Robustly Ranking Mechanisms

Tilman Börgers and Doug Smith*

For a mechanism designer with an objective such as welfare we propose a method for robustly ranking mechanisms. The method is based on eliminating weakly dominated strategies only, and thus does not require any assumptions about agents' beliefs about each other except full support. We illustrate the usefulness of this method in two examples. In both examples we show that there are mechanisms that are ranked by our method above dominant strategy mechanisms. These examples question the literature's focus on dominant strategy mechanisms in cases when such mechanisms yield undesirable outcomes.

I Informal Examples

Bilateral Trade. A seller and a buyer negotiate over an indivisible item. The seller's value v_S and the buyer's value v_B are privately observed variables taken from $[0, 1]$. If a sale takes place at price p, then the seller's utility is p, and the buyer's utility is $v_B - p$. Otherwise the utilities are v_S and 0 respectively. The only dominant strategy mechanisms for this setting that satisfy ex post individual rationality and budget balance are mechanisms where the mechanism designer sets a, possibly random, price $\bar{p}$, and trade takes place if and only if both agents agree to trade at this price (Kathleen Hagerty and William P. Rogerson, 1987).

We show that a mechanism designer who seeks efficiency can do better by giving the

*University of Michigan and Federal Trade Commission, respectively. Thanks to Yan-Min Choo, Natalia Lazatti, Moritz Meyer-ter-Vehn, Xianwen Shi, and Takuro Yamashita for comments. The views expressed in this paper are the authors' and do not necessarily reflect those of the Federal Trade Commission.

seller the option to lower (but not raise) the price to a value $p_S \leq \bar{p}$. The seller and the buyer then accept or reject trade at p_S, and trade takes place if both accept.[1] If $v_S > \bar{p}$ the seller will reject trade. If $v_S < \bar{p}$, any $p_S < v_S$ is weakly dominated, but prices $p_S \in (v_S, \bar{p}]$ are not weakly dominated. Any such price may indeed be uniquely optimal, depending on the seller's beliefs about the buyer's action. For the buyer it is weakly dominant to accept if $p_S < v_B$, and to reject if $p_S > v_B$. Thus, no inefficient trade occurs, as in the dominant strategy mechanism, and more efficient trade may take place under the modified mechanism than under the dominant strategy mechanism, namely if $p_S < \bar{p}$ is chosen.[2]

Voting. N voters pick one of three candidates: x, y, or z. (Our argument cannot be extended to more than 3 alternatives.) Each voter has a privately observed utility for each candidate. The only dominant strategy mechanism is to randomly make one of the voters a dictator, and let the dictator choose (Aanund Hylland, 1980). Suppose the mechanism designer has a Rawlsian welfare function, and in particular prefers compromises to dictatorial solutions, that is, when some agents' ranking is $x \succ y \succ z$, and some other agents' ranking is $z \succ y \succ x$, then the mechanism designer favors y over x and z. Such a mechanism designer can do better than making one agent a dictator by giving every voter the opportunity to first approve of one or more alternatives. If there are alternatives of which all voters approve, then a randomly selected voter picks among these alternatives. Otherwise, the randomly selected voter acts as dictator.

A voter who avoids weakly dominated strategies always approves her top alternative and never approves her bottom alternative. Depending on the voter's von Neumann Mor-

[1] A mechanism that gives the buyer the ability to raise, but not lower, the price has analogous properties.

[2] Note that the modified mechanism is also ex post individually rational and budget balanced.

genstern utility function, however, weak dominance does not rule out approving the middle alternative. When some agents' ranking is $x \succ y \succ z$, and all other agents' ranking is $z \succ y \succ x$, both types of agents may approve y, and in this case the mechanism designer obtains a more preferred outcome than under random dictatorship. In addition, as we shall verify below for a particular Rawlsian welfare function, the alternative mechanism never yields lower expected welfare than random dictatorship.

II Formal Framework

There are N agents: $i \in N = \{1, 2, \ldots, N\}$, a set of alternatives A, and, for each agent i, a set of possible payoff types Θ_i. The utility function of agent i is $u_i : A \times \Theta_i \to \mathbb{R}$.[3] The mechanism designer has a welfare function $w : A \times \Theta \to \mathbb{R}$, where $\Theta = \times_{i=1}^{N} \Theta_i$. The mechanism designer chooses a game form: $G = (S_1, S_2, \ldots, S_N; g)$, where S_i is i's strategy set, and $g : S \to \Delta(A)$, denoting by S the set $\times_{i=1}^{N} S_i$.

To compare two game forms G and G' for a given type vector $\theta \in \Theta$, the mechanism designer eliminates in both mechanisms for every agent all strategies that are weakly dominated if the agent has type θ_i. Denote by S_{i,θ_i} the strategies left over for type θ_i of agent i in mechanism G after this elimination, and by $A_\theta(G)$ the range of g if g is restricted to the domain $\times_{i=1}^{N} S_{i,\theta_i}$. We define $A_\theta(G')$ analogously. These sets are thus the possible outcomes if no agent chooses a weakly dominated strategy. The key definition of this paper is:

Definition 1 *The mechanism designer* weakly prefers G to G' at $\theta \in \Theta$ *if the minimum of welfare in $A_\theta(G)$ is at least as large as the maximum of welfare in $A_\theta(G')$. She* strictly

[3]We only consider private values.

prefers G to G' at θ *if she weakly prefers* G *to* G' *at* θ*, but does not weakly prefer* G' *to* G *at* θ*. The mechanism designer* prefers G to G' *if she weakly prefers* G *to* G' *at all* $\theta \in \Theta$*, and strictly prefers* G *to* G' *at some* $\theta \in \Theta$*.*

III The Examples Revisited

Bilateral Trade. The outcomes are no trade, or trade, and a payment from the buyer to the seller: $A = \{0, 1\} \times \mathbb{R}$. The seller's and buyer's payoff types are their values, and the mechanism designer's welfare function is the sum of the agents' payoffs. The dominant strategy mechanisms are "fixed price mechanisms," parametrized by a price $\bar{p}$ which is independent of the agents' payoff types (Hagerty and Rogerson, 1987).[4] The price $\bar{p}$ can be random. Here we restrict attention to a deterministic $\bar{p} \in (0, 1)$. The agents can accept or reject this price: $S_S = S_B = \{\mathcal{Y}, \mathcal{N}\}$, and trade at price $\bar{p}$ occurs if and only if $s_S = s_B = \mathcal{Y}$. The alternative mechanism G', the "downward flexible price mechanism," allows the seller to reduce the price: $S'_S = [0, \bar{p}] \cup \{\mathcal{N}\}$, where a choice p_S from $[0, \bar{p}]$ indicates the seller's willingness to trade at price p_S. The buyer chooses the highest acceptable price $p_B \in S'_B = [0, 1]$. Trade, at price p_S, takes place if and only if the seller does not choose $\mathcal{N}$, and $p_B \geq p_S$.[5]

Proposition 1 *For any given* $\bar{p} \in (0, 1)$*, the mechanism designer prefers the* downward flexible price mechanism *to the* fixed price mechanism*.*

[4]For the definition of dominant strategy mechanisms in this context see Hagerty and Rogerson (1987). Hagerty and Rogerson also require mechanisms to be ex post individually rational and budget balanced. We ignore these conditions here. Our alternative trading mechanism satisfies these two conditions trivially.

[5]This is a static version of the mechanism described in Section I.

We sketch the proof. Obviously, agents' weakly dominant strategy in the fixed price mechanism is to accept trade if and only if $v_S \leq \bar{p}$ or $v_B \geq \bar{p}$ respectively. Here, and below, we assume that agents who are indifferent between trading and not trading have a lexicographically second order preference in favor of trade per se. We obtain $A_\theta(G) = \{1\}$ if $v_S \leq \bar{p} \leq v_B$ and $A_\theta(G) = \{0\}$ otherwise. (For simplicity, we omit the price dimension of the outcome space.) In the downward flexible price mechanism, the buyer has a weakly dominant strategy: $p_B = v_B$, and the seller has the weakly dominant strategy $\mathcal{N}$ if $v_S > \bar{p}$. If $v_S \leq \bar{p}$, strategies $\mathcal{N}$ and $p_S < v_S$ are weakly dominated. On the other hand, any $p_S \in [v_S, \bar{p}]$ is the unique best response if the buyer chooses p_B exactly equal to p_S, and therefore such a p_S is not weakly dominated. We obtain: $A_\theta(G') = \{1\}$ if $v_S \leq \bar{p} \leq v_B$; $A_\theta(G') = \{0, 1\}$ if $v_S \leq v_B < \bar{p}$; and $A_\theta(G') = \{0\}$ otherwise. Thus the mechanism designer weakly prefers G' to G for all θ, and strictly prefers G' to G when $v_S \leq v_B < \bar{p}$.

Voting. $A = \{x, y, z\}$, and voters' payoff types are von Neumann Morgenstern utility functions with domain A. We assume that voters are not indifferent between any two alternatives. We normalize utilities so that the top alternative has utility 0 and the bottom alternative has utility 1. The welfare function is Rawlsian. The mechanism designer does not use agents' von Neumann Morgenstern utilities for interpersonal comparisons. Instead, she sets for $a \in \{x, y, z\}$: $w(a, \theta) = 1$ if $u_i(a, \theta_i) = 1$ for all $i \in N$; $w(a, \theta) = 0.5$ if $u_i(a, \theta_i) > 0$ for all $i \in N$ and $u_i(a, \theta_i) < 1$ for at least one $i \in N$; and $w(a, \theta) = 0$ if $u_i(a, \theta_i) = 0$ for at least one $i \in N$. This w reflects aversion to ex post inequality.

The only dominant strategy mechanisms G are random dictatorships where the probability of each agent being dictator is determined ex ante (Hylland, 1980).[6] For simplicity,

[6]For the definition of dominant strategy mechanisms in this context see Hylland (1980).

we focus on the case that the probability of each agent being dictator is $1/N$. We shall show that there is a mechanism G', "random dictatorship with compromise," that the mechanism designer prefers to random dictatorship. In G' each voter i names a non-empty subset of the set of alternatives which is the set of alternatives that she approves of, and a strict order of $\{x, y, z\}$. One voter is randomly selected as the decider, where each voter has probability $1/N$ of being selected. If the sets of alternatives that agents approve of has a non-empty intersection, the alternative ranked highest by the decider in that intersection is chosen. Otherwise the alternative ranked highest by the decider among all alternatives is chosen.

Proposition 2 *The mechanism designer prefers* random dictatorship with compromise *to* random dictatorship.

We sketch the proof. It is obviously in both mechanisms weakly dominant to reveal the preference order truthfully. In G' It is weakly dominated for voter i to include i's bottom alternative in, or to exclude i's top alternative from, the set of approved alternatives. Including i's middle alternative is weakly dominated if its von Neumann Morgenstern utility is $1/N$ or less. This is because even a lottery that gives with probability $1/N$ the most preferred alternative, and with the remaining probability the bottom alternative, is not worse than the middle alternative. If voter i's utility of the middle alternative is more than $1/N$, it is not weakly dominated to include the middle alternative in the set of approved alternatives. This is because it may be that all agents other than i rank i's bottom alternative top, agree with i on the middle alternative, and include their top two alternatives in the approval sets. It is then uniquely optimal for i to include his middle alternative, too.

We prove first that for all $\theta \in \Theta$ the mechanism designer weakly prefers random dictatorship with compromise to random dictatorship. We distinguish three cases. The first is that θ is such that the maximum welfare from any alternative in $\{x, y, z\}$ is zero. This case is trivial, because all possible outcomes from either of the two mechanisms yield expected welfare zero. Next, suppose that the largest welfare from any alternative in $\{x, y, z\}$ is one. Then all agents have the same preferred alternative. This alternative is chosen under random dictatorship with probability 1. It is also chosen with probability 1 under random dictatorship with compromise, because it will be included in all agents' sets of approved alternatives, and therefore will be chosen by all agents from the intersection of these sets.

The remaining case is that θ is such that the maximum welfare from any alternative in $\{x, y, z\}$ is 0.5. To deal with this case it is sufficient to prove that, if the intersection of the agents' approval sets is non-empty, all alternatives in this set yield expected welfare 0.5. But this is trivial because the intersection of the set of approved alternatives cannot include any agent's bottom alternative. This concludes the proof that for all $\theta \in \Theta$ the mechanism designer weakly prefers random dictatorship with compromise to random dictatorship.

It remains to show that there are $\theta \in \Theta$ such that the mechanism designer strictly prefers random dictatorship with compromise to random dictatorship. Consider a θ where some agents rank alternatives $x \succ y \succ z$, and all other agents' ranking is $z \succ y \succ x$, and where for each agent the utility from y is more than $1/N$. Random dictatorship yields expected welfare 0, whereas random dictatorship with compromise, when agents choose not weakly dominated strategies, may result in y, which gives welfare 0.5. This outcome arises if all agents include their top alternative and y in their approval sets. The mechanism designer for such θ strictly prefers random dictatorship with compromise to random dictatorship.

IV Related Literature

Implementation in Not Weakly Dominated Strategies. Tilman Börgers (1991) considers a voting problem similar to our second example, and focuses, as we do, on the implementation of outcome correspondences in not weakly dominated strategies. He observes that the mechanism designer may prefer a game form that implements a correspondence to a dominant strategy mechanism, if the mechanism designer values compromises among the agents. Two differences with this paper are that he only models agents' ordinal preferences, not their cardinal preferences, and that he does not completely specify the mechanism designer's welfare function. He answers (negatively) a question that is of interest also in our setting: Do mechanisms exist that do not just expand the set of possible outcomes, beyond the ones allowed by dominant strategy mechanisms, but that rule out some undesirable outcomes of dominant strategy mechanisms? We do not yet know whether his result holds in our setting.

Takuro Yamashita (2011) is relevant to the question we just mentioned. Like we do, he studies implementation in not weakly dominated strategies for the bilateral trade problem. He focuses on a worst case scenario: the ex post welfare minimizing undominated strategies are played. This yields a complete ordering of mechanisms. A mechanism can perform better than a dominant strategy mechanism in Yamashita's ordering only if it admits for some type vectors only outcomes in weakly undominated strategies that are strictly better than the outcomes of a dominant strategy mechanisms. Thus the mechanisms that we present here do not perform better in Yamashita's ordering than dominant strategy mechanisms. Yamashita attributes to the designer a fixed belief about payoff types, and thus allows the designer to make tradeoffs between higher welfare for some payoff type vectors, and lower welfare for others. He finds beliefs for which optimal mechanisms are dominant strategy mechanisms and other beliefs for which they are not.

In Börgers and Doug Smith (2011) we investigate the voting problem and consider Bayesian equilibria of the voting game on all finite type spaces, assuming that weakly dominated strategies are not played. We postulate a mechanism designer who only relies on Pareto comparisons. We find that no game form is preferred, in a sense similar to that of Definition 1, to random dictatorship. However, if the mechanism designer considers agents' interim expected utilities rather than, as in Definition 1, ex post utilities, she prefers random dictatorship with compromise to random dictatorship. These results highlight the distinction between interim and ex post welfare.

Robust Full Implementation. Dirk Bergemann and Stephen Morris (2009, 2011) investigate social choice functions that can be implemented in strategies that survive iterated elimination of interim strictly dominated strategies for each type of each player.[7] Three differences between Bergemann and Morris's work and ours are: First, they focus on iterated elimination of strictly dominated strategies, whereas we consider a single round of elimination of weakly dominated strategies. Second, they focus on social choice functions, whereas we allow correspondences. Third, they describe what can be implemented. If what can be implemented is the planners' first best, this solves the planner's problem. However, Bergemann and Morris, unlike us, do not consider the case that the first best cannot be

[7]Their interest in this implementation concept is motivated by the fact that full implementation in interim Bayesian equilibria on all type spaces is equivalent to implementation in iterated elimination of interim strictly dominated strategies.

implemented, and therefore the planner has to to pick a second best mechanism.

Regarding the first point, we note that it is easy to see that our results would not be true if we used iterated elimination of strictly dominated strategies, but that they would continue to hold if we used the elimination of all weakly dominated strategies, followed by iterated elimination of interim strictly dominated strategies. This latter claim is true because the second step of this procedure, iterated interim elimination of strictly dominated strategies, does not eliminate any further strategies in either of our examples. To see this note that in each case we proved that the remaining strategies were not weakly dominated by showing that they were unique best responses to combinations of strategies of the other players that are not weakly dominated for those players. Therefore, none of these strategies is eliminated by strict dominance.

The previous paragraph's modified version of our procedure for determining possible outcomes from a mechanism is an asymmetric information version of a procedure studied in Börgers (1994; see also the references cited there) for complete information games. An adaptation of the argument in Börgers (1994) shows that, for finite mechanisms, the procedure can be justified by assuming that players have full support beliefs, and that this fact, the set of possible payoff types, as well as players' rationality, are approximate common knowledge. By contrast, Bergemann and Morris's procedure assumes that rationality is (exact) common knowledge. We can now interpret the first of the three above differences between our work and Bergemann and Morris's work, for finite mechanisms, as the difference between these two epistemic assumptions.

Robust Partial Implementation. Bergemann and Morris (2005) follow the literature on mechanism design in ignoring multiple equilibria. They give abstract examples where a mechanism designer can achieve first best in Bayesian equilibria on every type space, but not in dominant, or ex post incentive compatible, strategies. Multiple equilibria do not seem to play any role in these examples. The examples are thus similar to our examples, in that the mechanism designer prefers certain mechanisms over dominant strategy mechanisms in the sense of our Definition 2.

Kim-Sau Chung and Jeff Ely (2007) provide another formulation of the optimization problem of a revenue maximizing auctioneer who does not make assumptions about agents' beliefs about each other. They fix the designer's beliefs about payoff types, and then invoke the revelation principle, thus. Using a maxmin approach they obtain a complete ordering of mechanisms. They describe cases in which dominant strategy mechanisms are optimal (their Proposition 1), and give an example in which they are not (their Proposition 2).

Smith (2011) considers a public good problem with a balanced budget requirement. Smith's focus is on robust mechanism design, in the sense that the mechanism designer focuses on some particular equilibrium of each game form that he considers. The preference order that he attributes to the mechanism designer is, adapted to his context, the same as the preference order that we attribute to the mechanism designer. Smith exhibits a mechanism to which the mechanism designer prefers no other mechanism. In the context of our examples, the investigation of such mechanisms remains an open issue.

References

Bergemann, Dirk and Stephen Morris. 2005. "Robust Mechanism Design." *Econometrica*, 73(6): 1771-1813.

Bergemann, Dirk and Stephen Morris. 2009. "Robust Implementation in Direct Mechanisms." *Review of Economic Studies*, 76 (4): 1175-1204.

Bergemann, Dirk and Stephen Morris. 2011. "Robust Implementation in General Mechanisms." *Games and Economic Behavior*, 71 (2): 261-281.

Börgers, Tilman. 1991. "Undominated Strategies and Coordination in Normalform Games." *Social Choice and Welfare*, 8(1): 65-78.

Börgers, Tilman. 1994. "Weak Dominance and Approximate Common Knowledge." *Journal of Economic Theory*, 64(1): 265 -276.

Börgers, Tilman and Doug Smith. 2011. "Robust Mechanism Design and Dominant Strategy Voting Rules." Working paper, University of Michigan.

Chung, Kim-Sau and Jeff Ely. 2007. "Foundations of Dominant-Strategy Mechanisms." *Review of Economic Studies*, 74(2): 447-476.

Hagerty, Kathleen and William P. Rogerson. 1987. "Robust Trading Mechanisms." *Journal of Economic Theory*, 42(1): 94-107.

Hylland, Aanund. 1980. "Strategy Proofness of Voting Procedures with Lotteries as Outcomes and Infinite Sets of Strategies." Working paper, University of Oslo.

Smith, Doug. 2011. "A Prior Free Efficiency Comparison of Mechanisms for the Public Good Problem." Working paper, University of Michigan.

Yamashita, Takuro. 2011. "A Necessary Condition for Robust Implementation: Theory and Applications." Working paper, Toulouse School of Economics.