Journal of Mathematical Economics 74 (2018) 108–118

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

Symmetric mechanism design

Yaron Azrieli, Ritesh Jain *

Department of Economics, The Ohio State University, 1945 North High Street, Columbus, OH 43210, USA

ARTICLE INFO

Article history:
Received 27 June 2017
Received in revised form 10 November 2017
Accepted 13 November 2017

Keywords:
Mechanism design
Symmetry
Fairness
Revelation principle

ABSTRACT

Designers of economic mechanisms can often benefit by using discriminatory mechanisms which treat agents asymmetrically. This paper studies the extent to which a policy prohibiting biased mechanisms is effective in achieving fair outcomes. Our main result is a characterization of the class of social choice functions that can be implemented by symmetric mechanisms. When the solution concept used is Bayes–Nash equilibrium, symmetry is typically not very restrictive and discriminatory social choice functions can be implemented by symmetric mechanisms. Our characterization in this case is based on a 'revelation principle' type of result, where we show that a social choice function can be symmetrically implemented if and only if a particular kind of (indirect) symmetric mechanism implements it. When implementation in dominant strategies is considered, only symmetric social choice functions can be implemented by symmetric mechanisms. We illustrate our results in environments of voting with private values, voting with a common value, and assignment of indivisible goods.

1. Introduction

Designers of economic mechanisms can often benefit by biasing the rules in favor of some of the participating agents. In auction design, for example, it is well known (Myerson, 1981) that if bidders are heterogeneous then the seller can intensify the competition by subsidizing the bids of weaker bidders. Thus, this 'affirmative action' type of bias may increase the seller's revenue.[1] A second example is the design of reward schemes whose goal is to incentivize agents to exert costly effort, where even in a completely symmetric environment the principle can benefit by introducing discriminatory rewards. This type of results have been obtained by Winter (2004) in a joint production setup with increasing returns to scale technology, and by Pérez-Castrillo and Wettstein (2015) in the setup of contests. Third, the admission criteria employed by academic institutions in order to achieve a diverse student body are often criticized as discriminatory and unfair.[2]

On the other hand, fairness is a high priority goal for policy makers. This is evident from the many existing acts and regulations whose goal is to guarantee that markets are not biased against certain groups in the population.[3] In the context of mechanism design, a natural candidate for an anti-discriminatory regulation is that the "rules of the game" are the same for all participants. In other words, the mechanism should be symmetric across agents. Symmetry is a normatively appealing property that has been used extensively in the social choice and mechanism design literature. It appears for example in the classic result of May (1952) on simple majority rule, in the theory of social welfare functions (e.g. Mas-Colell et al., 1995, Chapter 22), and in the theory of values of cooperative games (e.g. Shapley, 1953), among many others.

The goal of this paper is to analyze the extent to which symmetric mechanisms guarantee fair outcomes. We consider an abstract mechanism design environment with incomplete information, allowing for both correlated types and interdependent values. The output of a mechanism specifies a public outcome and a private outcome for each agent. For instance, in a public good provision problem (as in e.g. Ledyard and Palfrey, 1999) a mechanism determines whether the good is provided or not (the public outcome) and the transfer required from each of the agents (the private outcomes) as a function of the profile of messages sent to the mechanism.[4] Roughly speaking, a symmetric mechanism is one in which a permutation of the vector of messages results in no change

* Corresponding author.
E-mail addresses: azrieli.2@osu.edu (Y. Azrieli), jain.224@osu.edu (R. Jain).

1 In a recent paper, Deb and Pai (2017) show how the seller can implement the revenue-maximizing auction without explicitly biasing the rules in favor of weaker bidders. We discuss their result and its relation to our work in detail below. Another relevant recent paper is Knyazev, (2016) who studies auction design when the goal of the seller is to maximize the surplus of his 'favorite' bidder.

2 These policies have been challenged in courts, a recent example is the case of Fisher versus University of Texas, US supreme court, vol. 570 (2013).

3 For example, the U.S. Equal Employment Opportunity Commission (EEOC) enforces several anti-discriminatory labor-market laws. Another example is the Genetic Information Nondiscrimination Act of 2008 in health insurance markets.

4 Environments with only public outcomes (such as voting environments) and environments with only private outcomes (such as auctions) can of course be accommodated in our framework.

https://doi.org/10.1016/j.jmateco.2017.11.007
0304-4068/

to the public outcome and the corresponding permutation of the private outcomes.

Our first main result (Theorem 1) characterizes the class of social choice functions (i.e., mappings from type profiles to outcomes) that a designer can implement in Bayes–Nash equilibrium using symmetric mechanisms. In many environments this class is large and contains functions that clearly favor certain agents over others. In some environments this includes even the extreme case of dictatorial functions, as we illustrate in Section 2 below. Thus, a regulation requiring mechanisms to treat agents symmetrically (in the sense we have defined it) is not necessarily an effective way to achieve fairness. On the other hand, in many environments symmetry is not a vacuous condition; some implementable social choice functions can no longer be implemented once symmetry is required.

The characterization in Theorem 1 can be roughly described as follows. A given (incentive compatible) social choice function f can be implemented by a symmetric mechanism if and only if for every pair of agents i and j there exists another social choice function f_{ij} such that (1) f_{ij} treats i and j symmetrically; and (2) for any type of agent i, the expected utility i receives under f by truthfully revealing his type is weakly higher than the expected utility he receives under f_{ij} from any possible type report. When such functions f_{ij} exist, we explicitly show how to construct a symmetric mechanism that implements f. The construction is based on the idea that the equilibrium message of an agent encodes his identity as well as his type, and the mechanism uses f to determine the outcome when the message profile contains all identities. The function f_{ij} is used to incentivize agent i to reveal his true identity instead of 'pretending' to be agent j.[5] Conversely, if for some pair of agents i and j an appropriate function f_{ij} does not exist, then, not only that this construction does not work, symmetric implementation is impossible altogether. Thus, this result has the flavor of a 'revelation principle', in the sense that one should only consider a particular kind of (indirect) mechanisms to determine whether symmetric implementation is possible or not.

We emphasize that we do not argue that the type of mechanism described above is a practical way for designers to get around a rule requiring symmetry. It is merely a theoretical construction that enables us to prove the sufficiency part of our theorem and to provide an upper bound on what can be symmetrically implemented. It may very well be that simpler and more practical symmetric mechanisms can be found that implement a given social choice function, as is the case in Deb and Pai (2017) for example. However, we do think that an important contribution of the paper is to expose the role that indirect mechanisms have in overcoming exogenous constraints such as symmetry.[6]

It is important to point out that Theorem 1 concerns partial implementation, and that the mechanism used in its proof would typically have multiple equilibria that may generate different outcomes than the one the designer intended. We thus follow the standard approach of the mechanism design literature in assuming that the designer can induce the agents to play a particular equilibrium by making it focal.[7] Notice that, if the underlying environment is symmetric across agents, then in order to implement a biased social choice function with a symmetric mechanism the equilibrium itself must generate the asymmetry, i.e., the designer should make an asymmetric equilibrium the focal one. A regulator interested in preventing discrimination should therefore be concerned not only with the symmetry of the rules of the mechanism, but also with the way the mechanism is framed and with the communication between the designer and the participants.

A well-known criticism of implementation in Bayes–Nash equilibrium is that it assumes too much common knowledge among the agents and the planner (see Bergemann and Morris, 2005). One solution to this problem is to consider more robust solution concepts. Here, we consider symmetric implementation in dominant strategies. Dominant strategy mechanisms are typically studied in environments with private values, so we restrict attention to this type of environments. In Theorem 2 we prove that if the environment is symmetric then only symmetric social choice functions can be implemented in (weakly) dominant strategies by symmetric mechanisms. In other words, indirect mechanisms are not useful in overcoming the symmetry constraint. This stands in stark contrast to the results for implementation in Bayes–Nash equilibrium.

The current work is motivated by a recent paper of Deb and Pai (2017), who study the possibility of a seller to design a discriminatory auction in an independent private-values setup under the same symmetry constraint as in the current paper. However, the set of feasible mechanisms they consider is further restricted to include only mechanisms in which bidders submit a single number (the bid), and the highest bidder wins the object. This rules out the type of mechanisms that we use in our proof. Nevertheless, they are able to show that essentially any incentive compatible direct mechanism (social choice function in our terminology) can be implemented in Bayes–Nash equilibrium by some such symmetric auction.[8] In other words, symmetry puts almost no restriction on what can be implemented. In Corollary 1 below we show that the same is true in voting environments with independent private-values, i.e. *every* incentive compatible social choice function can be symmetrically implemented in Bayes–Nash equilibrium. However, as we show by examples, this is no longer true with correlated types or interdependent values.

In the following section we illustrate our results with three examples. The first is in an environment of voting with private values, the second in a common-value voting setup (as in the 'Condorcet Jury' theorem), and the third is in a model of assignment of indivisible goods. Section 3 presents the notation and the definition of symmetry that we use, and Section 4 contains Theorem 1 and its proof. The special case in which the outcome space contains only public outcomes (voting environment) is further analyzed in Section 5. Section 6 contains the analysis of implementation in dominant strategies. Section 7 concludes.

2. Motivating examples

2.1. Voting with private values

Consider a society of two agents which needs to choose an alternative from the set $\{x, y\}$ (they may also choose to randomize between the alternatives). Each agent can either be of type X or of type Y. Both agents have the same (private-values) utility function

[5] This construction is reminiscent of mechanisms used in the literature on implementation with complete information (e.g. Maskin, 1999), where each agent reports everyone's preferences and an agent is punished if his message disagrees with the messages of all other agents. There are important differences however: First, and most importantly, in our case the designer does not use the punishment scheme to elicit any new information from the agents; it is only used to create symmetry in cases where the original social choice function is not symmetric. Second, under complete information the construction works for *every* social choice function, while this is not the case in our setup. Third, our construction is meaningful even if there are only two agents, which is not the case with complete information.

[6] Other examples that illustrate the potential importance of indirect mechanisms include (Bester and Strausz, 2000) for the case where the designer lacks commitment power, Saran (2011) when agents have menu-dependent preferences, and Strausz (2003) when only deterministic mechanisms are allowed.

[7] The equilibrium used in the proof of Theorem 1 is focal in the sense that each agent reports his true name and type.

[8] The notion of implementation used in Deb and Pai (2017) is somewhat weaker than ours, since they only require that the *expected* payment in the indirect mechanism is equal to the payment in the direct mechanism. Our definition requires that they are equal ex-post.