Studies in Choice and Welfare

For further volumes:
http://www.springer.com/series/6869

Bezalel Peleg • Hans Peters

Strategic Social Choice

Stable Representations of Constitutions

Springer

Professor Dr. Bezalel Peleg
The Hebrew University of Jerusalem
Institute of Mathematics and
Center for the Study of Rationality
Feldman Building, Givat-Ram
Jerusalem 91904
Israel
pelegba@math.huji.ac.il

Professor Dr. Hans Peters
Department of Quantitative Economics
University of Maastricht
P.O.Box 616
6200 MD Maastricht
The Netherlands
h.peters@maastrichtuniversity.nl

Studies in Choice and Welfare ISSN 1614-0311
ISBN 978-3-642-13874-4 e-ISBN 978-3-642-13875-1
DOI 10.1007/978-3-642-13875-1
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010934384

Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This book was written mainly during the Spring periods of 2008 and 2009, when the first author was visiting Maastricht University. Financial support both from the Dutch Science Foundation NWO (grants 040.11.013 and 0.40.11.082) and from the research institute METEOR (Maastricht University) is gratefully acknowledged.

Jerusalem *Bezalel Peleg*
Maastricht *Hans Peters*
April 2010

Contents

Preview to this book

A *constitution* describes the *rights* and *obligations* of individuals and groups in a society. Thus, it implies to which social states or sets of social states these groups are entitled. In order to enforce such rights and obligations, laws and rules are required that set bounds to the behavior of individuals and groups: they restrict the choices that individuals can make. A few natural desiderata for such a system of laws and rules come to mind. First, these laws and rules should leave the society members enough space to be able to enforce those social states to which they are constitutionally entitled, but not more than that. Second, such a system of laws and rules should make a situation possible in which society is stable, that is, in some state of equilibrium. Third, such equilibrium social states should be collectively optimal if possible: there should be no social state that is better for all members of society.

In this monograph, following Gärdenfors (1981), we model rights and constitutions by *effectivity functions*, a term coined by Moulin and Peleg (1982). An effectivity function assigns to each group in society a collection of sets of social states. If a group S is *effective* for a set of social states B, then this means that S is constitutionally entitled to the prevailing social state being in B. As described above, we additionally need a set of laws and rules in order that S be able to 'enforce' the social state to be in B. This is formalized by a *game form*. A game form endows each individual with a set of strategies. To each profile of individual strategies, an outcome function assigns a social state. Thus, a game form makes it possible to impose the constitution in practice. The idea of society functioning as a game is well established. Friedman (1962, p. 25) writes:

> It is important to distinguish the day-to-day activities of people from the general customary and legal framework within which these take place. The day-to-day activities are like the actions of the participants in a game when they are playing it; the framework, like the rules of the game they play. [...]

Naturally, the rules and laws (game form) should reflect the constitution. More precisely, by jointly choosing their individual strategies the members of a group (coalition) S can make sure that the resulting social state (alternative, outcome) is in some set B. These combinations (S, B) induced by the game form should be exactly the same as the combinations determined by the constitution. In that case, we say that the game form *represents* the effectivity function which models the constitution.

However, it is not sufficient to have a game form representing the constitution: we also want the game form to possess a certain form of stability. Also this point is stressed by Friedman (1962, p. 25):

> In both games and society also, no set of rules can prevail unless most participants most of the time conform to them without external sanctions; unless that is, there is a broad underlying social consensus. But we cannot rely on custom or on this consensus alone to interpret and to enforce the rules; we need an umpire. These then are the basic roles of government in a free society: to provide a means whereby we can modify the rules, to mediate differences among us on the meaning of the rules, and to enforce compliance with the rules on the part of those few who would otherwise not play the game.

Friedman uses the need for stability (social consensus) as an argument to have an 'umpire' in the form of the government. In our approach, governments do not appear explicitly, but may be represented by a group of individuals in the game form. The same holds for 'law and order': policemen and judges can be players or groups in the game form. This approach does not favor any political points of view: these should be implicit in the constitution or perhaps in the individual preferences of the players. For stability we rely on *game-theoretic equilibrium* concepts. Individuals in society are characterized, not only by their rights and obligations, but also by their preferences. The minimal requirement that we impose is that, whatever preferences the individuals hold, the resulting game (i.e., game form *cum* preferences) has a *Nash equilibrium*.

The issue of representation of a given power structure by a game form can be traced back to von Neumann and Morgenstern (1944), who showed that a superadditive coalitional game can be represented by a (transferable utility) strategic game. This result was extended to games with nontransferable utility, see Aumann (1967) and Borm and Tijs (1992). The most basic theorem in this monograph, Theorem 2.4.7, is in fact the generalization of this result to coalitional game forms[1], that is, effectivity functions. The rest of the monograph originates, conceptually, mainly from work from the mid-seventies by the first author (Peleg, 1978a and 1978b).

The emphasis in this work is on strategic stability of game forms representing effectivity functions, and not so much on considerations of equity or fairness. One could say that these should be embodied in the constitution modelled by the effectivity function and represented by the game form. Nevertheless, just as in the work of Hurwicz and Schmeidler (1978) we shall also pay attention to *Pareto optimality* of equilibrium outcomes of our representing game forms. In fact, it will appear that in the prevalent representing

[1] This term is due to Abdou and Keiding (1991).

game forms constructed in this monograph there are always Pareto optimal equilibrium outcomes.

Summarizing, we study representations of effectivity functions by game forms that satisfy, at least, the minimal stability requirement of having a Nash equilibrium for any profile of individual preferences. Although our leading motivation is to view effectivity functions as modelling constitutions, this is certainly not their exclusive usage. The theory presented here applies also to societies on a smaller scale. In principle, it applies to any society (including, for instance, academic societies) where member rights and obligations are exercised through a set of rules or procedures (e.g., by-laws).

Before we proceed to a more detailed description of the parts and chapters by which this monograph is organized, a few remarks pertaining to both its extent and limitations are in order.

First, from a broader perspective and as mentioned, an effectivity function can be viewed as a general form of a cooperative game, in the spirit of characteristic function games as introduced by von Neumann and Morgenstern (1944). The representation problem is then tantamount to finding a non-cooperative game (form) that endows each player and coalition with the same 'power' as the given effectivity function. The challenge is to find representations that perform well in terms of existence of Nash equilibria (or strong Nash equilibria), and Pareto optimality of the resulting equilibrium outcomes. If possible, the representing game form should be 'nice', e.g., in the sense of some continuity properties. Thus, although constitutions of large and small societies form a leading motivation for this work, the actual scope is larger.

Second, our emphasis is on representation and game-theoretic stability issues, and we do not have the ambition to contribute substantially to the purely legal or philosophical literature on laws and constitutions. Rather, this work can be seen as a specific contribution to the economic literature on *mechanism design* if we take the latter in a wide sense: the design of game forms (mechanisms) in order to reach collective decisions on the basis of individual choices.

Third, the theory in this book should be distinguished from what is usually called *implementation theory*. Implementation theory is concerned with finding a game form associated with a social choice correspondence (or function) such that, for any profile of preferences, the (Nash, strong,...) equilibrium outcomes of the game coincide exactly with the outcomes prescribed by the social choice correspondence. Thus, in implementation theory, one could say that the representation problem is restricted to equilibrium outcomes: there are no restrictions on the outcomes resulting from non-equilibrium behavior. In contrast, in the theory of this book, the representation issue is not restricted to equilibrium outcomes, and, consequently, representing game forms can be constructed completely independently of admissible preference profiles.

This monograph consists of two parts. Part I (Chapters 1–7) is closest to the above general description of the book and considers representations of effectivity functions by game forms, strategic stability properties of those game forms, Pareto optimality of equilibrium outcomes of those game forms, and continuity properties of game forms. Part II (Chapters 8–11) specializes to *social choice functions*. A social choice function assigns an alternative to any profile of preferences and is, thus, a game form where the strategies of the players are their individual preferences. Such a social choice function or, equivalently, such a game form induces an effectivity function which, naturally, is represented by it. Thus, compared to Part I, Part II of the monograph focuses on a special kind of ('direct revelation') game forms, namely social choice functions. The well-known Gibbard-Satterthwaite Theorem says that in such a game form there is always a player who can manipulate, i.e., fares better by not playing the strategy of reporting his true preference. As a consequence, playing the game may lead to undesirable outcomes, and in particular not to the outcomes intended by the original social choice function. For this reason, we shall focus on strong Nash equilibria that result in the same final outcome which would ensue if each player reported his true preference.

The book is based on work that has appeared over the last thirty years, including some recent articles, but also contains new results, such as Nash consistency of upper semicontinuous representations (Chapter 7), and a strongly consistent representation result for topological spaces (Chapter 5). There are quite some new or improved proofs of existing results as well, for instance the proof of the representation theorem in Chapter 2, and many of the proofs in Chapter 3.

Part I: Representations of constitutions

After the introductory Chapter 1 we set off in Chapter 2 with a formal description of a constitution and the effectivity function that it induces. We follow Gärdenfors (1981) and Peleg (1998): the latter reference presents a more detailed formalization of the concept of rights. The main result of Chapter 2 is Theorem 2.4.7, which shows that every effectivity function (under the usual necessary conditions of monotonicity and superadditivity) can be represented by a game form. The game form used in the proof of this theorem is central: it is used and modified throughout Part I of the monograph.

In Chapter 3 we derive necessary and sufficient conditions on an effectivity function to be representable by a *Nash consistent* game form, i.e., a game form that has a Nash equilibrium for every profile of preferences. This is Theorem 3.2.3. In Theorem 3.3.10 we show that for the case where the number of alternatives (social states) is finite, these conditions can be phrased directly in terms of the original effectivity function. The crucial condition here is an intersection condition that limits the power of individuals. Most of the remainder of this chapter is devoted to the case where the set of alternatives is infinite and we have some topological structure on this set and on the effectivity function. To obtain Nash consistency we need to add a topological condition, but the mentioned intersection condition stays intact.

In this chapter we also discuss the relation between our results and some well-known ‘paradoxes’ in the social choice literature, notably the Gibbard Paradox (Gibbard, 1974), and the inconsistency of Pareto Optimality and Minimal Liberalism (Sen, 1970), also called the ‘liberal paradox’. As to the latter, we show that the game form used to prove the existence of Nash consistent representations – the same game form as used for the main representation result in Chapter 2, see above – is, in fact, *weakly acceptable*. This means that the game formed by the game form and any profile of preferences has a Nash equilibrium with a Pareto optimal outcome. In this sense, our results offer a partial resolution to the liberal paradox.

Chapter 3 is based mainly on Peleg, Peters, and Storcken (2002). The part on the topological case also owes to Abdou (1988).

Chapter 4 goes deeper into the issue of Pareto optimality. Specifically, an *acceptable* game form is a Nash consistent game form such that all Nash equilibrium outcomes for all preference profiles are Pareto optimal. This is clearly desirable: whenever a Nash equilibrium is played, we do not have to worry about its Pareto optimality. Acceptability is attained under the demanding extra condition that no two disjoint coalitions can veto the same alternative x: it is not possible that both coalitions S and T are effective for the set of all alternatives except x if S and T are disjoint. Chapter 4 is based mainly on Peleg (2004).

In Chapter 5 we consider representing game forms that are *strongly consistent*, i.e., admit a strong Nash equilibrium for any profile of preferences. A strong Nash equilibrium is a strategy profile that is resistant not only to deviations by individuals but also to deviations of all other coalitions. Strong consistency is a natural strengthening of Nash consistency in view of the fact that we might also expect coalitions to deviate, but it is attained only at the price of strong conditions on the effectivity function, specifically *maximality* and *core stability*. Maximality means that the effectivity function is equal to its polar: for the case of finitely many alternatives, this implies that for each set of alternatives and each coalition of individuals, either that coalition is effective for the given set of alternatives or the complement of that coalition is effective for the complement of the given set of alternatives. Core stability means that for any given profile of preferences there should be an undominated alternative, where an alternative x is undominated if no coalition S is effective for a set B such that all members of S prefer all alternatives of B over x. Stability can be replaced by convexity, which is a condition imposed directly on the effectivity function (see Corollary 5.3.4). The results of this chapter are collected from various sources, among which is Peleg (1998).

In Chapter 6 we reexamine the intersection condition necessary for Nash consistency of representing game forms, as established in Chapter 3. Restricting ourselves to the case of finitely many alternatives, we manage to avoid this restrictive condition at the price of allowing some uncertainty in the outcomes. Specifically, we show that adding equal chance lotteries over pure alternatives and assuming that players evaluate these by utility functions respecting stochastic dominance – a minimal and natural requirement – enables us to obtain Nash consistent representations without any extra condition. We call such an effectivity function, obtained by adding equal chance lotteries, a *lottery model* if it preserves the original effectivity function in the following sense: a coalition is effective for a set of lotteries if and only if it was originally effective for the set consisting of the union of the supports of all those lotteries. Chapter 6 may serve as a starting point for finding representations under incomplete information about player types (preferences). It is based on Peleg and Peters (2009).

In the final chapter of Part I we go deeper into the topological properties of representing game forms for the case where the set of alternatives is a compact metric space. Specifically we investigate *continuity* of the outcome function. Our motivation for this is that continuity of the outcome function is a desirable property: it would be undesirable if a small change in an individual strategy would result in an entirely different social state. Unfortunately, we have to start with an impossibility result, entirely due to the lack of continuity of set intersection. On the other hand, we establish some weaker continuity properties, like for instance (upper or lower) semicontinuity when the set of alternatives is a compact subset of the real line. In the analysis the Cantor (ternary) set plays an important role, due to the mathematical fact that there exists a continuous surjective function from the Cantor set to any compact metric space.

The approach in Chapter 7 is necessarily more technical. Nevertheless the main message is that, although completely continuous representations may not exist, there is still much continuity possible while maintaining Nash consistency. This chapter is based mainly on Keiding and Peleg (2006b).

Part II: Consistent voting

Chapter 8 is introductory to Part II and recalls the Gibbard-Satterthwaite Theorem for social choice functions. As explained above, a social choice function is a special kind of game form, in which the possible preferences of the players are their strategies – so each player reports a strategy – and the outcome function (the social choice function) assigns to each profile of (reported) preferences an alternative. The Gibbard-Satterthwaite Theorem states that there is always a profile of (true) preferences and a player for whom it is not optimal to report his true preference, unless there is a dictator or there are only two alternatives. As a result, the final outcome may not be the desired outcome (desired according to the social choice function applied to the true preference profile). In Part II we accept this state of affairs as a matter of fact and consider the following alternative approach: for a given profile of (true) preferences, can we find a strong Nash equilibrium of the associated game such that the resulting outcome is identical to the outcome we would obtain if each player were to report truthfully? A social choice function with this property is called *exactly and strongly consistent* (ESC). Chapter 8, moreover, very briefly reviews other approaches to escape the consequences of the

Gibbard-Satterthwaite Theorem. One of these is the concept of 'equilibrium with threats' (Peleg and Procaccia, 2007).

Chapter 9 starts the investigation of ESC social choice functions using *feasible elimination procedures*, already treated in Peleg (1978a). The main results of this chapter imply that a(n anonymous) social choice function is ESC if and only if it always selects an alternative that can be obtained by a feasible elimination procedure. Such a feasible elimination procedure is based on a given set of positive integer weights assigned to alternatives. For a given preference profile, an alternative can be eliminated if it is the bottom alternative for a coalition of cardinality at least the weight of the alternative. This is reminiscent of the core of an associated effectivity function and, indeed, a relation between feasible elimination procedures and the core is established (Theorem 9.3.6).

In Chapter 10 the concept of a feasible elimination procedure is extended to apply to effectivity functions. An effectivity function is called *elimination stable* if the set of alternatives resulting from applying feasible elimination procedures is non-empty. The chapter contains characterizations of elimination stable effectivity functions in terms of conditions that can be checked independently of preference profiles, and is based mainly on Holzman (1986b).

In the final chapter (Chapter 11), based on Peleg and Peters (2006), we extend our results to the case where the number of alternatives is (still) finite but the set of individuals is a *continuum*: this is the prevalent framework of (for instance, national) elections. We start by extending the Gibbard-Satterthwaite Theorem to this model and establish, similar to Kirman and Sondermann (1972), that non-manipulability is only possible if there is a so-called 'invisible dictator'. Next, we extend the concept of feasible elimination procedures and find that for subsets of alternatives we have to distinguish between 'e-sets' and 'i-sets': the latter can only be blocked – and hence, eliminated – by coalitions of size strictly larger than the total weight of the 'i-set', whereas for 'e-sets' equality is sufficient for blocking. Using these tools, we are able to extend most of the results of Chapters 9 and 10 to this framework with a continuum of voters. In particular, we obtain an almost complete characterization of anonymous ESC social choice functions in this model.