

A REPRESENTATION THEOREM FOR VOTING WITH LOGICAL CONSEQUENCES*

PETER GÄRDENFORS
Lund University

This paper concerns voting with logical consequences, which means that anybody voting for an alternative x should vote for the logical consequences of x as well. Similarly, the social choice set is also supposed to be closed under logical consequences. The central result of the paper is that, given a set of fairly natural conditions, the only social choice functions that satisfy social logical closure are oligarchic (where a subset of the voters are decisive for the social choice). The set of conditions needed for the proof include a version of Independence of Irrelevant Alternatives that also plays a central role in Arrow's impossibility theorem.

1. INTRODUCTION

Arrow's (1951/63) famous impossibility result concerns the amalgamation of preference orderings. In the research on voting methods, the set of alternatives has been assumed to be a set of unstructured and independent alternatives. However, this is an unnatural assumption, since the alternatives we normally vote about are propositions and they therefore have logical relations to other propositions.

The tradition of social choice theory that Arrow's theorem belongs to can be contrasted with the tradition of deliberative democracy (Elster, ed.

* Much of this article was written while the author was a fellow at the Swedish Collegium for Advanced Study in the Social Sciences (SCASSS) in Uppsala. I want to thank the Collegium for providing me with excellent working conditions. Wlodek Rabinowicz and other fellows gave me valuable comments at a seminar at SCASSS when an early version of the paper was presented. I also want to thank Luc Bovens, Franz Dietrich, Christian List and an anonymous referee for their excellent comments on a later version. The final version was prepared during a stay at Oxford University for which I am grateful to the British Academy.

1998). In social choice theory, the democratic problem involves aggregation of views, interests, or preferences across individuals, not deliberation over their content. In deliberative democracy, the essence of democratic legitimacy is the capacity to deliberate in the production of a decision. Deliberation involves discussion in which individuals are amenable to scrutinizing and changing their preferences in light of persuasion from other participants.

One problem that appears in the deliberative tradition is the so-called “doctrinal paradox” (Kornhauser 1992). This paradox is illustrated by the following example that involves a court considering the following three propositions:

- x: there is a valid contract
- y: the client broke the contract
- z: the client is liable (logical consequence of x and y)

Assume that the jury contains three members i, j, and k, who vote for the propositions in the following way:

| | x | y | z |
|----------|-----|-----|-----|
| i | yes | no | no |
| j | no | yes | no |
| k | yes | yes | yes |
| Majority | yes | yes | no |

In this setting, the alternatives to be voted for are assumed to have a propositional structure and thereby have logical consequences (they form a Boolean algebra). A very natural condition on voting then becomes that if somebody votes for an alternative x, then she should vote for the logical consequences of x as well. In other words, the set of supported alternatives should be closed under logical consequences. All three jury members fulfil this assumption.

If one furthermore assumes that the social choice set should also be closed under logical consequences, one obtains a very interesting condition on voting functions. The “doctrinal paradox” is that simple majority voting does not satisfy this criterion, as is illustrated by the example above.

The question then becomes: Which voting functions satisfy the condition of social logical closure? The purpose of this paper is to show that, given a set of fairly natural conditions, the only social choice functions

that satisfy social logical closure are oligarchic (where a subset of the voters are decisive for the social choice). The set of conditions needed for the proof include a version of Independence of Irrelevant Alternatives that also plays a central role in Arrow's theorem.

2. THE FORMAL SETUP

It will be assumed that there is set X of *alternatives* (voting issues). The alternatives can also be called *propositions* since they stand in logical relations to each other.¹ Technically, it will be assumed that X forms a Boolean algebra closed under \neg , \vee and $\&$ (negation, disjunction and conjunction) and with top element T (truth) and bottom element F (falsity).² (Readers who are not familiar with Boolean algebras can think of the alternatives as a set of propositions closed under the operations of propositional logic. Under this reading, T can be taken to be any tautology and F to be any contradiction.) If $\neg x \vee y = T$, then y is said to be a *logical consequence* of x .

The set of *voters* will be denoted V and it consists of individuals i, j , etc.³ For each alternative x in X , each voter approves or disapproves or abstains. Thus, in contrast to the Arrow tradition, no preference rankings of the alternatives are assumed. The set of alternatives voter i approves of will be called *i's choice set* and it will be denoted C_i . I write $C_i x$ to denote that alternative x is in C_i , that is, that i approves of x .

Because a voter may abstain, neither $C_i x$ nor $C_i - x$ may hold. In earlier technical papers about the doctrinal paradox (e.g. List and Pettit 2002; List and Pettit 2004; Pauly and van Hees forthcoming; Dietrich forthcoming; van Hees 2004), it has been assumed that the choice sets are *complete* so that for every alternative x , either $C_i x$ or $C_i - x$ holds.

Each individual choice set is supposed to satisfy the following two conditions:

Individual logical closure (ILC): If $C_i x$ and y is a logical consequence of x , then $C_i y$; and if $C_i x$ and $C_i y$, then $C_i(x \& y)$.

Individual consistency (IC): Not both $C_i x$ and $C_i - x$ hold.

It follows from ILC that if x and y are logically equivalent, then $C_i x$ if and only if $C_i y$ for all voters i . Thus it is not necessary to distinguish an alternative from its logical equivalence class.

A vector $S = \langle C_i \rangle$, where i ranges over all voters in V , is called a *voting situation*. A *voting function* is a function from voting situations S to a *social*

¹ In the social choice tradition, alternatives are supposed to be mutually exclusive. Obviously, I am not making this assumption here.

² For an introduction to Boolean algebras, see e.g. Bell and Slomson (1969).

³ Technically, there is no restriction on the cardinality of V , but for all interesting voting functions, there will be at least two voters.

choice set C_S . Again, it is not assumed that C_S is *complete* so that for every alternative x , either $C_S x$ or $C_S -x$ holds. In the earlier technical papers mentioned above, C_S has been assumed to be *complete*. This is a strong and unnatural assumption that is crucial for the impossibility theorems that have been proved in these papers. In many voting situations, for example when everybody is indifferent between x and $-x$ or has no opinion about either x or $-x$, there is no reason to socially decide that either x or $-x$ should be chosen. I see it as an advantage that the results in this paper do not depend on the completeness of the social choice set.

3. CONDITIONS ON VOTING FUNCTIONS

I next turn to the set of conditions on voting functions that will be used for the representation theorem. The first two are familiar from the Arrow tradition:

Universal domain (UD): The voting function is defined for all possible voting situations satisfying ILC and IC.

Pareto optimality (PO): For all x in X , if $C_i x$ for all i , then $C_S x$.

The following two conditions are the social parallels of ILC and IC:

Social logical closure (SLC): For all x and y in X , if $C_S x$ and y follows logically from x , then $C_S y$; and if $C_S x$ and $C_S y$, then $C_S(x \& y)$.

Social consistency (SC): For all x in X , not both $C_S x$ and $C_S -x$ hold.

Imposing SLC and SC on voting functions, means that the logical structure of the alternatives will be important for the evaluation of a voting function. These conditions on voting functions are what really extend the framework that has previously been applied in the tradition following Arrow's theorem.

The next condition to be formulated is the correspondence of Arrow's "independence of irrelevant alternatives." There is one strong and one weak version:

Strong independence of irrelevant alternatives (SIIA): For all x in X , if situations S and S' are such that $C_i x$ if and only if $C'_i x$ for all i , then $C_S x$ if and only if $C_{S'} x$.

However, this condition is intuitively too strong since it requires that even if the voting situation changes so that more voters will approve $-x$, then the social choice with respect to x should not change as long as the same set of voters support x . The weak version accounts for this drawback:

Weak independence of irrelevant alternatives (WIIA): For all x in X , if situations S and S' are such that $C_i x$ if and only if $C'_i x$ and $C_i - x$ if and only if $C'_i - x$ for all i , then $C_S x$ if and only if $C_{S'} x$.

It should be noted that if all individual choice sets are complete, the distinction between strong and weak IIA disappears.

Define the majority voting function as the function where $C_S x$ if and only if $C_i x$ holds for a majority of the voters i in V . As the example above shows, the majority voting function does not satisfy SLC (but it satisfies all the remaining conditions). This is what is called the “doctrinal paradox.”

4. THE REPRESENTATION THEOREM

List and Pettit (2002, 2004), Pauly and van Hees (forthcoming), Dietrich (forthcoming) and van Hees (2004) have proved some impossibility results for voting with logical consequences. However, their results are strongly dependent on the assumption that social choice sets are complete. In contrast, the conditions that were introduced in the previous section are consistent, as will be shown soon. The question I focus on is *which* voting functions satisfy the criteria. In order to prepare for the central representation theorem of this paper, we need the following definition, where J is a subset of the set V of voters:

The central theorem to be proved below concerns which sets of voters have the power to decide the social choice set. Say that a set of voters D is *decisive* for a voting function, if and only if, for all x in X , in any voting situation S where $C_j x$ for all j in D , $C_S x$ holds. A voting function satisfying UD is called *oligarchic*, if and only if, there exists a non-empty smallest decisive set J ; the voting function is then called *oligarchic with respect to J*.

Special cases of oligarchic functions are *unanimous* ($J = V$) and *dictatorial* ($J = \{i\}$) functions. Note that according to the definition, C_S may contain more alternatives than those that are approved by all members of J in S . Let us say that an oligarchic function is *strict* if, for all situations, C_S consists of exactly the alternatives that are approved by all members of J , that is $C_S = \cap C_j$ for all j in J .

The following result establishes that the conditions introduced in the previous section are consistent:

Proposition: All strict oligarchic voting functions satisfy conditions PO, SLC, SC and SIIA.

Proof. More or less trivial. Conditions SLC and SC follow from the fact that the intersection of logically closed consistent sets is a logically closed consistent set.

For the representation theorem, one needs one more technical assumption:

Logically independent alternatives (LIA): For any alternative x that is not the top or bottom of the algebra (that is $x \neq T$ and $x \neq F$) there exists an

alternative y such that x is not a logical consequence of y , x is not a logical consequence of $\neg y$, $\neg x$ is not a logical consequence of y , and $\neg x$ is not a logical consequence of $\neg y$ (i.e., x is logically independent of y).

A Boolean algebra X is said to be *atomless* if and only if, for all x in X such that $x \neq F$, there is some alternative y , $F \neq y \neq x$, such that x is a logical consequence of y (i.e., $\neg y \vee x = T$). The definition of an atomless algebra says that for any alternative x (that is not a contradiction), there are more fundamental propositions that have x as a consequence.

Lemma: Condition LIA is equivalent to the condition that the Boolean algebra X of alternatives is atomless.

Sketch of proof: To show that LIA implies that X is atomless, for any alternative x , take an alternative z that is logically independent of x (such a z exists by LIA) and let y be $x \& z$. The difficult part is to show that an atomless algebra satisfies LIA. The idea is that for any x such that $T \neq x \neq F$, there exists a z such that $F \neq z \neq x$ and x is a logical consequence of z and a w such that $F \neq w \neq x$ and $\neg x$ is a logical consequence of w . It can then be shown that the alternative $y = (z \vee \neg x) \& (w \vee x)$ is logically independent of x .

One way of interpreting the condition that X is atomless is to say that all alternatives can be argued for. This is a kind of *deliberative* principle. Of course, there are many kinds of arguments for an alternative x that are not of the kind that they have x as a logical consequence. Nevertheless, the fact that the lemma shows that LIA is equivalent to this condition gives some further support for LIA from a deliberative point of view.

Atomless algebras have an infinite number of elements. Thus, condition LIA forces the set X of alternatives to be infinite. This may be seen as a strong assumption, but can be motivated as a technical idealization of the freedom a society has in creating new alternatives to be considered as arguments or counterarguments to the issues that are to be voted for. In any case, the condition mainly plays a technical role in the proofs to follow. An analogy is that in economic theory it is necessary to assume an infinite number of consumers with preferences over different sets of goods in order to derive the existence of a *unique* price vector.

The proof of the theorem below will use the mathematical notion of a *filter* (for a presentation of the role of this notion in model theory see Bell and Slomson 1969). A filter over a set V is a set $F(V)$ of subsets of V that satisfies the following conditions: (i) $V \in F(V)$, (ii) if $A \in F(V)$ and A is a subset of B , then $B \in F(V)$, and (iii) if $A \in F(V)$ and $B \in F(V)$, then $A \cap B \in F(V)$. If there is some subset J of V such that $F(V)$ consists of J and all supersets of J , then the filter is said to be *generated by* J . A filter that does not contain all subsets of V is said to be *proper*.

Theorem: If a voting function satisfies conditions UD, PO, SLC, WIIA and LIA, then the set of decisive sets forms a filter over the set V of voters.

Proof. The three conditions in the definition of a filter must be verified. (i) The set of all voters V is a decisive set by PO. (ii) It is trivial that if D is decisive and D is a subset of D' , then D' is decisive. (iii) The central part of the proof is to show that if A and B are decisive sets, then so is $A \cap B$. Assume that $C_j x$ for all j in $A \cap B$ in some situation S . To show that $C_S x$, change S to S' in the following way: Keep $C_i x = C'_i x$ and $C_i - x = C'_i - x$ for all i in V . For some y that is logically independent of x , change the choice sets so that $C'_i y$ for all i in $A - B$ and $C'_i - y$ for all i in $B - A$. Such a y exists by LIA. By UD, the voting function is defined for S' . Then $C'_i(y \vee x)$ for all i in A , by ILC, and hence $C_{S'}(y \vee x)$ since A is decisive. Similarly $C'_i(-y \vee x)$ for all i in B and hence $C_{S'}(-y \vee x)$ since B is decisive. It follows by SLC that $C_{S'} x$, because x is a logical consequence of $y \vee x$ and $-y \vee x$, and hence $C_S x$ by WIIA. Hence $A \cap B$ is decisive, which completes the proof.

If y is not logically independent of x , one cannot always change to $C'_i y$ and $C'_i - y$ for all i without effecting $C'_i x$ or $C'_i - x$ by ILC and IC and then WIIA cannot be applied. This is the reason why LIA is needed as an assumption. Note that only WIIA but not SIIA is needed for the proof. Furthermore, SC is not assumed, but if a voting function satisfies SC, then the empty set of voters cannot be decisive and the filter will be proper, that is, not contain all subsets of V .

Corollary 1: If a voting function satisfies conditions UD, PO, SLC, SC, WIIA and LIA and the set of voters is finite, then the function is oligarchic.

Proof. By the theorem the set of decisive sets form a proper filter. It is well known that all proper filters on finite sets are generated by some non-empty subset J of V (see Bell and Slomson 1969). Hence, the voting function is oligarchic with respect to J .⁴

As a matter of fact, there is nothing in the proof of the theorem that depends on the assumption that individual choice sets are not complete. This means that it is possible to prove an even stronger result that involves further restricted voting functions. Let us say that a voting function is *restricted to individually complete choice sets*, if it is defined only for voting situations where all the individual choice sets are complete. Condition UD in the following corollary should then be interpreted in relation to this restricted set of voting situations. (As noted above, WIIA becomes equivalent to SIIA in this setting.)

Corollary 2: If a voting function that is restricted to individually complete choice sets satisfies conditions UD, PO, SLC, SC, WIIA and LIA and the set of voters is finite, then the function is oligarchic.

⁴ If the set of voters is infinite, then the so-called Fréchet filter, consisting of all subsets of V with a finite complement, could also be used to define a voting function that satisfies the conditions in the theorem.

This result accentuates the crucial role of *social* completeness in earlier papers that prove impossibility results (List and Pettit 2002, 2004; Pauly and van Hees forthcoming; Dietrich forthcoming; van Hees 2004). In my opinion, the impossibility results are consequences of an unnaturally strong restriction on the outcomes of a voting function.

The following additional condition may be considered as a strong democratic desideratum, requiring that all voter be treated equally:

Anonymity (A): If the voting situation S' is obtained from S by a permutation of the individual choice sets, then, $C_{S'x} = C_{Sx}$.

Corollary 3: If a voting function satisfies conditions UD, PO, SLC, SC, WIIA, LIA and A, then only V is a decisive set.

In other words, if all voters are treated equally, then only unanimous choices will be decisive.⁵ A parallel result can be proved for voting functions that are restricted to individually complete choice sets. Unlike Arrow's theorem, the results in this section do not require preference rankings, only choice sets with no restrictions, except for that of logical consequence and consistency.

5. RELATED RESULTS

There exist some related results in the literature. I will here only make a brief comparison between the formal framework of this paper and those of the earlier works. List and Pettit (2002) prove an "impossibility result" concerning voting with logical consequences. Their proof is constructed around a condition they call Systematicity, which requires that for any propositions x and y in X , if every individual in V makes exactly the same judgment (acceptance/rejection) on x as she makes on y , then the collective judgment on x should also be the same as that on y , and the same pattern of dependence of collective judgments on individual ones should hold for all voting situations in the domain of the voting function (List and Pettit 2002: 99). This condition turns out to be very strong and it has been criticized by Chapman (2002).

In contrast to List and Pettit's result, I don't assume that choice sets are complete, i.e. either $C_i x$ or $C_i - x$. Furthermore, my condition WIIA is much weaker than their Systematicity. On the other hand, they do not need to assume LIA.

Similarly, Pauly and van Hees (forthcoming) assume a multi-valued logic that makes choices complete in the two-valued case and do not make contradictions necessarily have value F in the multi-valued case. Since they assume that the set of alternatives is "atomically closed" their framework is not directly comparable to the one used here. However, their theorem

⁵ Corollary 3 does not exclude that alternatives that are not unanimously chosen will also be in the social choice set.

basically states that if the voting functions are restricted to a finite set of voters and individual choice sets and a social choice set that are complete, then a voting function that satisfies conditions UD, PO, SLC, SC, WIIA (or SIIA) is *dictatorial* (i.e. there is a unique individual that forms a decisive set).⁶

On the surface, Pauly and van Hees' theorem seems to be stronger than the one proved in this paper. However, their result is again heavily dependent on social choices being complete. In my opinion, this assumption is unnatural and should be avoided in discussions around methods for voting with logical consequences.

6. IMPLICATIONS

What do the technical results concerning voting with logical consequences tell us about democracy and collective rationality? Following Dryzek and List (2003), I began this article by distinguishing between the social choice theory and deliberative democracy. In social choice theory, the democratic problem involves aggregation of preferences across individuals, while deliberation involves discussion in which individuals may change their preferences to conform to those of other participants.

The theorem proved in this paper relates to both perspectives. The formal setup obviously belongs to social choice theory. However, the requirements of Individual logical closure (ILC) and Social logical closure (SLC) (and, to some extent, Logically independent alternatives LIA), bring in elements of the deliberative tradition. The upshot of the theorem is that if we accept ILC, SLC, LIA, Weak independence of irrelevant alternatives, and some other more innocent conditions, then we will be constrained to oligarchic voting functions. And since an oligarchy will only be fully democratic in the limiting case when it consists of all members of the voting community, the theorem, via corollary 3, points to unanimous voting functions as the only acceptable ones.

Given the tension that exists between social choice theory and deliberative democracy, there are three main ways to react to this outcome (and other results from social choice theory).

1. At one extreme, it can be claimed that social choice theory with all its impossibility results shows that complete democracy can never be obtained. In line with this, Churchill says: "Democracy is the worst form of government except all those other forms that have been tried."
2. At the other extreme, one may neglect social choice theory and claim that all problems will be circumvented, if only we let the voters deliberate in a rational way. In other words, processes of deliberation should be encouraged with the aim of reducing the level

⁶ See also Corollary 2 of Theorem 2 of Dietrich (forthcoming).

- of conflict between different people's opinions and bringing about greater agreement (List 2001: section 2). In his account of deliberative democracy, Elster (1986: 112) presents this view quite succinctly: "The core of the theory [of deliberative democracy] . . . is that rather than aggregating or filtering preferences, the political system should be set up with a view to changing them by public debate and confrontation. The input to the social choice mechanism would then not be the raw, quite possibly selfish or irrational, preferences . . . but informed and other-regarding preferences. Or rather, there would not be any need for an aggregation mechanism, since a rational discussion would tend to produce unanimous preferences."
3. A middle road between these two extremes is taken by Dryzek and List (2003). Social choice theory shows the constraints that a deliberative procedure must fulfil in order to accomplish a rational and democratic outcome: "Deliberation facilitates pursuit of several escape-routes from the impossibility results invoked by social-choice theoretic critics of democracy. . . . Thus social choice theory shows exactly what deliberation must accomplish in order to render collective decision making tractable and meaningful, suggesting that democracy must in the end have a deliberative aspect." From this perspective, the theorem of this paper can be seen as providing yet another constraint from social choice theory.

REFERENCES

- Arrow, K. J. 1951/63. *Social choice and individual values*, 2nd edn. Wiley
- Bell, J. L. and Slomson, A. B. 1969. *Models and ultraproducts: an introduction*. North-Holland
- Chapman, B. 2002. Rational aggregation. *Politics, Philosophy and Economics* 1(3):337–54
- Dietrich, F. Forthcoming. Judgement aggregation: (im)possibility theorems. *Journal of Economic Theory*
- Dryzek, J. and List, C. 2003. Social choice theory and deliberative democracy: a reconciliation. *British Journal of Political Science* 33(1):1–28
- Elster, J. 1986. The market and the forum. In *Foundations of social choice theory*, ed. J. Elster and A. Hylland, 103–32. Cambridge University Press
- Elster, J., ed. 1998. *Deliberative democracy*. Cambridge University Press
- Kornhauser, L. A. 1992. Modelling collegial courts. II. Legal doctrine. *Journal of Law, Economics and Organization* 8:441–70
- List, C. 2001. Two concepts of agreement. *The Good Society* 11(1):72–9
- List, C. and Pettit, P. 2002. Aggregating sets of judgements: an impossibility result. *Economics and Philosophy* 18:89–110
- List, C. and Pettit, P. 2004. Aggregating sets of judgments: two impossibility results compared. *Synthese* 140(1–2):207–35
- Pauly, M. and van Hees, M. Forthcoming. Logical constraints on judgment aggregation. *Journal of Philosophical Logic*
- Van Hees, M. 2004. Limits of epistemic democracy. Department of Philosophy, University of Groningen, unpublished manuscript, 31 August 2004