2, \theta_i).$$

The above conditions imply that the Agents' preferences are single peaked, and symmetric around the peak θ_i. We further impose the following additional monotonicity condition: for all non-decreasing $a(\theta_i, \theta_{-i})$,

$$\frac{\partial^2 u(a(\theta_i, \theta_{-i}), \theta_i)}{\partial \theta_i \partial a}(1 - \theta_i) \text{ and } -\frac{\partial^2 u(a(\theta_i, \theta_{-i}), \theta_i)}{\partial \theta_i \partial a}\theta_i \text{ are non-increasing} \qquad \text{(Monotone Hazard)}$$

This is the same condition as the one used by Athey et al (2005). As an example, note that, if $u(a, \theta_i) = -(a - \theta_i)^2$, the condition is trivially satisfied.

We follow Athey et al (2005) and in a slight abuse of terminology refer to the condition as monotone hazard condition.

[7] Melumad and Shibano (1991) also allow for the possibility of the receiver's and the sender's preferences coincide for some states.

3 An Optimal Monotonic Mechanism:

Before knowing their private types, the agents specify the rules of the mechanism by which the joint action will be chosen. Since we do not have quasilinear utilities and side payments, the agents' bargaining power will play an important role in the determination of the optimal mechanism. Given that the agents are ex-ante symmetric, we focus our analysis on the case in which, ex-ante, they choose the mechanism that maximizes the equally weighted sum of their utilities[8]

$$\sum_i E\left[u\left(a,\theta_i\right)\right].$$

The allocation rule (an enforceable contract) maps the players' reported types $\tilde{\theta} \equiv \left(\tilde{\theta}_1;\tilde{\theta}_2\right)$, and an independent randomization device $x \sim U[0,1]$ into an action:[9],[10]

$$a\left(\tilde{\theta}_i,\tilde{\theta}_{-i};x\right):[0,1]^2\times[0,1]\rightarrow\Re.$$

Before proceeding towards solving the problem above, it is convenient to recast it in terms of the agents' virtual utilities.[11]

3.1 Preliminaries

The objective is to find the incentive compatible allocation rule $a(\theta;x)$ that solves the following problem:

$$\begin{aligned} &\max_{a(\theta;x)} E_{\theta,x}\left(u(\theta_1,a)+u(\theta_2,a)\right)\\ s.t.\theta_i &\in \arg\max_{\tilde{\theta}_i} E_{\theta_{-i},x}\left[u\left(a\left(\widehat{\theta}_i,\theta_{-i};x\right),\theta_i\right)\right]\ \forall i\end{aligned}$$

In order to get the virtual utility representation, we start by making use of the following standard result:

Lemma 1 (IC Representation) *Letting*

$$U(\theta_i)=\max_{\tilde{\theta}} E_{\theta_{-i},x}\left[u\left(a\left(\widehat{\theta}_i,\theta_{-i};x\right),\theta_i\right)\right]=E_{\theta_{-i},x}\left[u\left(a\left(\theta_i,\theta_{-i};x\right),\theta_i\right)\right],$$

Incentive Compatibility is equivalent to:[12]

$$U(\theta_i)=\begin{cases} U(\frac{1}{2})+\int_{\frac{1}{2}}^{\theta_i} E_{\theta_{-i},x}\left[u_{\theta_i}\left(a\left(\tau,\theta_{-i};x\right),\tau\right)\right]d\tau, & \text{if } \theta_i>\frac{1}{2}\\ U(\frac{1}{2})-\int_{\theta_i}^{\frac{1}{2}} E_{\theta_{-i},x}\left[u_{\theta_i}\left(a\left(\tau,\theta_{-i};x\right),\tau\right)\right]d\tau, & \text{if } \theta_i<\frac{1}{2}\end{cases}$$

[8] Harsanyi (1955) argues that if society has preferences that satisfy the von-Neumann and Morgenstern axioms, social choices should be made according to an utilitarian criterion. This, along with the fact that the players are identical ex-ante, would justify our criterion.

[9] We can, without loss, restrict attention to Direct Mechanisms. This follows from the Revelation Principle (Myerson, 1981).

[10] Even though we allow for general stochastic mechanisms, the randomization device is only used to preserve the symmetry in the allocation. There is no need for randomization in the allocation from an efficiency standpoint.

[11] See Myerson (1981) and Myerson's notes on virtual utility at http://home.uchicago.edu/~rmyerson/research/virtual.pdf

[12] The symmetry of the problem makes it natural to pick $\theta=\frac{1}{2}$ as the reference type. The reader might be more used to seeing highest or lowest type picked as the reference type. Note that this choice is in general arbitrary and made for convenience.

with $E_{\theta_{-i},x}\left[u_{\theta_i}\left(a\left(\tau,\theta_{-i};x\right),\theta_i\right)\right]$ *non-decreasing in* τ .

The proof follows from Milgrom and Segal (2002) and is detailed in the Appendix.

Incentive compatibility only requires $E_{\theta_{-i},x}\left[u_{\theta_i}\left(a\left(\theta_i,\theta_{-i};x\right),\theta_i\right)\right]$ to be non-decreasing. Given the difficulty of the problem, and in order to make progress, we restrict a bit further the set feasible allocations by considering only schedules that are non-decreasing, that is, schedules that satisfy:

$$a\left(\theta_i,\theta_{-i};x\right) \text{ non-decreasing in } \theta_i \text{ for all } \theta_{-i},x. \qquad \text{(Monotonicity)}$$

After integrating by parts the representation of $U(\theta_i)$ induced by Lemma (1), one can define our objective function as:

$$V\left(a\right)=\sum_{i=1}^{2}\left\{\begin{array}{c} E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)\right] \\ +\frac{1}{2}E_{\theta,x}\left[u_{\theta_i}\left(a\left(\theta_i,\theta_{-i};x\right),\theta_i\right)\left(1-\theta_i\right)|\theta_i>\frac{1}{2}\right] \\ -\frac{1}{2}E_{\theta,x}\left[u_{\theta_i}\left(a\left(\theta_i,\theta_{-i};x\right),\theta_i\right)\theta_i|\theta_i<\frac{1}{2}\right] \end{array}\right\} \qquad (1)$$

and recast the program of interest as:

$$\max_{a(\theta;x)} V\left(a\right) \qquad (2)$$

$$s.t.\ a\left(\theta;x\right)\ being\ Incentive\ Compatible$$

$$and\ satisfying\ Monotonicity. \qquad (3)$$

As mentioned in the introduction, we do not allow for transfers. This greatly increases the difficulty of the problem. In settings with quasi-linear preferences and side payments, any $a\left(\theta,x\right)$ that satisfies expected monotonicity can be made IC by a proper choice of a transfer function. Therefore, in such cases, it is correct to proceed by simply maximizing pointwise the Incentive Compatible representation of the players' utility, as it can always be made equivalent to the players' expected utility by choosing the right side payments. The fact that players cannot make side payments forces us to consider the Incentive Compatibility Constraints explicitly, which, in turn, makes the problem fairly hard.[13]

Given the difficulties outlined above instead of trying to characterize the whole allocation directly in one step we proceed by showing that it is without loss to set the allocation to 1/2 when the players are on different sides of 1/2. This is shown formally in Lemma (2). In turn, we show using Lemma (3) that this then allows us to completely describe the optimal mechanism.

[13] Among other difficulties, the set of IC schedules is not convex. Hence, we cannot directly rely on Lagrangian Theorems to find the optimal contract.

3.2 An Optimal Monotonic Mechanism

Using the Law of Iterated Expectations, we can write the objective functional in (2) as:

$$V(a) = \underbrace{\frac{1}{4}\left(E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i},x\right),\frac{1}{2}\right)|\theta_{-i}\geq\frac{1}{2},x<\frac{1}{2}\right]+E_{\theta,x}\left[u_{\theta_i}\left(a\left(\theta_i,\theta_{-i},x\right),\theta_i\right)\left(1-\theta_i\right)|\theta\geq\left(\frac{1}{2},\frac{1}{2}\right)\right]\right)}_{A}$$

$$\underbrace{\frac{1}{4}\left(E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i},x\right),\frac{1}{2}\right)|\theta_{-i}\leq\frac{1}{2},x<\frac{1}{2}\right]-E_{\theta,x}\left[u_{\theta_i}\left(a\left(\theta_i,\theta_{-i},x\right),\theta_i\right)\theta_i|\theta\leq\left(\frac{1}{2},\frac{1}{2}\right)\right]\right)}_{B}$$

$$+\underbrace{\frac{1}{4}\left(E_{\theta,x}\left[u_{\theta_i}\left(a\left(\theta_i,\theta_{-i},x\right),\frac{1}{2}\right)\left(1-\theta_i\right)|\theta_i>\frac{1}{2}>\theta_{-i}\right]-E_{\theta,x}\left[u_{\theta_i}\left(a\left(\theta_i,\theta_{-i},x\right),\frac{1}{2}\right)\theta_i|\theta_{-i}>\frac{1}{2}>\theta_i\right]\right)}_{C}$$

$$+\underbrace{\frac{1}{4}\left(E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i},x\right),\frac{1}{2}\right)|\theta_{-i}\geq\frac{1}{2},x>\frac{1}{2}\right]+E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i},x\right),\frac{1}{2}\right)|\theta_{-i}\leq\frac{1}{2},x>\frac{1}{2}\right]\right)}_{D}.$$

Now, for any non-decreasing allocation $a(\theta_i,\theta_{-i},x)$, consider replacing $a(\theta_i,\theta_{-i};x)$ over the region $\left[0,\frac{1}{2}\right)\times\left(\frac{1}{2},1\right]$ by

$$\widetilde{a}\left(\theta_i,\theta_{-i},x;\alpha\right)=\left(1-\alpha\right)a\left(\theta_i,\theta_{-i};x\right)+\alpha E_{\theta,x}\left[a\left(\theta_i,\theta_{-i};x\right)|\theta\in\left[0,\frac{1}{2}\right)\times\left(\frac{1}{2},1\right]\right]$$

where $\alpha\in[0,1]$.

We argue that this replacemnet has a positive effect on term C for α small. Indeed, note that

$$\begin{aligned}
&\frac{\partial}{\partial\alpha}\left[E_{\theta,x}\left[u_{\theta_i}\left(\widetilde{a}\left(\theta_i,\theta_{-i},x;\alpha\right),\theta_i\right)\left(1-\theta_i\right)|\theta_i>\frac{1}{2}>\theta_{-i}\right]\right]|_{\alpha=0}\\
=\;& E_{\theta,x}\left[\left(\begin{array}{c}\left[u_{\theta_i a}\left(a\left(\theta_i,\theta_{-i};x\right),\theta_i\right)\left(1-\theta_i\right)\right]\cdot\\ \left[E_{\theta,x}\left[a\left(\theta_i,\theta_{-i};x\right)|\theta\in\left[0,\frac{1}{2}\right]\times\left[\frac{1}{2},1\right]\right]-a\left(\theta_i,\theta_{-i};x\right)\right]|\theta_i>\frac{1}{2}>\theta_{-i}\end{array}\right)\right]\\
\geq\;& \left[\left(\begin{array}{c}E_{\theta,x}\left[u_{\theta_i a}\left(a\left(\theta_i,\theta_{-i};x\right),\theta_i\right)\left(1-\theta_i\right)|\theta_i>\frac{1}{2}>\theta_{-i}\right]\cdot\\ E_{\theta,x}\left[E_{\theta,x}\left[a\left(\theta_i,\theta_{-i};x\right)|\theta\in\left[0,\frac{1}{2}\right]\times\left[\frac{1}{2},1\right]\right]-a\left(\theta_i,\theta_{-i};x\right)|\theta_i>\frac{1}{2}>\theta_{-i}\right]\end{array}\right)\right]\\
=\;& 0
\end{aligned}$$

where the inequality follows from the Monotone Hazard condition along with $\left[E_{\theta,x}\left[a\left(\theta_i,\theta_{-i};x\right)|\theta\in\left[0,\frac{1}{2}\right]\times\left[\frac{1}{2},1\right]\right]-a(\theta\right.$ being decreasing in θ_i, A similar argument can be made for the other component of term C.

This discussion suggests that the term C – that is associated with the region in which $\theta_i<1/2<\theta_{-i}$ – would be maximized by an incentive compatible mechanism that selects constant action. Note that a constant action of $\frac{1}{2}$ in this region also maximizes the term D in the objective. From the work in social choice theory by Moulin (1980), we know that if we required instead ex-post incentive compatibility the optimal allocation would also have 1/2 of diagonals.[14]

Nonetheless, once we consider Bayesian implementation, it is not obvious that it is optimal to set 1/2 as the allocation in these regions. We could expect that by perturbing the allocation slightly in these off-diagonal regions one could improve the attainable values on the on-diagonals $\left(\left(\frac{1}{2},1\right)^2 \textit{ and } \left(0,\frac{1}{2}\right)^2\right)$, once

[14] See also Barberà and Jackson (1994), and Barberà (2001) for detailed discussions of strategy-proof social choice functions.

incentive constraints are taken explicitly into account. Suppose we were to carry out such a perturbation in the region where $\theta_i < \frac{1}{2} < \theta_{-i}$. Note that we start from $a(\theta) = \frac{1}{2} < \theta_{-i}$ hence, if we were to make the allocation strictly increasing in θ_{-i} in this region player $-i$ would have more incentives to claim his type is higher than it actually is. This would not help us bring the allocation in $\left(\frac{1}{2}, 1\right)^2$ any closer to first best, since the problem with the first best allocation is exactly that types in this region would want to pretend they are higher than they actually are. Therefore, within the class of weakly increasing allocation rules it is best to set a constant $\left(a(\theta) = \frac{1}{2}\right)$ on the off-diagonals. The following lemma establishes this formally.

Lemma 2 (1/2 off-diagonals) *Given any symmetric incentive compatible allocation $a(\theta; x)$ that satisfies Monotonicity we can find an alternative incentive compatible allocation $\tilde{a}(\theta; x)$ which is weakly better and satisfies $\tilde{a}(\theta; x) = \frac{1}{2}$ for all x, when $\theta_i > \frac{1}{2} > \theta_{-i}$, and $\tilde{a}\left(\frac{1}{2}, \theta_{-i}; x\right) = \frac{1}{2}$ when $x > \frac{1}{2}$.*

This is a very powerful result towards the full characterization of the optimal allocation. Once one knows that setting $\frac{1}{2}$ off-diagonals is optimal – so that this region plays no role in terms of providing incentives over the main diagonal – the problem is separable. The following result states this in a precise way.

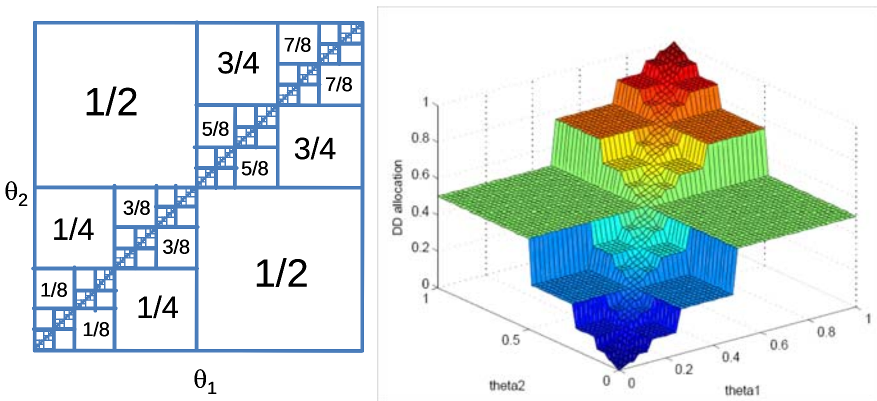
Lemma 3 (Separability) *Let $a^*(\theta, x)$ be an allocation that solves the program of interest. If $a^*(\theta; x) = \frac{1}{2}$ for all $\theta \in \left[0, \frac{1}{2}\right) \times \left(\frac{1}{2}, 1\right]$, and $a^*\left(\frac{1}{2}, \theta_{-i}, x\right) = \frac{1}{2}$ for all θ_{-i} whenever $x > \frac{1}{2}$ then:*

$$\begin{aligned} for\ \theta &\in [0, 1/2]^2,\ a^*(\theta, x) \in \arg\max_{a()} \sum_i E\left[u(a, \theta_i) \,|\theta \in (0, 1/2)^2\right] \\ s.t.\ IC\ for\ i &= 1, 2\ given\ \theta_{-i} \in (0, 1/2)\ \ and\ monotonicity \end{aligned}$$

and

$$\begin{aligned} for\ \theta &\in [1/2, 1]^2,\ a^*(\theta, x) \in \arg\max_{a()} \sum_i E\left[u(a, \theta_i) \,|\theta \in (1/2, 1)^2\right] \\ s.t.\ IC\ for\ i &= 1, 2\ given\ \theta_{-i} \in (1/2, 1)\ \ and\ monotonicity \end{aligned}$$

Furthermore, the problem over $\left[0, \frac{1}{2}\right]^2$, and respectively $\left[\frac{1}{2}, 1\right]^2$ is, subject to rescaling, exactly the same as the original problem (the problem over $[0, 1]^2$). So we can sequentially apply appropriately rescaled versions of Lemmas (2) and (3) to those regions. The resulting allocation from this iteration is graphed below.



We label this allocation the Divide and Discard allocation because of the following alternative to implementing this allocation as a direct revelation mechanism with the agents making just one report of their type. Instead of fully revealing their type in one round of communication the allocation can be attained by having agents simultaneously report if they prefer an allocation below or above the midpoint of the interval of possible types ($\frac{1}{2}$ for the initial interval $[0,1]$). If their reports fall on different sides of the midpoint, then the midpoint is implemented, if they both report to be on the same side of the midpoint then the process is restarted considering only the interval they both reported. For example the $[0,\frac{1}{2}]$ interval if they both reported "below" or the $[\frac{1}{2},1]$ interval if they reported "above". This process is iterated until agents eventually report to be on different sides of the relevant midpoint.[15]

In the next section we study the dynamic implementation of the DD allocation in detail but first, we finish this section by summarizing our main result:

Theorem 1 *The DD allocation is optimal in the class of non-decreasing Incentive Compatible allocations.*

3.2.1 Beyond the Uniform Case

The self-similarity and the precise cutoff values of the DD allocation depend crucially on the fact that uniform distribution is symmetric everywhere. Nonetheless, the forces behind having a constant action in Lemma 2 are more general. As we show below, with more general distributions, the optimal monotonic allocation will retain the property of being a step function if a monotone hazard condition holds.

Indeed, consider the case in which θ_i is drawn from $F(.)$, with density $f(.)$. In this environment, the Monotone Hazard is

$$\frac{\partial^2 u\left(a\left(\theta_i,\theta_{-i}\right),\theta_i\right)}{\partial\theta_i\partial a}\frac{\left(1-F\left(\theta_i\right)\right)}{f\left(\theta_i\right)} \text{ and } -\frac{\partial^2 u\left(a\left(\theta_i,\theta_{-i}\right),\theta_i\right)}{\partial\theta_i\partial a}\frac{F\left(\theta_i\right)}{f\left(\theta_i\right)} \qquad \text{(Monotone Hazard)}$$

are non-increasing in θ_i, for all non-decreasing $a(.)$.[16] For the special case in which $u(a,\theta_i)=-(a-\theta_i)^2$, this condition amounts to $\frac{(1-F(\theta_i))}{f(\theta_i)}$ and $-\frac{F(\theta_i)}{f(\theta_i)}$ being non-increasing, which is the standard monotone hazard condition used extensively in the mechanism design literature.

The next result provides a partial characterization of an optimal monotonic allocation for the non-uniform case:

Proposition 1 *There exists a partition of* $[0,1]^2$ *into sets* $\{A_{jk}\}_{j,k\in\{1,\ldots\}}$, *and numbers* $\{a_{jk}\}_{j,k\in\{1,\ldots\}}$ *such that an optimal monotonic allocation* $a(\theta,x)$ *has*

$$a(\theta,x)=a_{jk} \text{ if } \theta\in A_{jk}.$$

That is, for each set A_{jk}, *the chosen action does not depend on the players' types.*

Although, much as in the DD, an optimal monotonic allocation will be constant over certain regions, a full characterization of the relevant cutoffs (and, therefore, of the allocation itself) is very hard with general distributions.

[15] With zero probability the types are the same. In that case, the procedure described above would never stop, then simply take the limiting point as the allocation. Also, if at any point a player is indifferent between reporting either "below" or "above" we assume he flips a fair coin to decide.

[16] Again, this is the condition used by Athey et al (2005).

4 Dynamic Implementation of the DD Allocation

We now return to the case in which types are uniformly distributed. The DD captures in a simple way the property that negotiations often take place in rounds, and choices are sequentially eliminated until an agreement is reached.

We now verify that the agents have incentives to report truthfully in every round.

Proposition 2 *Thruthtelling is a dominant strategy at each stage of the "dynamic" implementation of the DD allocation rule.*

The proof is in the Appendix but the result is straightforward. At each stage, reporting the truth on average brings the decision closer to the Agent's preferred point. The only types who are indifferent are the cutoff types.

Additionally, this allocation rule can be implemented without commitment. Indeed, suppose, as is the case in ADM, that it is two managers of some firm that must report to the CEO, who will in turn decide. If the CEO cares equally about both divisions he does not need to commit in order to follow the dynamic implementation of the DD allocation.

Once it is common knowledge that the agents are on different sides of a midpoint, the mediator cannot extract any more beneficial information from them so he will choose the last midpoint as the allocation.

Proposition 3 *Suppose* $\underline{\theta} \leq \theta_i < \frac{\underline{\theta}+\bar{\theta}}{2} < \theta_{-i} \leq \bar{\theta}$*, then*:

$$\begin{aligned} a(\theta) &= \frac{\underline{\theta}+\bar{\theta}}{2} \in \arg\max_{a(\tilde{\theta})} E\left[\sum_i u\left(a(\theta), \theta_i\right) | \underline{\theta} \leq \theta_i < \frac{\underline{\theta}+\bar{\theta}}{2} < \theta_{-i} \leq \bar{\theta}\right] \\ & s.t \\ \theta_i &\in \arg\max_{\underline{\theta} \leq \tilde{\theta}_i < \frac{\underline{\theta}+\bar{\theta}}{2}} E_{\theta_{-i}}\left[u\left(a\left(\tilde{\theta}_i, \theta_{-i}\right), \theta_i\right) | \frac{\underline{\theta}+\bar{\theta}}{2} < \theta_{-i} \leq \bar{\theta}\right] \\ \theta_{-i} &\in \arg\max_{\frac{\underline{\theta}+\bar{\theta}}{2} < \tilde{\theta}_{-i} \leq \bar{\theta}} E_{\theta_i}\left[u\left(a\left(\tilde{\theta}_{-i}, \theta_i\right), \theta_{-i}\right) | \underline{\theta} \leq \theta_i < \frac{\underline{\theta}+\bar{\theta}}{2}\right] \end{aligned}$$

The Proposition above is stronger than required since it establishes that even with the ability to commit, once the agents know that they are on opposite sides of the midpoint, it is not efficient for the principal to extract any more information from them. It is then easy to verify that if all the mediator knows is that the players are in different sides of a given midpoint, his optimal choice for an allocation is the midpoint itself.

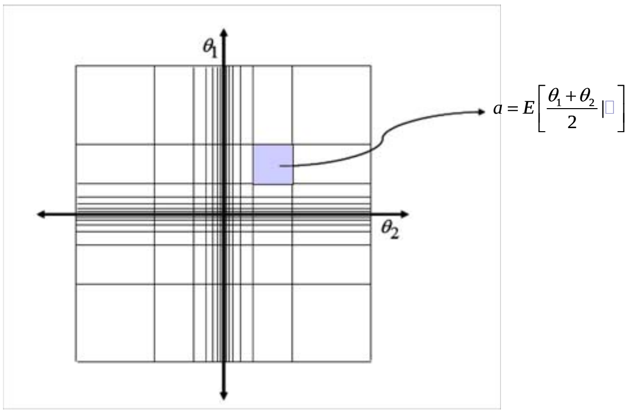
Fully revealing their types directly in one round of communication would fail if the mediator cannot commit. Once she learns the agents' type she would then want to deviate from whatever was promised and implement the first best allocation $a = \frac{\theta_i+\theta_{-i}}{2}$. Restricting the amount of information conveyed to the mediator is a way to prevent her from perturbing the allocation rule ex-post. Similar forces are behind the partition equilibria, in Crawford and Sobel (1982). In the next section we compare the DD allocation to the one that ADM obtain using partition equilibria.

Short and Long Cheap Talk In recent work, ADM studied the optimal allocation without commitment restricting their analysis to only one round of cheap talk by the Agents. As they acknowledge:

"It is well known in the literature on cheap talk games that repeated rounds of communication may expand the set of equilibrium outcomes even if only one player is informed. However, even for a simple cheap talk game such as the leading example in Crawford and Sobel (1982), it is still an open question as to what is the optimal communication protocol."

Surprisingly, for the quadratic case, the value attained with the allocation they characterize (for the extreme case of prohibitively high miscoordination costs, that corresponds to our model) is exactly the same as the one we attain with the DD allocation.[17] This implies that, in this environment, one round of communication is actually sufficient.[18] Hence, some of the conclusions derived by ADM are actually much stronger since the limitation of their analysis to one round of communication is actually of no consequence in terms of ex-ante payoffs.

Nonetheless, we believe that the dynamic implementation of the DD allocation with several rounds of communication has some additional appealing features. Before making our case, it is useful to recall what the allocation characterized by ADM looks like. Essentially, the type space is partitioned, each agent reports the element of the partition to which his favorite action belongs, and then the Principal implements as an allocation the average type given the reported rectangle. Partitions are very fine close to the middle of the interval since incentive constraints are not very binding for those types and progressively become coarser towards the extremes. Below, we replicate Figure 2 from their paper.



ADM Allocation

The first important thing to note is that the DD allocation is renegotiation proof, while the ADM allocation is not. For example, if both players report to be in the shaded area in the figure above they would have a strong incentive to communicate further. Essentially, they are facing the same situation they were facing originally albeit within a smaller range. So, although the CEO is not allowed to commit to an allocation, it is important that he *can* commit not to keep on talking. Instead, with the DD allocation, as long as agents report to be in the same quadrant, they would keep on refining their reports until it is clear their interests are in conflict. This happens when there is a value (the midpoint) that objectively separates both types.

Second, although ADM provide a very simple difference equation to characterize their allocation we find the simplicity of the DD mechanism very appealing.

[17] Simple computations show that both allocations deliver an ex-ante expected utility of $-\frac{2}{21}$. For comparison, the first best value is $-\frac{2}{24}$.

[18] See for example Aumann and Hart (2003) or Krishna and Morgan (2004) for more on the potential benefits of long cheap talk.

5 Concluding Remarks

This paper considered a setting in which two agents have to take a commonly agreed action for the case in which the players' preferences over actions are private information, there are no transfers, the action to be chosen rather than being binary belongs to an interval, and the welfare criterion is utilitarian.

The main results are as follows. The optimal monotonic allocation (which we label the DD allocation) can be implemented by simultaneously asking the players if they are to the left or to the right of the midpoint over the remaining choice set. If they both agree on the side of the coarse partition they prefer, we discard the section of the interval which none preferred and continue dividing the remaining interval in this way until one chooses left and the other right. In that case, the midpoint of the remaining interval is implemented. The DD allocation can also be implemented by a Principal without commitment, and, surprisingly, yields the same expected value as the one attained with just one round of cheap talk between the Agents, as in the mechanism in ADM.

References

[1] Alonso, R., Dessein, W. and Matouschek, N., 2008. "When Does Coordination Require Centralization?", *American Economic Review,* March 98(1), 145-179.

[2] Alonso, R., Dessein, W. and Matouschek, N., 2008. "Centralization versus Decentralization: An Application to Price Setting by a Multi-Market Firm." *Journal of the European Economic Association,* Papers and Proceedings, April-May.

[3] Arrow, K., 1979. "The Property Rights Doctrine and Demand Revelation under Incomplete Information", *Economics and Human Welfare.* Academic Press.

[4] Athey, S., Atkenson, A., and Kehoe, P., 2005. "The Optimal Degree of Monetary Policy Discretion". *Econometrica* 73 (5), September (2005), 1431-1476.

[5] Aumann, R.J. and Hart, S., 2003. "Long Cheap Talk", *Econometrica* Vol. 71 (6), pp. 1619–1660.

[6] Ausubel, L. and Deneckere, R., 1993. "A Generalized Theorem of the Maximum" , *Economic Theory*, Vol. 3, No. 1, pp. 99-107.

[7] Barberà, S. and Jackson, M., 1994. "A Characterization of Strategy-proof Social Choice Functions for Economies with Pure Public Goods", *Social Choice and Welfare,* 11, pp. 241-252.

[8] Barberà, S., 2001. "An Introduction to Strategy-proof Social Choice Functions", *Social Choice and Welfare,* pp. 619-653.

[9] Crawford, V. and Sobel, J., 1982. "Strategic Information Transmission", *Econometrica*, Vol. 50(6), pp. 1431-1451.

[10] d'Apresmont, C. and Gerard-Varet, L., 1979. "Incentives and Incomplete Information", *Journal of Public Economics* 11, pp. 25-45.

[11] Fleckinger, P. 2008. "Bayesian improvement of the phantom voters rule: An example of dichotomic communication", Mathematical Social Sciences 55 (2008).

[12] Goltsman, M., Hörner, J., Pavlov, G., and Squintani, F., 2008, "Arbitration, Mediation and Cheap Talk", mimeo, Department of Economics, Northwestern University.

[13] Harsanyi, J., 1955. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility", *The Journal of Political Economy,* Vol. 63, No. 4, pp. 309-321.

[14] Hart, O. and Moore, J., 2004. "Agreeing Now to Agree Later: Contracts that rule out but do not rule in", mimeo, Department of Economics, Harvard University.

[15] Krishna, V. and Morgan, J., 2004. "The art of conversation: eliciting information from experts through multi-stage communication", *Journal of Economic Theory*, Vol. 17(2), pp. 147-179.

[16] Maggi, G. and Morelli, M., 2006. "Self-Enforcing Voting in International Organizations", *American Economic Review,* Vol. 96(4), pp. 1137-1158.

[17] Melumad, N., and Shibano, T., 1991, "Communication in Settings with No Transfers", *Rand Journal of Economics,* Vol. 22, No. 2 (Summer, 1991), pp. 173-198

[18] Milgrom, P. and Segal, I., 2002. "Envelope Theorems for Arbitrary Choice Sets", *Econometrica* Vol. 70(2), pp. 583-601.

[19] Moulin, H., 1980. "On Strategy Proofness and Single Peakedness", *Public Choice,* 35, pp. 437–55.

[20] Myerson, R., 1981. “Optimal Auction Design”, *Mathematics of Operations Research*, Vol. 6(1) , pp. 58-73.

[21] Persico, N., 2004. "Committee Design with Endogenous Information", *Review of Economic Studies,* 71, pp. 165-91.

[22] Rosenlicht, M., 1999. *Introduction to Analysis,* Prentice-Hall.

[23] Royden, H.L., 1988, *Real Analysis,* Third Edition, Englewood Cliffs: Prenctice Hall.

[24] Segal, I.,2006. Communication in Economic Mechanisms, in Advances in Economics and Econometrics: Theory and Application, *Ninth World Congress (Econometric Society Monographs)*, ed. by Richard Blundell, Whitney K. Newey, and Torsten Persson, Cambridge University Press

6 Appendix

Appendix A: Preliminary Results

In this appendix, we prove some of the preliminary results we will need to prove the optimality of the DD, and some of the other results in the text.

Proof Lemma (IC Representation)

Proof. The proof is standard. For necessity, just notice that the integral formula is implied by Milgrom and Segal's (2002) Envelope Theorem. Moreover, if $a(\theta, x)$ is Incentive Compatible, for all $\theta' > \theta''$, one must have

$$E_{\theta_{-i},x}\left[u\left(a(\theta',\theta_{-i},x),\theta'\right)\right] \geq E_{\theta_{-i},x}\left[u\left(a(\theta'',\theta_{-i},x),\theta'\right)\right] \tag{IC$_{\theta'\theta''}$}$$

and

$$E_{\theta_{-i},x}\left[u\left(a(\theta'',\theta_{-i},x),\theta''\right)\right] \geq E_{\theta_{-i},x}\left[u\left(a(\theta',\theta_{-i},x),\theta''\right)\right] \tag{IC$_{\theta''\theta'}$}$$

Summing both expressions up,

$$E_{\theta_{-i},x}\left[u\left(a(\theta',\theta_{-i},x),\theta'\right)\right]-E_{\theta_{-i},x}\left[u\left(a(\theta',\theta_{-i},x),\theta''\right)\right] \geq E_{\theta_{-i},x}\left[u\left(a(\theta'',\theta_{-i},x),\theta'\right)\right]-E_{\theta_{-i},x}\left[u\left(a(\theta'',\theta_{-i},x),\theta''\right)\right]$$

Hence,

$$\int_{\theta''}^{\theta'} \left[E_{\theta_{-i},x}\left[u_{\theta_i}\left(a(\theta',\theta_{-i},x),\tau\right] - E_{\theta_{-i},x}\left[u_{\theta_i}\left(a(\theta'',\theta_{-i},x),\tau\right]\right] d\tau \geq 0$$

for all $\theta' > \theta''$. Since $u_{\theta_i a}(a, \theta_i) \geq 0$, the expected monotonicity condition must then hold.

For sufficiency*, let* $\theta_i > 0.5$*, and consider* $0.5 < \widehat{\theta} < \theta_i$.

$$\begin{aligned}
U_i(\theta_i) - U_i(\widehat{\theta}) &= \int_{\widehat{\theta}}^{\theta_i} E_{\theta_{-i},x}\left[u_{\theta_i}\left(a\left(\tau,\theta_{-i},x\right),\tau\right)\right] d\tau \geq \\
\int_{\widehat{\theta}}^{\theta_i} E_{\theta_{-i},x}\left[u_{\theta_i}\left(a\left(\widehat{\theta},\theta_{-i},x\right),\tau\right)\right] d\tau &= E_{\theta_{-i},x}\left[u\left(a(\widehat{\theta},\theta_{-i},x),\theta_i\right)\right] - E_{\theta_{-i},x}\left[u\left(a(\widehat{\theta},\theta_{-i},x),\widehat{\theta}\right)\right] \\
&= E_{\theta_{-i},x}\left[u\left(a(\widehat{\theta},\theta_{-i},x),\theta_i\right)\right] - U_i(\widehat{\theta})
\end{aligned}$$

where the first inequality follows from the expected monotonicity of the allocation.

Therefore,

$$U_i(\theta_i) \geq E_{\theta_{-i},x}\left[u\left(a(\widehat{\theta},\theta_{-i},x),\theta_i\right)\right].$$

The analysis for all other cases is analogous. ■

Proof of DD is IC. Let a^{DD} denote the DD allocation. Fix a given type $\theta_i \in [0,1]$ and let[19]

$$C(\theta_i) = \{\theta_{-i} : \theta_i \text{ is a cut-off point given the DD allocation}\}.$$

From the way the DD allocation is constructed,

$$E_{\theta_{-i},x}\left[u\left(a^{DD}(\theta_i+\gamma,\theta_{-i},x),\theta_i\right)|\theta_{-i}\in C(\theta_i)\right] = E_{\theta_{-i},x}\left[u\left(a^{DD}(\theta_i-\gamma,\theta_{-i},x),\theta_i\right)|\theta_{-i}\in C(\theta_i)\right],$$

for all $\gamma > 0$.

Hence,

$$\begin{aligned}
0 &= \frac{1}{\gamma}\left(E_{\theta_{-i},x}\left[u\left(a^{DD}(\theta_i+\gamma,\theta_{-i},x),\theta_i\right)|\theta_{-i}\in C(\theta_i)\right] - E_{\theta_{-i},x}\left[u\left(a^{DD}(\theta_i-\gamma,\theta_{-i},x),\theta_i\right)|\theta_{-i}\in C(\theta_i)\right]\right) \\
&= \frac{1}{\gamma}\left[\begin{array}{c} \left(E_{\theta_{-i},x}\left[u\left(a^{DD}(\theta_i+\gamma,\theta_{-i},x),\theta_i\right)|\theta_{-i}\in C(\theta_i)\right] - E_{\theta_{-i},x}\left[u\left(a^{DD}(\theta_i,\theta_{-i},x),\theta_i\right)|\theta_{-i}\in C(\theta_i)\right]\right) + \\ E_{\theta_{-i},x}\left[u\left(a^{DD}(\theta_i,\theta_{-i},x),\theta_i\right)|\theta_{-i}\in C(\theta_i)\right] - E_{\theta_{-i},x}\left[u\left(a^{DD}(\theta_i-\gamma,\theta_{-i},x),\theta_i\right)|\theta_{-i}\in C(\theta_i)\right] \end{array}\right].
\end{aligned}$$

[19] As examples, $C\left(\frac{1}{2}\right) = [0,1]$, and $C\left(\frac{3}{4}\right) = \left(\frac{1}{2},1\right]$.

Taking the limit as $\gamma \to 0$,

$$\frac{\partial E_{\theta_{-i},x}\left[u\left(a^{DD}\left(\widehat{\theta},\theta_{-i},x\right),\theta_i\right)|\theta_{-i}\in C\left(\theta_i\right)\right]}{\partial\widehat{\theta}}|_{\widehat{\theta}=\theta_i}=0.$$

Now, whenever $\theta_{-i}\in[0,1]\backslash C\left(\theta_i\right)$, $\frac{da^{DD}(\theta_i,\theta_{-i},x)}{d\theta_i}=0$, so that

$$\frac{\partial E_{\theta_{-i},x}\left[u\left(a^{DD}\left(\widehat{\theta},\theta_{-i},x\right),\theta_i\right)|\theta_{-i}\in[0,1]\backslash C\left(\theta_i\right)\right]}{\partial\widehat{\theta}}|_{\widehat{\theta}=\theta_i}=0.$$

Therefore,

$$\frac{\partial E_{\theta_{-i},x}\left[u\left(a^{DD}\left(\theta_i,\theta_{-i},x\right),\theta_i\right)\right]}{\partial\widehat{\theta}}=\left(\begin{array}{c}\Pr\left(\theta_{-i}\in C\left(\theta_i\right)\right)\frac{\partial E_{\theta_{-i},x}\left[u\left(a^{DD}\left(\theta_i,\theta_{-i},x\right),\theta_i\right)|\theta_{-i}\in C\left(\theta_i\right)\right]}{\partial\widehat{\theta}}\\+\left(1-\Pr\left(\theta_{-i}\in C\left(\theta_i\right)\right)\right)\frac{\partial E_{\theta_{-i},x}\left[u\left(a^{DD}\left(\theta_i,\theta_{-i},x\right),\theta_i\right)|\theta_{-i}\in[0,1]\backslash C\left(\theta_i\right)\right]}{\partial\widehat{\theta}}\end{array}\right)=0.$$

As the first order condition for truthtelling is satisfied, the Integral Representation of Lemma 6 holds. Since the DD allocation is non-decreasing, it also satisfies the expected monotonicty condition of Lemma 6.
■

Proof of Proposition 2 (Dynamic IC). Suppose that after a number of rounds the remaining segment of types under consideration is $\left[\underline{\theta},\bar{\theta}\right]$. Given any announcement of the other player, we characterize, in the table below, the allocation from reporting truthfully vs. lying. For concreteness, we will assume the true type of player one is smaller than the midpoint of the remaining segment.

Allocation	*Player* 2 $<\frac{\underline{\theta}+\bar{\theta}}{2}$	*Player* 2 $>\frac{\underline{\theta}+\bar{\theta}}{2}$
Report Below (Truth)	$a<\frac{\underline{\theta}+\bar{\theta}}{2}$	$a=\frac{\underline{\theta}+\bar{\theta}}{2}$
Report Above (Lie)	$a=\frac{\underline{\theta}+\bar{\theta}}{2}$	$a>\frac{\underline{\theta}+\bar{\theta}}{2}$

It can be seen clearly from the table above that reporting the truth is a dominant strategy since it always leads to an allocation that is closer to the agents preferred allocation and therefore a higher expected payoff. The only types that are indifferent are the midpoint types which we assume randomize in their reports. ■

Proof of Proposition 3. From Milgrom and Segal (2002), and the fact that the players' payoff satisfy a single crossing condition, it follows, using standard arguments (exactly the same as those in Lemma 6) that

$$\theta_1\in\arg\max_{\underline{\theta}\leq\tilde{\theta}_1<\frac{\underline{\theta}+\bar{\theta}}{2}}E_{\theta_2,x}\left[u\left(a\left(\tilde{\theta}_1,\theta_2,x\right),\theta_1\right)|\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\leq\bar{\theta}\right]$$

is equivalent to

$$E_{\theta_2,x}\left[u\left(a\left(\theta_1,\theta_2,x\right),\theta_1\right)|\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\leq\bar{\theta}\right]=\left\{\begin{array}{c}E_{\theta_2,x}\left[u\left(a\left(\frac{\underline{\theta}+\bar{\theta}}{2},\theta_2,x\right),\frac{\underline{\theta}+\bar{\theta}}{2}\right)|\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\leq\bar{\theta}\right]\\-\int_{\theta_1}^{\frac{\underline{\theta}+\bar{\theta}}{2}}E_{\theta_2,x}\left[u_{\theta_1}\left(a\left(\tau,\theta_2,x\right),\tau\right)|\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\leq\bar{\theta}\right]d\tau\end{array}\right\}\tag{ICLocal1}$$

and $E_{\theta_2,x}\left[u_{\theta_1}\left(a\left(\theta_1,\theta_2,x\right),\theta_1\right)|\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\leq\bar{\theta}\right]$ non-decreasing in θ_1, whereas

$$\theta_2\in\arg\max_{\frac{\underline{\theta}+\bar{\theta}}{2}<\tilde{\theta}_2\leq\bar{\theta}}E_{\theta_1,x}\left[u\left(a\left(\theta_1,\tilde{\theta}_2,x\right),\theta_2\right)|\underline{\theta}\leq\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}\right]$$

is equivalent to

$$E_{\theta_1,x}\left[u\left(a\left(\theta_1,\theta_2,x\right),\theta_2\right)|\underline{\theta}\leq\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}\right]=\left\{\begin{array}{c} E_{\theta_1,x}\left[u\left(a\left(\theta_1,\frac{\underline{\theta}+\bar{\theta}}{2},x\right),\frac{\underline{\theta}+\bar{\theta}}{2}\right)|\underline{\theta}\leq\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}\right] \\ +\int_{\frac{\underline{\theta}+\bar{\theta}}{2}}^{\theta_2}E_{\theta_1,x}\left[u_{\theta_2}\left(a\left(\theta_1,\tau,x\right),\tau\right)|\underline{\theta}\leq\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}\right]d\tau \end{array}\right\} \tag{ICLocal2}$$

and $E_{\theta_1,x}\left[u_{\theta_2}\left(a\left(\theta_1,\theta_2,x\right),\theta_2\right)|\underline{\theta}\leq\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}\right]$ non-decreasing in θ_2.
Integration of ICLocal1 and ICLocal2 by parts allows us to write the objective as

$$E_{\theta_2,x}\left[u\left(a\left(\frac{\underline{\theta}+\bar{\theta}}{2},\theta_2,x\right),\frac{\underline{\theta}+\bar{\theta}}{2}\right)|\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\leq\bar{\theta}\right]-E_{\theta,x}\left[u_{\theta_1}\left(a\left(\theta_1,\theta_2,x\right),\theta_1\right)\left[\theta_1+\underline{\theta}\right]|\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\right]$$
$$+E_{\theta_1,x}\left[u\left(a\left(\theta_1,\frac{\underline{\theta}+\bar{\theta}}{2},x\right),\frac{\underline{\theta}+\bar{\theta}}{2}\right)|\underline{\theta}\leq\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}\right]+E_{\theta,x}\left[u_{\theta_2}\left(a\left(\theta_1,\theta_2,x\right),\theta_2\right)\left(\bar{\theta}-\theta_2\right)|\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\right]$$

As we show in the First Step of Lemma (1/2 off-diagonals) below, for all non-decreasing $a\left(\theta_1,\theta_2,x\right)$,

$$\begin{aligned} & E_{\theta,x}\left[u_{\theta_2}\left(a\left(\theta_1,\theta_2,x\right),\theta_2\right)\left(\bar{\theta}-\theta_2\right)|\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\right] \\ \leq\; & E_{\theta}\left[u_{\theta_2}\left(E_{\theta,x}\left[a\left(\theta_1,\theta_2,x\right)|\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\right],\theta_2\right)\left(\bar{\theta}-\theta_2\right)|\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\right] \end{aligned}$$

Moreover, for any-non-decreasing $a\left(\theta_1,\theta_2,x\right)$, we have that

$$\begin{aligned} & -E_{\theta,x}\left[u_{\theta_1}\left(a\left(\theta_1,\theta_2,x\right),\theta_1\right)\left[\theta_1+\underline{\theta}\right]|\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\right] \\ \leq\; & -E_{\theta}\left[u_{\theta_1}\left(E_{\theta,x}\left[a\left(\theta_1,\theta_2,x\right)|\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\right],\theta_1\right)\left[\theta_1+\underline{\theta}\right]|\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\right]. \end{aligned}$$

Hence, it follows that both terms are maximized if one picks a constant. The best constant over those regions is $\frac{\underline{\theta}+\bar{\theta}}{2}$. Such constant also maximizes the terms

$$E_{\theta_2,x}\left[u\left(a\left(\frac{\underline{\theta}+\bar{\theta}}{2},\theta_2,x\right),\frac{\underline{\theta}+\bar{\theta}}{2}\right)|\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2\leq\bar{\theta}\right]$$

and

$$E_{\theta_1,x}\left[u\left(a\left(\theta_1,\frac{\underline{\theta}+\bar{\theta}}{2},x\right),\frac{\underline{\theta}+\bar{\theta}}{2}\right)|\underline{\theta}\leq\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}\right]$$

of the objective.
Hence, setting $a\left(\theta\right)=\frac{\underline{\theta}+\bar{\theta}}{2}$ for $\theta_1<\frac{\underline{\theta}+\bar{\theta}}{2}<\theta_2$ is optimal, as claimed. ■

Appendix B: The Optimality of the DD

In this appendix, we show that the DD is optimal among the class on non-decreasing Incentive Compatible allocations. Our strategy of proof involves two main steps.

The first step shows that, for any given class of allocations, if it is optimal to set $\frac{1}{2}$ off-diagonals, then the DD must be an optimal allocation within that class. The idea of the proof is straightforward. Given an uniform distribution, the problem of finding an optimal allocation, *conditional* on both agents having their types on the same region – say, above $\frac{1}{2}$ – and given that a constant action is being chosen when they are in different regions, is exactly the same as the original problem. In other words, if a constant action is chosen off-diagonals, so that the schedule off-diagonals plays no role in the provision of incentives, the problem self-replicates. Hence, the DD must be optimal.

In the second step, we show that if a non-decreasing schedule does not have $\frac{1}{2}$ off-diagonals, it can be weakly improved upon. We proceed in the following way. Starting with an arbitrary Incentive Compatible, non-decreasing, and continuous schedule $a(.)$, we show, using the same arguments as in the text, and ignoring Incentive Compatibility issues, that an improvement can be attained if one sets $\frac{1}{2}$ off-diagonals. We then move towards showing that further modifications of $a(.)$ along the main diagonals can be made so to guarantee both Incentive Compatibility, and that the new Incentive Compatible allocation still fares better than $a(.)$.[20] We then show, using limiting arguments, that an improvement can also be achieved if the initial non-decreasing allocation is not continuous. For this case, the improvement is not necessarily strict.

Steps 1 and 2 together prove Theorem1.

Separability and the Optimality of $\frac{1}{2}$ Off-Diagonals

Proof of Lemma Separability. The program of interest is

$$\max_{a(.,.)\text{ is IC over }[0,1]^2}\left(\begin{array}{c}\sum_i \frac{1}{4}\left[E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)|\theta_{-i}\leq\frac{1}{2},x<\frac{1}{2}\right]-E_{\theta,x}\left[u_{\theta_i}\left(a\left(\theta;x\right),\theta_i\right)\theta_i|\theta_i,\theta_{-i}<\frac{1}{2}\right]\right]\\ \sum_i \frac{1}{4}\left[E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)|\theta_{-i}>\frac{1}{2},x<\frac{1}{2}\right]+E_{\theta,x}\left[u_{\theta_i}\left(a\left(\theta;x\right),\theta_i\right)(1-\theta_i)|\theta_i,\theta_{-i}>\frac{1}{2}\right]\right]\\ \sum_i \frac{1}{4}\left[E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)|\theta_{-i}\leq\frac{1}{2},x>\frac{1}{2}\right]+E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)|\theta_{-i}>\frac{1}{2},x>\frac{1}{2}\right]\right]\\ \sum_i \frac{1}{4}\left[E_{\theta,x}\left[u_{\theta_i}\left(a\left(\theta;x\right),\theta_i\right)(1-\theta_i)|\theta_i>\frac{1}{2},\theta_{-i}<\frac{1}{2}\right]-E_{\theta,x}\left[u_{\theta_i}\left(a\left(\theta;x\right),\theta_i\right)\theta_i|\theta_i\leq\frac{1}{2},\theta_{-i}>\frac{1}{2}\right]\right]\end{array}\right)$$

Given $a^*(\theta,x)$ sets 1/2 off-diagonals what remains to be proven is that

$$a^*(\theta,x)\in\arg\max_{a()}\left(\begin{array}{c}\sum_i \frac{1}{4}\left[E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)|\theta_{-i}\leq\frac{1}{2},x<\frac{1}{2}\right]-E_{\theta,x}\left[u_{\theta_i}\left(a\left(\theta;x\right),\theta_i\right)\theta_i|\theta_i,\theta_{-i}<\frac{1}{2}\right]\right]\\ \sum_i \frac{1}{4}\left[E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)|\theta_{-i}>\frac{1}{2},x<\frac{1}{2}\right]+E_{\theta,x}\left[u_{\theta_i}\left(a\left(\theta;x\right),\theta_i\right)(1-\theta_i)|\theta_i,\theta_{-i}>\frac{1}{2}\right]\right]\\ s.t.\ IC\ for\ i=1,2\ given\ \theta_{-i}\in(0,1)\ and\ monotonicity\\ and\ a(\theta,x)=1/2\text{ off-diagonals}\end{array}\right.$$

Since any schedule which is incentive compatible over $[0,1]^2$ and has a constant off-diagonals must, in fact, be incentive compatible over $\left[0,\frac{1}{2}\right]^2$ and $\left[\frac{1}{2},1\right]^2$. The program above implies:

$$\begin{aligned} for\ \theta &\in (0,1/2)^2,\ a^*(\theta,x)\in\arg\max_{a()}\sum_i E\left[u_i(a,\theta_i)|\theta\in(0,1/2)^2\right]\\ s.t.\ IC\ for\ i &= 1,2\ given\ \theta_{-i}\in(0,1/2)\ and\ monotonicity\end{aligned}$$

and

$$\begin{aligned} for\ \theta &\in (1/2,1)^2,\ a^*(\theta,x)\in\arg\max_{a()}\sum_i E\left[u_i(a,\theta_i)|\theta\in(1/2,1)^2\right]\\ s.t.\ IC\ for\ i &= 1,2\ given\ \theta_{-i}\in(1/2,1)\ and\ monotonicity\end{aligned}$$

[20] One concern that might arise is that maximizing $V(a)$ see equation (1) is equivalent to maximizing the sum of agents's utilities only when a is IC. Nonetheless, the improvement we attain by setting 1/2 off-diagonals does carry meaning since we continue modifying the allocation to restore IC and we always use the same objective function.

as desired. ■

Lemma (1/2 off-diagonals): *Given any incentive compatible allocation $a(\theta;x)$ that satisfies Monotonicity, we can find an alternative incentive compatible allocation $\tilde{a}(\theta;x)$ which is weakly better and satisfies $\tilde{a}(\theta;x)=\frac{1}{2}$ for all x, when $\theta_i>\frac{1}{2}>\theta_{-i}$, and $\tilde{a}\left(\frac{1}{2},\theta_{-i};x\right)=\frac{1}{2}$ when $x>\frac{1}{2}$.*

Proof Lemma (1/2 off-diagonals): Without loss of generality we focus on the case in which the starting $a(\theta;x)$ is symmetric across players and around $\frac{1}{2}$ i.e.

$$\begin{aligned} a(\theta_i,\theta_{-i};x) &= a(\theta_{-i},\theta_i;x) \\ a(\theta_i,\theta_{-i};x) &= 1-a(1-\theta_i,1-\theta_{-i};x) \end{aligned}$$

We first consider the case where the starting $a(\theta;x)$ is continuous and show later (in Step 3) the result extends to non-continuous allocations. Furthermore, let us first point out that it is without loss to start with a schedule that is not constant off-diagonals. Indeed, if $a(\theta;x)=c\in\Re$ for all θ in $\left[0,\frac{1}{2}\right]\times\left[\frac{1}{2},1\right]$ and $\left[\frac{1}{2},1\right]\times\left[0,\frac{1}{2}\right]$, we could set $\frac{1}{2}$ off-diagonals (and $\frac{1}{2}$ fifty percent of time whenever a player announces $\frac{1}{2}$) without affecting incentives and such a change would generate a gain.

We now proceed to prove the result through a sequence of steps.

Step 1 (1/2 off-diagonals generates an improvement): *In this first step, we show that a strict improvement can be attained if we replace the off-diagonal values in the original allocation by* 1/2. Formally,

$$a_{1/2}(\theta,x)=\begin{cases}\frac{1}{2} & \text{if } \theta_i>\frac{1}{2}>\theta_{-i} \ \ or\ if\ \theta_i=\frac{1}{2}\ and\ x>\frac{1}{2} \\ a(\theta;x) & \text{otherwise}\end{cases}$$

We should note however that $a_{1/2}(\theta,x)$ may not be IC (we will address this in the next step).

We can write the objective functional as:

$$V(a)=\left(\begin{array}{c}
\underbrace{\sum_i\frac{1}{4}\left[E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)|\theta_{-i}\leq\frac{1}{2},x<\frac{1}{2}\right]-E_{\theta,x}\left[u_{\theta_i}\left(a(\theta;x),\theta_i\right)\theta_i|\theta_i,\theta_{-i}<\frac{1}{2}\right]\right]}_{A} \\
\underbrace{\sum_i\frac{1}{4}\left[E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)|\theta_{-i}>\frac{1}{2},x<\frac{1}{2}\right]+E_{\theta,x}\left[u_{\theta_i}\left(a(\theta;x),\theta_i\right)(1-\theta_i)|\theta_i,\theta_{-i}>\frac{1}{2}\right]\right]}_{B} \\
\underbrace{\sum_i\frac{1}{4}\left[E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)|\theta_{-i}\leq\frac{1}{2},x>\frac{1}{2}\right]+E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)|\theta_{-i}>\frac{1}{2},x>\frac{1}{2}\right]\right]}_{C} \\
\underbrace{\sum_i\frac{1}{4}\left[E_{\theta,x}\left[u_{\theta_i}\left(a(\theta;x),\theta_i\right)(1-\theta_i)|\theta_i>\frac{1}{2},\theta_{-i}<\frac{1}{2}\right]-E_{\theta,x}\left[u_{\theta_i}\left(a(\theta;x),\theta_i\right)\theta_i|\theta_i\leq\frac{1}{2},\theta_{-i}>\frac{1}{2}\right]\right]}_{D}
\end{array}\right)$$

For any non-decreasing allocation $a(\theta_i,\theta_{-i},x)$, consider replacing $a(\theta_i,\theta_{-i};x)$ over the region $\left[0,\frac{1}{2}\right)\times\left(\frac{1}{2},1\right]$ by

$$\widetilde{a}(\theta_i,\theta_{-i},x;\alpha)=(1-\alpha)\,a(\theta_i,\theta_{-i};x)+\alpha E_{\theta,x}\left[a(\theta_i,\theta_{-i};x)\,|\theta\in\left[0,\frac{1}{2}\right)\times\left(\frac{1}{2},1\right]\right]$$

where $\alpha\in[0,1]$.

We argue that this replacemnet has a positive effect on term D for α small.

Indeed, regarding its first term, note that

$$
\begin{aligned}
& \frac{\partial}{\partial \alpha}\left[E_{\theta,x}\left[u_{\theta_i}\left(\widetilde{a}\left(\theta_i,\theta_{-i},x;\alpha\right),\theta_i\right)\left(1-\theta_i\right)|\theta_i>\frac{1}{2}>\theta_{-i}\right]\right]|_{\alpha=0} \\
= \; & E_{\theta,x}\left[\left[u_{\theta_i a}\left(a\left(\theta_i,\theta_{-i};x\right),\theta_i\right)\left(1-\theta_i\right)\right]\left[E_{\theta,x}\left[a\left(\theta_i,\theta_{-i};x\right)|\theta\in\left[0,\frac{1}{2}\right)\times\left(\frac{1}{2},1\right]\right]-a\left(\theta_i,\theta_{-i};x\right)\right]|\theta_i>\frac{1}{2}>\theta_{-i}\right] \\
> \; & \left(\begin{array}{c} E_{\theta,x}\left[u_{\theta_i a}\left(a\left(\theta_i,\theta_{-i};x\right),\theta_i\right)\left(1-\theta_i\right)|\theta_i>\frac{1}{2}>\theta_{-i}\right]\cdot \\ E_{\theta,x}\left[\left[E_{\theta,x}a\left(\theta_i,\theta_{-i};x\right)|\theta\in\left[0,\frac{1}{2}\right)\times\left(\frac{1}{2},1\right]\right]-a\left(\theta_i,\theta_{-i};x\right)|\theta_i>\frac{1}{2}>\theta_{-i}\right]\end{array}\right) \\
= \; & 0
\end{aligned}
$$

where the strict inequality follows from the Monotone Hazard condition along with $\left[E_{\theta,x}\left[a\left(\theta_i,\theta_{-i};x\right)|\theta\in\left[0,\frac{1}{2}\right)\times\left(\frac{1}{2},1\right]\right]\right.$ being decreasing (and non-constant) in θ_i, A similar argument can be made for the other component of term D.

Since the initial $a(.)$ was arbitrary, one has that, over

$$
\left[0,\frac{1}{2}\right)\times\left(\frac{1}{2},1\right]\cup\left(\frac{1}{2},1\right]\times\left[0,\frac{1}{2}\right),
$$

the optimal schedule is constant. The best among the constant actions for this region is $\frac{1}{2}$. Moreover, by setting the constant to $\frac{1}{2}$ when one announces $\frac{1}{2}$, and $x>\frac{1}{2}$, the term C, which is defined by

$$
\left[E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)|\theta_{-i}\leq\frac{1}{2},x>\frac{1}{2}\right]+E_{\theta_{-i},x}\left[u\left(a\left(\frac{1}{2},\theta_{-i};x\right),\frac{1}{2}\right)|\theta_{-i}>\frac{1}{2},x>\frac{1}{2}\right]\right]
$$

is also maximized.

Therefore the modified allocation has (a) a constant action off-diagonals and (b) prescribes, at least half of the time, type $\frac{1}{2}'s$ most preferred actions. Hence,

$$
V\left(a_{1/2}\right)>V\left(a\right).
$$

■

Step 2 (Restoring IC): In this step we show that we can modify $a_{1/2}\left(\theta,x\right)$ in a way that restores IC, preserves 1/2 off-diagonals and does strictly better than the original allocation $a\left(\theta,x\right)$. Throughout, we specify the changes only for the top quadrant ($\left[\frac{1}{2},1\right]^2$) as the required changes over the bottom quadrant are similar.

We start by constructing an alternative schedule as follows.

For an integer N, consider the partition of $\left[\frac{1}{2},1\right]$ given by $\{A_i\}_i$, where $i\in\{1,...,N\}$; and, for $i\leq N-1$, $A_i=\left[\frac{1}{2}+\frac{i-1}{2N},\frac{1}{2}+\frac{i}{2N}\right)$ and $A_N=\left[\frac{1}{2}+\frac{N-1}{2N},1\right]$. Let $a_{ij}=E_{\theta,x}\left[a\left(\theta;x\right)|\theta\in A_i\times A_j\right]$ denote the expected action in each square $A_i\times A_j$ under the original allocation. Now consider, the schedule

$$
\overline{a}_N\left(\theta,x\right)=\begin{cases} \frac{1}{2} & \text{if } \theta_i>\frac{1}{2}>\theta_{-i} \;\; or\ if\ \theta_i=\frac{1}{2}\ and\ x>\frac{1}{2} \\ a_{ij} & \text{if } \theta\in A_i\times A_j\subset\left(\frac{1}{2},1\right]^2 \\ a_{1j} & \text{if } \theta_i=\frac{1}{2},\ \theta_j\in A_j\ and\ x<\frac{1}{2} \end{cases} \tag{A1}
$$

There exists a $\overline{N}$ and a $\gamma>0$ so that for all $N>\overline{N}$,

$$
V\left(\overline{a}_N\right)>V\left(a\right)+\gamma.
$$

This follows from noting that $\overline{a}_N(\theta,x)$ converges to $a(\theta;x)$ over $\left(\frac{1}{2},1\right]^2$ when N goes to the infinity, so that:

$$\lim_{N\to\infty}\overline{a}_N(\theta,x)=a_{1/2}(\theta,x),$$

Since

$$V\left(a_{1/2}\right)>V(a),$$

making use of the Dominated Convergence Theorem, it then follows that there exists a $\overline{N}$ and a $\gamma>0$ so that if $N>\overline{N}$

$$V(\overline{a}_N)>V(a)+\gamma$$

as stated.

Although better than $a(\theta;x)$ for N large, $\overline{a}_N(\theta,x)$ might not be IC.

To re-establish incentive compatibility, we need all the "cutoff types" $\frac{1}{2}+\frac{i}{2N}$, $i\in\{1,..,N-1\}$, to be indifferent between between reporting "left" or "right" together with expected monotonicity of the allocation. As a first step, we show that it is sufficient to modify the allocation by adding constants δ_i along the diagonal squares $A_i\times A_i$, $i\geq 1$, to satisfy the indifference conditions. As a second step, we show that such δ_i can be chosen to be positive and such that the resulting allocation fares strictly better than $a(\theta;x)$. Finally, the last step shows that expected monotonicity is indeed satisfied.

Step 2A:

In order for the IC constraints to be satisfied, the $\{\delta_i\}_i$ must be chosen to guarantee

$$\begin{aligned}&\frac{1}{2N}\sum_{j=1,j\neq i}^{N}u\left(a_{i,j},\frac{1}{2}+\frac{i}{2N}\right)+\frac{1}{2N}u\left(a_{ii}+\delta_i,\frac{1}{2}+\frac{i}{2N}\right)\\ =\;&\frac{1}{2N}\sum_{j=1,j\neq i+1}^{N}u\left(a_{i+1,j},\frac{1}{2}+\frac{i}{2N}\right)+\frac{1}{2N}u\left(a_{i+1i+1}+\delta_{i+1},\frac{1}{2}+\frac{i}{2N}\right),\end{aligned}\tag{A2}$$

so that the cut-off types are indifferent between reporting left and right.

The above condition implicitly defines δ_{i+1} as a function of δ_i. We next show that, for a properly chosen δ_1, one can find a sequence of $\{\delta_i\}_i$ with $\delta_i\geq 0$ for all i.

Step 2B: Let $p>0$ be such that, by letting δ_1 to be $O(N^p)$, $\frac{1}{2N}u\left(a_{11}+\delta_1,\frac{1}{2}+\frac{1}{2N}\right)$ is $O(1)$. There exists a sequence of non-negative numbers $\{\delta_i\}$ that (i) solve Equation (A2), and (ii) lead to a schedule that improves strictly upon the initial $a(\theta;x)$. Moreover, the δ_i can be chosen to be $O(N^p)$ for all i.

We establish this result in 3 Claims.

Claim 1: If δ_1 is strictly positive and $O(N^p)$, one can find, for all $i\geq 2$, a sequence of strictly positive$\{\delta_i\}_{i\geq 2}$, where each δ_i is also $O(N^p)$.

Proof: Note that (i) $a(\theta,x)$ is uniformly continuous (since $a(\theta,x)$ is a continuous function over a compact set), and (ii) $u(a(\theta,x),\theta_i)$ is uniformly continuous (by the same reasons). Hence, for any $\varepsilon>0$, there exists a N'so that, if $N>N'$,

$$\left|u\left(a_{i,j},\frac{1}{2}+\frac{i}{2N}\right)-u\left(a_{i+1,j},\frac{1}{2}+\frac{i}{2N}\right)\right|<2\varepsilon.$$

for *all* i,j.

Therefore,

$$\frac{1}{2N}\left|\sum_j\left(u\left(a_{i,j},\frac{1}{2}+\frac{i}{2N}\right)-u\left(a_{i+1,j},\frac{1}{2}+\frac{i}{2N}\right)\right)\right|\leq\frac{1}{2N}\sum_j\left|u\left(a_{i,j},\frac{1}{2}+\frac{i}{2N}\right)-u\left(a_{i+1,j},\frac{1}{2}+\frac{i}{2N}\right)\right|<\varepsilon$$

This implies that, for all i,j,

$$\frac{1}{2N}\sum_j\left|u\left(a_{i,j},\frac{1}{2}+\frac{i}{2N}\right)-u\left(a_{i+1,j},\frac{1}{2}+\frac{i}{2N}\right)\right|=O\left(\frac{1}{N}\right)$$

Now, consider 4 when $i=1$. It can be rewritten as

$$\begin{aligned}
&\frac{1}{2N}u\left(a_{22}+\delta_2,\frac{1}{2}+\frac{1}{2N}\right)\\
=\;&\frac{1}{2N}u\left(a_{11}+\delta_1,\frac{1}{2}+\frac{1}{2N}\right)+\frac{1}{2N}\sum_j\left[u\left(a_{1j},\frac{1}{2}+\frac{1}{2N}\right)-u\left(a_{2,j},\frac{1}{2}+\frac{i}{2N}\right)\right]\\
&+\frac{1}{2N}\left[u\left(a_{22},\frac{1}{2}+\frac{1}{2N}\right)-u\left(a_{11},\frac{1}{2}+\frac{1}{2N}\right)\right]\\
=\;&\frac{1}{2N}u\left(a_{11}+\delta_1,\frac{1}{2}+\frac{1}{2N}\right)+O\left(\frac{1}{N}\right)+O\left(\frac{1}{N^2}\right)
\end{aligned}$$

If δ_1 is $O(N^p)$, the right hand side is $O(1)$. Therefore, the left hand side must also be $O(1)$. This, in turn, calls for δ_2 being $O(N^p)$. As $\delta_1>0$, δ_2 can also be made positive. Proceeding inductively, the result follows.■

Denote by $\hat{N}$ the value such that, for $N>\hat{N}$, the sequence $\{\delta_i\}$ is such that all elements are $O(N^p)$ and positive.

Claim 2: There exists a strictly positive δ_1, which is $O(N^p)$ and so that the schedule defined by

$$\widetilde{a}_1(\theta,x)=\begin{cases}\overline{a}_N(\theta,x)+\delta_1 \text{ if } \theta\in A_1\times A_1\\ \overline{a}_N(\theta,x)\ \textit{otherwise}\end{cases}$$

satisfies

$$V(\widetilde{a}_1)\geq V(a).$$

Proof: This follows immediately from Step 1. In fact, for all $N>\max\left(\overline{N},\hat{N}\right)$,

$$V(\overline{a}_N)>V(a)+\gamma$$

for a strictly positive γ. By adding a strictly positive number over $A_1\times A_1$, one will decrease type $\frac{1}{2}'$s payoff. Since, from type $\frac{1}{2}'s$ perspective, the harm caused by such change will occur with probability $\frac{1}{2N}$, and

$$V(\overline{a}_N)>V(a)+\gamma$$

the positive δ_1 necessary to satisfy

$$V(\widetilde{a_1})\geq V(a).$$

can be made $O(N^p)$.■

Claim 3: One can find a schedule that satisfies the indifference condition (Equation (A2)) and fares strictly better than $a(\theta;x)$.

Proof: For some $N > \overline{\overline{N}} = \max\left(\overline{N}, \hat{N}\right)$, define a new schedule that is equal to $\overline{a}_N(\theta, x)$ except at the squares $A_i \times A_i$, where it is equal to $a_{ii} + \delta_i$ for $i \geq 1$, where the δ_i is given by the sequence defined in Claim 1, for the δ_1 in Claim 2. Denoting this schedule by $\widetilde{a}(., x)$ one has that

$$V(\widetilde{a}) > V(\widetilde{a}_1) \geq V(a).$$

This follows because (i) $V(.)$ is linear in $a(.)$ for $\theta \in \left[\frac{1}{2}, 1\right]^2$, and, finally, (ii) the adding of the $\delta_i' s$ for $i \geq 2$ does *not* affect the utility of type $\frac{1}{2}$.■

We have just shown that starting from an arbitrary continuous and non-decreasing schedule a, we can construct an alternative schedule $\widetilde{a}$ that has $\frac{1}{2}$ off-diagonals, satisfies Local Incentive Compatibility and fares better than the initial a. What is left to show is that $\widetilde{a}$ satisfies expected monotonicity. We know argue that this is in fact the case.

Step 2C (Monotonicity): For all i, there is $\widetilde{N}$ so that, for $N > \widetilde{N}$,

$$\sum_{j=1}^{N} \widetilde{a}_{i+1j} \geq \sum_{j=1}^{N} \widetilde{a}_{ij}$$

Proof: First note that the indifference condition Eq.(A2) can be read as

$$\begin{aligned} 0 &= \\ & \frac{1}{2N} \sum_{j=1, j\neq i}^{N} u\left(a_{i,j}, \frac{1}{2} + \frac{i}{2N}\right) + \frac{1}{2N} u\left(a_{ii} + \delta_i, \frac{1}{2} + \frac{i}{2N}\right) \qquad (4) \\ & -\frac{1}{2N} \sum_{j=1, j\neq i+1}^{N} u\left(a_{i+1,j}, \frac{1}{2} + \frac{i}{2N}\right) - \frac{1}{2N} u\left(a_{i+1i+1} + \delta_{i+1}, \frac{1}{2} + \frac{i}{2N}\right) \qquad (5) \end{aligned}$$

Doing a Taylor series expansion of $u\left(a_{i+1i+1} + \delta_{i+1}, \frac{1}{2} + \frac{i}{2N}\right)$ around $a_{ii} + \delta$, we have

$$\begin{aligned} u\left(a_{i+1i+1} + \delta_{i+1}, \frac{1}{2} + \frac{i}{2N}\right) &= u\left(a_{ii} + \delta_i, \frac{1}{2} + \frac{i}{2N}\right) + \\ & u_a\left(a_{ii} + \delta_i, \frac{1}{2} + \frac{i}{2N}\right)\left[a_{i+1i+1} + \delta_{i+1} - a_{ii} - \delta_i\right] + O\left(\left[a_{i+1i+1} + \delta_{i+1} - a_{ii} - \delta_i\right]^2\right) \end{aligned}$$

Also, doing a Taylor series expansion of $u\left(a_{i+1,j}, \frac{1}{2} + \frac{i}{2N}\right)$ around a_{ij}, we have

$$u\left(a_{i+1,j}, \frac{1}{2} + \frac{i}{2N}\right) = u\left(a_{i,j}, \frac{1}{2} + \frac{i}{2N}\right) + u_a\left(a_{i,j}, \frac{1}{2} + \frac{i}{2N}\right)\left[a_{i+1,j} - a_{i,j}\right] + O\left(\left[a_{i+1,j} - a_{i,j}\right]^2\right).$$

Hence,

$$\begin{aligned} & \sum_{j=1, j\neq i+1}^{N} u\left(a_{i+1,j}, \frac{1}{2} + \frac{i}{2N}\right) \\ = & \left[\sum_{j=1, j\neq i}^{N} u\left(a_{i,j}, \frac{1}{2} + \frac{i}{2N}\right)\right] + \left[u\left(a_{ii}, \frac{1}{2} + \frac{i}{2N}\right) - u\left(a_{i+1i+1}, \frac{1}{2} + \frac{i}{2N}\right)\right] \\ & + \sum_{j=1, j\neq i+1}^{N} \left[\left(u_a\left(a_{i,j}, \frac{1}{2} + \frac{i}{2N}\right)\left[a_{i+1,j} - a_{i,j}\right]\right) + O\left(\left[a_{i+1,j} - a_{i,j}\right]^2\right)\right]. \end{aligned}$$

Therefore, one can write the indifference condition as

$$\begin{aligned}0 \;=\;& \\ &-\frac{1}{2N}\left[u_a\left(a_{ii}+\delta_i,\frac{1}{2}+\frac{i}{2N}\right)[a_{i+1i+1}+\delta_{i+1}-a_{ii}-\delta_i]+O\left([a_{i+1i+1}+\delta_{i+1}-a_{ii}-\delta_i]^2\right)\right] \\ &-\frac{1}{2N}\left[\sum_{j=1,j\neq i}^{N}\left(u_a\left(a_{i,j},\frac{1}{2}+\frac{i}{2N}\right)(a_{i+1,j}-a_{i,j})+O\left([a_{i+1,j}-a_{i,j}]^2\right)\right)\right] \\ &-\frac{1}{2N}\left[u\left(a_{ii},\frac{1}{2}+\frac{i}{2N}\right)-u\left(a_{i+1i+1},\frac{1}{2}+\frac{i}{2N}\right)\right]\end{aligned}$$

or

$$\begin{aligned}0 \;=\;& \qquad\qquad \text{(Indiff)} \\ &\frac{1}{2N}\left[u_a\left(a_{ii}+\delta_i,\frac{1}{2}+\frac{i}{2N}\right)[(\delta_i-\delta_{i+1})+(a_{ii}-a_{i+1i+1})]+O\left([a_{i+1i+1}+\delta_{i+1}-a_{ii}-\delta_i]^2\right)\right] \\ &-\frac{1}{2N}\left[\sum_{j=1,j\neq i}^{N}\left(u_a\left(a_{i,j},\frac{1}{2}+\frac{i}{2N}\right)(a_{i+1,j}-a_{i,j})+O\left([a_{i+1,j}-a_{i,j}]^2\right)\right)\right] \\ &+\frac{1}{2N}\left[u_a\left(a_{ii},\frac{1}{2}+\frac{i}{2N}\right)(a_{i+1,i+1}-a_{ii})+O\left((a_{i+1}-a_i)^2\right)\right],\end{aligned}$$

where we have done a Taylor series expansion of $u\left(a_{i+1i+1},\frac{1}{2}+\frac{i}{2N}\right)$ around a_{ii}.

Now, assume, towards a contradiction, that expected monotonicity is violated, i.e., there is an i such that for all N :

$$\sum_{j=1}^{N}\widetilde{a}_{ij}>\sum_{j=1}^{N}\widetilde{a}_{i+1j}$$

This, in turn, implies that

$$\delta_i-\delta_{i+1}>\sum_{j=1}^{N}(a_{i+1j}-a_{ij})\geq 0.$$

Note that,

$$-\frac{1}{N}\left[\sum_{j=1,j\neq i}^{N}u_a\left(a_{i,j},\frac{1}{2}+\frac{i}{2N}\right)(a_{i+1,j}-a_{i,j})\right]<-\frac{1}{N}\min_j u_a\left(a_{i,j},\frac{1}{2}+\frac{i}{2N}\right)\sum_{j=1,j\neq i}^{N}(a_{i+1,j}-a_{i,j}).\qquad\text{(Ineq)}$$

Hence, applying (Ineq) to Equation 6, we get

$$
\begin{aligned}
0 &= \\
&\frac{1}{2N}\left[u_a\left(a_{ii}+\delta_i,\frac{1}{2}+\frac{i}{2N}\right)\left[(\delta_i-\delta_{i+1})+(a_{ii}-a_{i+1i+1})\right]+O\left(\left[a_{i+1i+1}+\delta_{i+1}-a_{ii}-\delta_i\right]^2\right)\right] \\
&-\frac{1}{2N}\left[\sum_{j=1,j\neq i}^{N}\left(u_a\left(a_{i,j},\frac{1}{2}+\frac{i}{2N}\right)(a_{i+1,j}-a_{i,j})+O\left(\left[a_{i+1,j}-a_{i,j}\right]^2\right)\right)\right] \\
&+\frac{1}{2N}\left[u_a\left(a_{ii},\frac{1}{2}+\frac{i}{2N}\right)(a_{i+1,i+1}-a_{ii})+O\left((a_{i+1}-a_i)^2\right)\right] \\
< &-\frac{1}{2N}\left[u_a\left(a_{ii}+\delta_i,\frac{1}{2}+\frac{i}{2N}\right)\left[(\delta_i-\delta_{i+1})+(a_{ii}-a_{i+1i+1})\right]+O\left(\left[a_{i+1i+1}+\delta_{i+1}-a_{ii}-\delta_i\right]^2\right)\right] \\
&-\frac{1}{2N}\min_j u_a\left(a_{i,j},\frac{1}{2}+\frac{i}{2N}\right)\sum_{j=1,j\neq i}^{N}(a_{i+1,j}-a_{i,j})-\frac{1}{N}\sum_{j=1,j\neq i}^{N}\left[O\left((a_{i+1}-a_i)^2\right)\right] \\
&+\frac{1}{2N}\left[u_a\left(a_{ii},\frac{1}{2}+\frac{i}{2N}\right)(a_{i+1,i+1}-a_{ii})+O\left((a_{i+1}-a_i)^2\right)\right]
\end{aligned}
$$

We now show that, for large N, the right hand side of this inequality is smaller than zero, which leads to the desired contradiction. Towards this note that for large N the following are true:

1. $$\frac{1}{2N}\left[u_a\left(a_{ii}+\delta_i,\frac{1}{2}+\frac{i}{2N}\right)\left[(\delta_i-\delta_{i+1})\right]\right]<0,$$
   since, for δ_i large, $u_a\left(a_{ii}+\delta_i,\frac{1}{2}+\frac{i}{2N}\right)<0$. Moreover, this term is of order $O(1)$ for both $\frac{1}{N}u_a\left(a_{ii}+\delta_i,\frac{1}{2}+\frac{i}{2N}\right)$, and $\delta_i-\delta_{i+1}$ – which is of the same order as $\sum_{j=1,j\neq i}^{N}(a_{i+1,j}-a_{i,j})$ – are $O(1)$.
2. $\frac{1}{N}\left[u_a\left(a_{ii}+\delta_i,\frac{1}{2}+\frac{i}{2N}\right)(a_{ii}-a_{i+1i+1})\right] = \frac{1}{N}\left[u_a\left(a_{ii}+\delta_i,\frac{1}{2}+\frac{i}{2N}\right)\left[(a_{ii}-a_{i+1i})+(a_{i+1i}-a_{i+1i+1})\right]\right]$ is $O\left(\frac{1}{N}\right)$.
3. $-\frac{1}{N}\min_j u_a\left(a_{i,j},\frac{1}{2}+\frac{i}{2N}\right)\sum_{j=1,j\neq i}^{N}(a_{i+1,j}-a_{i,j})$ is $O\left(\frac{1}{N}\right)$.
4. $\frac{1}{N}\left[u_a\left(a_{ii},\frac{1}{2}+\frac{i}{2N}\right)(a_{i+1,i+1}-a_{ii})\right]$ is $O\left(\frac{1}{N^2}\right)$
5. $\frac{1}{N}O\left(\left[a_{i+1i+1}+\delta_{i+1}-a_{ii}-\delta_i\right]^2\right)$ is $O\left(\frac{1}{N}\right)$, $\frac{1}{N}O\left((a_{i+1}-a_i)^2\right)$ is $O\left(\frac{1}{N^2}\right)$, and $\frac{1}{N}\sum_{j=1,j\neq i}^{N}\left[O\left((a_{i+1}-a_i)^2\right)\right]$ is $O\left(\frac{1}{N}\right)$.

Hence, there exists an N large such that the right hand side of the inequality is negative which implies, that for all i, there must be an $\widetilde{N}$ so that, for $N>\widetilde{N}$,

$$\sum_{j=1}^{N}\widetilde{a}_{i+1j}\geq\sum_{j=1}^{N}\widetilde{a}_{ij}.\blacksquare$$

All the results above establish the proof for the case in which the initial $a(.)$ is continuous.

In Step 3 below, we deal with the case in which $a(.)$ is non-decreasing but potentially discontinuous.

Step 3: For any $k \in \Re$, define the set:

$$A(k) = \left\{ a(\theta) : \begin{array}{c} a(\theta) \;\; non-decreasing \;\; and \\ -k \leq \left(\begin{array}{c} E_{\theta_{-i}} \left[u\left(a\left(\theta_i, \theta_{-i}\right), \theta_i\right)\right] - \\ E_{\theta_{-i}} \left[u\left(a\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right)\right] \end{array} \right) \forall \widehat{\theta} \times \theta_i \end{array} \right\}.$$

Similarly, define the set

$$C(k) = \{a(\theta) \in A(k) : a(.) \;\; is \;\; continuous\}.$$

Using Helly's selection theorem, it follows that we can show that for a given k :

i) $A(k)$ is compact valued.

ii) $A(k)$ is upper hemi-continuous.

Now define $g(k)$ as:

$$\begin{aligned} g(k) &= \max_{a(.)} E_\theta \left[\sum_{i=1}^{2} u\left(a\left(\theta_i, \theta_{-i}\right), \theta_i\right) \right] \\ & \quad s.t \;\; a(.) \in A(k). \end{aligned}$$

Given the properties of $A(k)$ stated above, Theorem 2 of Ausubel and Deneckre (1993) can be used to show that $g(\kappa)$ is continuous.

Next, define $h(k)$ as:

$$\begin{aligned} h(k) &= \sup_{a(.)} E_\theta \left[\sum_{i=1}^{2} u\left(a\left(\theta_i, \theta_{-i}\right), \theta_i\right) \right] \\ & \quad s.t \;\; a(.) \in C(k). \end{aligned}$$

Since $C(k) \subset A(k)$, $h(k) \leq g(k) \; \forall k$.

Also, for $k' > k$ since $A(k) \subset A(k')$ and $C(k) \subset C(k')$ it follows that $g(k) \leq g(k')$ *and* $h(k) \leq h(k')$

Lemma 4 *For any $\kappa_2 > \kappa_1$, if $a(.) \in A(\kappa_1)$, there exists a sequence of continuous functions $\{f_n(\theta)\}$ that converge pointwise to $a(\theta)$ and such that, for finite, but large n, $f_n(\theta)$ is in $C(\kappa_2)$.*

Proof. As $a(\theta) \in A(\kappa_1)$,

$$E_{\theta_{-i}}\left[u\left(a(\theta), \theta_i\right)\right] + \kappa_1 \geq E_{\theta_{-i}}\left[u\left(a\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right)\right] \text{ for all } \theta_i \times \widehat{\theta}.$$

so that

$$E_{\theta_{-i}}\left[u\left(a(\theta), \theta_i\right)\right] + \kappa_2 > E_{\theta_{-i}}\left[u\left(a\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right)\right] \text{ for all } \theta_i \times \widehat{\theta}.$$

Now, take the sequence of continuous non-decreasing functions $\{f_n\}_n$ such that

$$f_n(\theta) \rightarrow a(\theta) \text{ for all } \theta$$

(see Lemma 7 below).

Using the continuity of $u(., \theta_i)$, the above implies that

$$\begin{aligned} u\left(f_n(\theta), \theta_i\right) &\rightarrow u\left(a(\theta), \theta_i\right) \text{ for all } \left(\theta_i, \theta_{-i}\right) \\ u\left(f_n\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right) &\rightarrow u\left(a\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right) \text{ for all } \left(\widehat{\theta}, \theta_{-i}\right) \end{aligned}$$

Since f_n is continuous and defined over a compact set, it is bounded. Since $u(., \theta_i)$ is continuous, there is a $k < \infty$ such that

$$\left|u\left(f_n\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right)\right| < k \text{ for all } n.$$

Therefore, the Dominated Convergence Theorem implies:

$$E_{\theta_{-i}}\left[u\left(f_n\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right)\right] \rightarrow E_{\theta_{-i}}\left[u\left(a\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right)\right] \text{ for all } \widehat{\theta} \times \theta_i$$

It then follows from Proposition 23, in page 72, of Royden (1988) that, for any $\varepsilon > 0$ there exists an $N(\varepsilon)$ such that for all $n > N(\varepsilon)$

$$\left|E_{\theta_{-i}}\left[u\left(f_n\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right) - u\left(a\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right)\right]\right| < \frac{\varepsilon}{2} \text{ for almost all } \widehat{\theta} \times \theta_i$$

$$\left|E_{\theta_{-i}}\left[\begin{array}{c} \left[u\left(f_n\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right) - u\left(f_n\left(\theta_i, \theta_{-i}\right), \theta_i\right)\right] \\ \left[u\left(a\left(\widehat{\theta}, \theta_{-i}\right) - \theta_i\right) - u\left(a\left(\theta_i, \theta_{-i}\right), \theta_i\right)\right] \end{array}\right]\right| < \varepsilon \text{ for almost all } \widehat{\theta} \times \theta_i.$$

Hence, for $k_1 < k_2$, if

$$E_{\theta_{-i}}\left[u\left(a\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right) - u\left(a\left(\theta_i, \theta_{-i}\right), \theta_i\right)\right] \leq k_1 \text{ for all } \widehat{\theta} \times \theta_i$$

there exists an ε and an $N(\varepsilon)$ s.t.

$$E_{\theta_{-i}}\left[u\left(f_n\left(\widehat{\theta}, \theta_{-i}\right), \theta_i\right) - u\left(f_n\left(\theta_i, \theta_{-i}\right), \theta_i\right)\right] \leq k_2 \text{ for almost all } \widehat{\theta} \times \theta_i$$

Furthermore, if the IC constraints hold for almost all $\widehat{\theta} \times \theta_i$ then they must hold for all. Suppose, in search of a contradiction, that there were some θ_i which would benifit by ε more than k_2 from claiming to be $\widehat{\theta} \neq \theta_i$. Now, for any such ε, since all f_n is continuous and the agent's utility function is continuous, there must exist a ball of mass $q > 0$ of types around θ_i for which the IC constraint is strictly violated by more than $\frac{\varepsilon}{2}$. Of course, this contradicts the statement that for almost all $\widehat{\theta} \times \theta_i$ the allocation was IC. Hence, for n large (but finite), $f_n \in C(k_2)$, as desired. ■

Lemma 5 *For any $\kappa_2 > \kappa_1$, $h(\kappa_2) \geq g(\kappa_1)$.*

Proof. Fix a $\kappa_2 > \kappa_1$, and pick an arbitrary $\epsilon > 0$. Let $a(\theta)$ be such that $g(\kappa_1) = E_\theta[u(a(\theta), \theta_i)]$. Consider a sequence of continuous functions $\{f_n\}$ that converge pointwise to $a(.)$. By the Dominated Convergence Theorem, we can find N_1 large enough such that, if $n > N_1$,

$$E_\theta\left[\sum_{i=1}^{2} u\left(f_n(\theta), \theta_i\right)\right] > E_\theta\left[\sum_{i=1}^{2} u\left(a(\theta), \theta_i\right)\right] - \epsilon = g(\kappa_1) - \epsilon.$$

Moreover, using Lemma(4) above, we can find N_2 large enough for which,

$$f_n \in C(\kappa_2)$$

for all finite n that are larger than N_2. Hence, taking $\overline{N} = \max\{N_1, N_2\}$, for (finite) $n > \overline{N}$,

$$h(\kappa_2) \geq E_\theta\left[\sum_{i=1}^{2} u\left(f_n(\theta), \theta_i\right)\right] > E_\theta\left[\sum_{i=1}^{2} u\left(a(\theta), \theta_i\right)\right] - \epsilon = g(\kappa_1) - \epsilon,$$

so that

$$h(\kappa_2) > g(\kappa_1) - \epsilon.$$

As ϵ was arbitrary, the claim follows. ■

Next we show that $h(.)$ is continuous. Since $h(.)$ is increasing, if it were discontinuous, there would be a κ such that

$$\lim_{n\to\infty} h\left(\kappa - \frac{1}{n}\right) < \lim_{n\to\infty} h\left(\kappa + \frac{1}{n}\right).$$

Now, from Lemma (5) above, for all n,

$$h\left(\kappa - \frac{1}{n}\right) \geq g\left(\kappa - \frac{2}{n}\right).$$

Moreover,

$$g\left(\kappa + \frac{2}{n}\right) \geq h\left(\kappa + \frac{1}{n}\right).$$

Hence, one would have

$$\lim_n g\left(\kappa + \frac{2}{n}\right) \geq \lim_n h\left(\kappa + \frac{1}{n}\right) > \lim_n h\left(\kappa - \frac{1}{n}\right) \geq \lim_n g\left(\kappa - \frac{2}{n}\right).$$

which would contradict the continuity of $g(.)$. Hence, $h(.)$ must be continuous.

Now, we can finally establish that:

$$h(0) = g(0)$$

So far, we have shown that, for any $\kappa > 0$, we have:

$$g(\kappa) \geq h(\kappa) \geq g(0).$$

where the first inequality follows from the definitions of g and h, and the second inequality follows from Lemma (5).

Finally, taking the limits as $\kappa \to 0$, and using the continuity of $g(.)$ and $h(.)$,

$$g(0) \geq h(0) \geq g(0).$$

Hence, showing we can achieve a strict improvement over any continuous allocation is sufficient since by allowing the original set of allocations to be discountinuous does not allow us to do strictly better which is what $h(0) = g(0)$ implies. Note though that to be able to actually attain the optimal value we might need to use a discountinuous allocation.

We now show that, for any non-decreasing function $g:[0,1]^2 \to \Re$, we can find a sequence of continuous non-decreasing functions $\{g_m\}_m$, which converge pointwise to $g(.)$.

In order to do so, we first use the following result, which proof can be found in Rosenlicht (1968, pages 237 and 238).

Lemma 6 *For a given N, consider the partition of $[0,1]$ given by $\{A_i\}_i$, where $A_i = \left[\frac{1}{2}+\frac{i}{2N}, \frac{1}{2}+\frac{i+1}{2N}\right)$ whenever $i \in \{-N,...,N-2\}$, and $A_{N-1} = \left[\frac{1}{2}+\frac{N-1}{2N}, 1\right]$. Consider a function $g:[0,1]^2 \to \Re$, of the following form*

$$g(\theta) = c_{ij} \in \Re \text{ whenever } \theta \in A_i \times A_j,$$

with $c_{i+1j+1} \geq c_{i+1j} \geq c_{ij}$. That is, $g(.)$ is a non-decreasing simple function. One can then find a sequence of non-decreasing continuous functions that converge to $g(.)$ pointwise.

We can now prove

Lemma 7 *Let $g:[0,1]^2 \to \Re$ be a non-decreasing function. There exists a sequence of continuous non-decreasing functions that converge pointwise to $g(.)$*

Proof. For a given N, consider the partition of $[0,1]$ given by $\{A_i\}_i$, where $A_i = \left[\frac{1}{2}+\frac{i}{2N}, \frac{1}{2}+\frac{i+1}{2N}\right)$ whenever $i \in \{-N,...,N-2\}$, and $A_{N-1} = \left[\frac{1}{2}+\frac{N-1}{2N}, 1\right]$. Define g_N as follows:

$$g_N(\theta) = E_\theta\left[g(\theta)\,|\theta \in A_i \times A_j\right].$$

Clearly, for all θ,

$$|g_N(\theta) - g(\theta)| \to 0$$

as $N \to \infty$.

Now, fixing a N, one has that $g_N(.)$ is a non-decreasing simple function. Hence, by 6, one can find, for each N, a sequence of continuous non-decreasing functions $\{g_N^m(.)\}_m$ so that, for all θ,

$$|g_N^m(\theta) - g_N(\theta)| \to 0$$

as $m \to \infty$. Since,

$$|g_N^m(\theta) - g(\theta)| \leq |g_N^m(\theta) - g_N(\theta)| + |g_N(\theta) - g(\theta)|,$$

we have that, for all θ,

$$|g_N^m(\theta) - g(\theta)| \to 0$$

as $m, N \to \infty$. ■

Proof of Proposition 1. The proof follows exactly the same steps as the ones in Lemma 1/2 off-diagonals. Hence, we just sketch the main points here.[21] We consider the case in which $f(.)$ is symmetric around $\frac{1}{2}$. We argue later that this is without loss.

We start by noting that, if one replaces $\frac{1}{N}$ by $\int_{A_{-i}} f(\theta_{-i})\,d\theta_{-i}$ in the proof of Lemma 1/2 off-diagonals, all the steps go through for the general distribution case. Hence, it is without loss of optimality to assume that $a(\theta; x) = \frac{1}{2}$ for all x, when $\theta_i > \frac{1}{2} > \theta_{-i}$, and $a\left(\frac{1}{2}, \theta_{-i}; x\right) = \frac{1}{2}$ when $x > \frac{1}{2}$.

[21] A detailed proof can be obtained from the authors upon request.

Given that setting $\frac{1}{2}$ off-diagonals is optimal, if an allocation $a(\theta, x)$ does not satisfy the the condition in the Proposition, there are sets $B_1 \subset \left[0, \frac{1}{2}\right)^2$ and $B_2 \subset \left(\frac{1}{2}, 1\right]^2$ with strictly positive probability for which

$$\frac{\partial a(\theta, x)}{\partial \theta_i} > 0 \text{ for all } \theta \in B_1 \cup B_2.$$

An improvement can then be attained as follows. We focus on the region $\left(\frac{1}{2}, 1\right]^2$ – the argument for $\left[0, \frac{1}{2}\right)^2$ is symmetric.We consider, without loss, the case in which $a(.)$ is continuous over that region (see Claim 3 in Lemma 1/2 off-diagonals). For an integer N, consider the partition of $\left[\frac{1}{2}, 1\right]$ given by $\{A_i\}_i$, where $i \in \{1, ..., N\}$; and, for $i \leq N-1$, $A_i = \left[\frac{1}{2} + \frac{i-1}{2N}, \frac{1}{2} + \frac{i}{2N}\right)$ and $A_N = \left[\frac{1}{2} + \frac{N-1}{2N}, 1\right]$. Let $a_{ij} = E_{\theta,x}\left[a(\theta; x) | \theta \in A_i \times A_j\right]$ denote the expected action in each square $A_i \times A_j$ under the original allocation.

Now consider, the schedule

$$\overline{a}_N(\theta) = \begin{cases} \frac{1}{2} & \text{if } \theta_i > \frac{1}{2} > \theta_{-i} \;\; or \;\; if \; \theta_i = \frac{1}{2} \; and \; x > \frac{1}{2} \\ a_{ij} & \text{if } \theta \in A_i \times A_j \subset \left(\frac{1}{2}, 1\right]^2 \\ a_{1j} & \text{if } \theta_i = \frac{1}{2}, \; \theta_j \in A_j \; and \; x < \frac{1}{2} \end{cases}.$$

The allocation $\overline{a}_N(\theta)$ improves strictly upon $a(\theta, x)$ since it will take "averages" over the region B_2 (the arguments are exactly the same as those in Lemma 1/2 off-diagonals. It may not be Incentive Compatible, though. One can then proceed exactly as Steps 2 and 3 of Lemma 1/2 off-diagonals, and add non-negative numbers $\{\delta_i\}_{i \geq 1}$ over $A_i \times A_i$, $1 \geq 1$, in such a way that Incentive Compatibility is restored and the gains for the objective are preserved. The allocation resulting from this procedure will then have the property stated in the Proposition.

Now, if the density is not symmetric around $\frac{1}{2}$, one just need to apply, over $[0, 1]^2$, the same procedure as the one used over $\left[\frac{1}{2}, 1\right]^2$ for the case in which $f(.)$ is symmetric around $\frac{1}{2}$. ■