AER: Insights 2020, 2(4): 459–474
https://doi.org/10.1257/aeri.20200178

Strategy-Proofness, Independence of Irrelevant Alternatives, and Majority Rule†

By Partha Dasgupta and Eric Maskin*

We show that strategy-proofness, the Pareto principle, anonymity, neutrality, independence of irrelevant alternatives, and decisiveness uniquely characterize majority rule on any domain of preferences for which there exists a voting rule satisfying these axioms. In our formulation, strategy-proofness includes manipulations by coalitions. However, we demonstrate that the characterization still holds when coalitions are restricted to arbitrarily small size. We also show that when coalitions can manipulate outside the domain, there is an extension of majority rule that satisfies these axioms on any domain without Condorcet cycles. (*JEL* D72)

Current election methods leave room for improvement. In the United States, Donald Trump won none of his first 17 victories in the 2016 Republican primaries by a majority: mainstream Republicans "canceled" each other out by splitting the anti-Trump vote. In France, far-right candidate Marine Le Pen made it to the runoff of the 2017 presidential election even though, almost certainly, she would have lost to François Fillon (eliminated in the first round) in a head-to-head contest. Better election methods would probably have prevented these anomalies.

There are many possible election methods, called *voting rules*, from which to choose. Here are some examples.

In *plurality rule* (used to elect members of Parliament in the United Kingdom and members of Congress in the United States), each citizen votes for a candidate, and the winner is the candidate with the most votes,[1] even if short of majority.[2]

In *runoff voting*, there are two rounds. First, each citizen votes for one candidate. If some candidate gets a majority, she wins. Otherwise, the top two vote-getters face each other in a runoff determining the winner.[3]

* Dasgupta: Faculty of Economics and Politics, University of Cambridge (email: partha.dasgupta@econ.cam.ac.uk); Maskin: Higher School of Economics and Department of Economics, Harvard University (email: emaskin@fas.harvard.edu). Larry Samuelson was coeditor for this article. This is a revised version of a manuscript first drafted in 2010. We thank Drew Fudenberg, Jerry Green, Sergiu Hart, Shengwu Li, and audiences at Stony Brook, Harvard-MIT, CNRS, TSE, Princeton, and Stanford for useful comments as well as Cecilia Johnson for research assistance. We thank the Rilin Fund for research support.

† Go to https://doi.org/10.1257/aeri.20200178 to visit the article page for additional materials and author disclosure statement(s).

[1] There can be a tie for the most votes, an issue that also arises for the other voting rules we mention and that is dealt with formally in our discussion of decisiveness in Section II.

[2] Plurality rule was the method the Republican Party adopted for many of its 2016 primaries.

[3] Runoff voting is used for presidential elections in France, Russia, Brazil, and many other countries.

Under *majority rule*—advocated by the Marquis de Condorcet (Condorcet 1785)—each voter ranks the candidates in order of preference. The winner is then the candidate who, according to the rankings, beats each opponent in a head-to-head contest.

In *rank-order voting* (the Borda count)—proposed by Condorcet's intellectual archrival Jean-Charles Borda (Borda 1781)—voters again rank the candidates. With n candidates, a candidate gets n points for every voter who ranks her first, $n - 1$ points for a second-place vote, and so on. The winner is the candidate with the most points.

Each voting rule so far is *ordinal* in the sense that the way a citizen votes can be deduced from his ordinal preferences over candidates (we define ordinality formally in Section II).[4] Next are two voting rules that are cardinal (i.e., dependent on more than just ordinal preferences).

In *approval voting*, each citizen approves as many candidates as he wants. The winner is the candidate with the most approvals. The voting rule fails ordinality because a citizen's preference ordering doesn't by itself determine the boundary between "approved" and "unapproved" candidates.[5]

In *range voting*, a citizen grades each candidate on, say, a ten-point scale ("1" denotes dreadful, and "10" denotes superb). A candidate's points are then summed over citizens, and the candidate with the biggest total score wins.[6]

Faced with many possibilities, how should society decide what voting rule to adopt? Ever since Arrow (1951), a standard answer is for society to consider what it wants in a voting rule, that is, to (i) posit a set of principles or axioms that any good voting rule should satisfy and (ii) determine the voting rule(s) with which they are consistent.

We use the axiomatic approach here (Section II gives precise definitions of our axioms, all familiar from the literature). Specifically, we suppose that there is a large number of voters and examine the *Pareto principle* (Pareto, for short): if all citizens prefer candidate x to y, then y should not be elected; *anonymity*: all citizens' votes should count equally; neutrality: all candidates should be treated equally; decisiveness: the election should result in a clear-cut winner; *independence of irrelevant alternatives* (IIA): if x is the winner among a set of candidates Y (the "ballot"), then x must still be the winner if the ballot is reduced from Y to Y' by dropping some losing (irrelevant) candidates;[7] and *ordinality*: the winner should depend only on citizens' *ordinal* rankings and not on preference intensities or other cardinal information.[8]

Of these six axioms, IIA is arguably the least "obvious." Still, it has strong appeal, in large part because it implies that a voting rule should not be vulnerable to vote splitting. Vote splitting arises when candidate x would beat y in a one-on-one contest

[4] More accurately, the way the citizen votes can be deduced if he is voting nonstrategically. We consider *strategic voting*—a major theme of this paper—later in this introduction.

[5] Behind approval voting is the idea that a minimum quality level—a cardinal concept—determines the boundary.

[6] Two variants of range voting are (i) *majority judgment* (Balinski and Laraki 2010), which is the same as range voting except that the winner has the biggest median (not total) score, and (ii) *budget voting*, in which a citizen has a set number of votes that he can allocate in any way to the different candidates. The winner is, again, the candidate with the biggest total.

[7] Arrow (1951) and Nash (1950) formulate (nonequivalent) axioms with the name "independence." In this paper and its predecessor (Dasgupta and Maskin 2008), we adopt the Nash formulation but could have used Arrow's version instead (indeed, an early paper in this line of work, Maskin 1995, does just that).

[8] Arrow (1951) notes that a citizen's ordinal preference between x and y can be ascertained from a simple experiment: give the citizen the choice between x and y. However, he argues that there is no reliable way of eliciting preference *intensities*. We support this view with Theorem 1 showing that cardinal voting rules can't be strategy-proof.

40%	35%	25%
Trump	Rubio	Kasich
Kasich	Kasich	Rubio
Rubio	Trump	Trump

FIGURE 1. VOTE SPLITTING

Notes: For the rankings above, Trump wins (with 40 percent) under plurality rule. However, both John Kasich and Marco Rubio would defeat Trump in a head-to-head contest (and, indeed, there is evidence from 2016 polls to back up this hypothetical). They lose in a three-way race because they split the anti-Trump vote. Hence, plurality rule violates IIA. Runoff voting does too: Rubio wins in a three-way race (first Kasich is dropped, and then Rubio defeats Trump in the runoff), but Kasich wins head to head against Rubio.

35%	33%	32%
x	y	z
y	z	x
z	x	y

FIGURE 2. CONDORCET CYCLES

Notes: Given the rankings above, candidate z can't be the winner under majority rule because a majority of voters (68 percent) prefer y. Moreover, y can't be the winner because a majority (67 percent) prefer x. But x can't win because a majority (65 percent) prefer z. The three rankings constitute a Condorcet cycle.

but loses to y when z runs too (because z splits off some of the vote that otherwise would go to x). See Figure 1 for an illustration that both plurality rule and runoff voting violate IIA because of vote splitting.

The Arrow impossibility theorem establishes that there is *no* voting rule satisfying all six axioms with at least three candidates and unrestricted voter preferences (see Theorem A in Section II).[9] In particular, majority rule doesn't even produce a winner for all rankings that voters might have, as Condorcet himself showed in a famous example of a "Condorcet cycle" (see Figure 2).

Thus, in Dasgupta and Maskin (2008), we argue that the natural follow-up question to Arrow is, which voting rule satisfies these axioms for the widest class of *restricted domains* of preferences? That paper shows that there is a sharp answer to this question: majority rule. Specifically, Theorem B states that majority rule satisfies the six axioms when preferences are drawn from a given domain if and only if that domain does not contain a Condorcet cycle. More strikingly, if some voting rule satisfies the six axioms on a given domain, then majority rule must also satisfy the axioms on that same domain. And, unless the original voting rule is itself majority rule, there exists another domain on which majority rule satisfies the axioms and the original voting rule does not (Theorem C).

[9] Arrow (1951) considers social welfare functions (mappings from profiles to social rankings) rather than voting rules. However, making the translation from one kind of mapping to the other is straightforward. Arrow (1951) uses a weak form of anonymity called nondictatorship and doesn't require decisiveness or neutrality; thus his version of the impossibility theorem is stronger than Theorem A.

In this paper, we consider an additional, often-invoked axiom: *strategy-proofness*—a voting rule should induce citizens to vote according to their true preferences, not strategically. There are at least two justifications for strategy-proofness. First, if citizens do vote strategically, then the voting rule in question doesn't produce the outcomes intended; since the rule's inputs are distorted, so are the outputs. Second, strategic voting imposes a burden on citizens. It is hard enough for a conscientious citizen to determine his own preferences: he has to study the candidates' characters, backgrounds, and positions. If, additionally, he must know *other* citizens' preferences and strategies in order to react strategically to them, his decision problem becomes much more difficult. For example, consider Figure 1. In a plurality rule election, Kasich supporters can stop Trump from winning by voting for Rubio, but this requires them to know this and coordinate on that manipulation.

Our first new result (Theorem 1 in Section III) establishes that any voting rule satisfying strategy-proofness and decisiveness on a given domain must be ordinal. The proof is straightforward; indeed, the argument ruling out range voting is especially simple: suppose there are two candidates x and y running, and a particular citizen judges them both to be quite good. If he were grading honestly, he would give x a grade of 8 and y a grade of 7. But, in an election, he has the incentive to give x a grade of 10 and y a grade of 1 to maximize x's chance of winning, a violation of strategy-proofness.

Just as we ran into the Arrow impossibility theorem in our previous work, we collide with the Gibbard-Satterthwaite impossibility theorem (Gibbard 1973, Satterthwaite 1975) once we impose strategy-proofness. A fortiori (in view of Theorem A), there exists no voting rule that satisfies all seven axioms when voters' preferences are unrestricted. Indeed, there is no voting rule that even satisfies all of anonymity, neutrality, decisiveness, and strategy-proofness (Theorem E). Hence, we turn our attention again to restricted domains and show that majority rule satisfies the seven axioms on any restricted domain without a Condorcet cycle (Theorem 2 in Section IV).

Implicit in Theorem 2 is the assumption that voters are confined to the restricted domain in question when they misrepresent their preferences. As we argue in our discussion of public goods in Section IV, this assumption makes sense in some circumstances but not all. Yet, when voters can misrepresent *freely*, a majority (Condorcet) winner may not exist. Thus, we must extend majority rule so that it always produces a well-defined outcome. Specifically, we use the *Smith set* (Smith 1973, Fishburn 1977), the (unique) minimal subset of candidates that beat any other candidate by a majority.[10] When a majority winner does not exist, we choose a random member of the Smith set as the outcome. Theorem 3 establishes that this extension of majority rule satisfies the seven axioms[11] on any domain without Condorcet cycles.

Theorem C from Dasgupta-Maskin (2008) shows that majority rule dominates other voting rules in the sense of satisfying Pareto, anonymity, neutrality,

[10] We could alternatively work with a refinement of the Smith set, e.g., the uncovered set (see Miller 1977), but this wouldn't change our analysis or results significantly.

[11] Strategy-proofness needs to be modified slightly to accommodate misrepresentations outside the domain in question.

decisiveness, IIA, and ordinality more often. When we add strategy-proofness to the mix, majority rule is, in fact, *uniquely* characterized. Theorem 4 establishes that a voting rule satisfying Pareto, anonymity, neutrality, decisiveness, IIA, and strategy-proofness (ordinality is redundant) on some domain can *only* be majority rule, a result that generalizes May (1952).

Our definition of strategy-proofness considers misrepresentations by coalitions. Intuitively, coordinating misrepresentations in a large coalition is harder than in a small one. Hence, a result that relies on large coalitions may not be entirely convincing. Accordingly, Theorem 5 strengthens Theorem 4 by showing it holds when coalitions are restricted to be of arbitrarily small size.

I. Model

There is a finite set X of potential candidates for a given office.[12] The electorate is a continuum of voters, taken to be the unit interval $[0,1]$ (the continuum makes the probability that there is a tie for the winner negligible, an issue discussed in Section II).

Each voter $i \in [0,1]$ is described by his *utility function* $u_i : X \to \mathbb{R}$. To simplify analysis, we rule indifference out by *assumption*. That is, for all $x, y \in X$, if $x \neq y$, then $u_i(x) \neq u_i(y)$. Let U_X consist of all utility functions on X without indifference. A *profile* $u_{(\cdot)}$ on U $(\subseteq U_X)$ is a specification of a utility function $u_i \in U$ for each voter $i \in [0,1]$. A ballot is a subset Y $(\subseteq X)$ consisting of the candidates who are actually running for the office. Let ΔY consist of the probability distributions over Y.

A *voting rule* is a function that, for each profile $u_{(\cdot)}$ and each ballot Y, selects a winner $F(u_{(\cdot)}, Y) \in \Delta Y$, where ΔY is the set of randomizations over Y. This formulation allows for election methods in which the winner is determined partly by chance.

To facilitate our analysis, we focus henceforth on voting rules that are finitely based in the sense that a voter's set of possible utility functions can be partitioned into a finite number of equivalence classes.[13] Formally, F is *finitely based* provided there exist a finite set S (the base set) and, for each voter i, a mapping $h_i : U_X \to S$ such that, for all profiles $u_{(\cdot)}$ and $u'_{(\cdot)}$, if $h_i(u_i) = h_i(u'_i)$ for all $i \in [0,1]$, then for all $Y \subseteq X, F(u_{(\cdot)}, Y) = F(u'_{(\cdot)}, Y)$. All the voting rules discussed in the introduction are finitely based (e.g., for an ordinal voting rule, S is just the set of rankings; for range voting, S is the set of possible mappings from candidates to grades between 1 and 10).

With a continuum of voters, we can't literally count the *number* of voters with a particular preference; we must instead consider *proportions* of voters. For that purpose, we can use Lebesgue measure μ on [0,1]. Thus, for profile $u_{(\cdot)}$, $\mu(\{i \mid u_i(x) > u_i(y)\})$ is the proportion of voters who prefer candidate x to candidate y.[14]

[12] A "potential" candidate is one who could conceivably run for the office in question but, in the end, might not.

[13] This focus makes it easier to define the concept of a generic profile: see the discussion of decisiveness in Section II.

[14] To be accurate, we must confine attention to profiles for which the set $\{i \mid u_i(x) > u_i(y)\}$ is *measurable* with respect to μ.

We can now formally define the voting rules mentioned in the introduction. We suppose that if there is tie, it is broken randomly. That is, if W is the set of candidates who tie for first, the outcome is $q(W)$, a random selection (with equal probabilities) from W. Here is the definition of plurality rule.

DEFINITION (Plurality Rule (First Past the Post)):

$$F^P(u_{(\cdot)},Y)$$
$$= q\Big(\Big\{x \in Y|\, \mu\big(\{i\,|\,u_i(x) > u_i(y)\; \textit{for all}\; y \neq x,\, y \in Y\}\big)$$
$$\geq \mu\big(\{i\,|\,u_i(x') > u_i(y)\; \textit{for all}\; y \neq x',\, y \in Y\}\big)\; \textit{for all}\; x' \in Y\Big\}\Big).$$

In words, candidate x wins if a higher proportion of voters rank x first than they do any other candidate x'. If there are multiple such x, one is selected at random. Similarly, here is the formal definition of majority rule.

DEFINITION (Majority Rule (Condorcet's Method)):

$$F^C(u_{(\cdot)},Y) = q\Big(\Big\{x \in Y \Big|\, \mu\big(\{i\,|\,u_i(x) > u_i(y)\}\big) \geq \frac{1}{2}, \textit{for all}\; y \neq x,\, y \in Y\Big\}\Big).$$

That is, candidate x is a Condorcet winner if, for any other candidate y, a majority prefer x to y; ties among Condorcet winners are resolved randomly (there can be multiple Condorcet winners because the inequality in the definition of F^C is weak).[15]

II. Axioms

We now define our axioms, which with two small exceptions are standard in the voting literature.[16] We say that a voting rule satisfies a given axiom on domain U if the axiom holds for all profiles $u_{(\cdot)}$ drawn from U.

Pareto on U: For all $u_{(\cdot)}$ on U, $Y \subseteq X$ and $x, y \in Y$, if $u_i(x) > u_i(y)$ for all i, then $y \notin F(u_{(\cdot)},Y)$. That is, if everyone prefers x to y and x is on the ballot, then y can't be elected.

Anonymity on U: Fix any measure-preserving[17] permutation of the electorate $\pi: [0,1] \rightarrow [0,1]$. For any $u_{(\cdot)}$ on U, let $u^{\pi}_{(\cdot)}$ be the profile such that, for all i, $u^{\pi}_i = u_{\pi(i)}$. Then, for all Y, if $x = F(u_{(\cdot)},Y)$, we have $x = F(u^{\pi}_{(\cdot)},Y)$. In words, if we permute a profile so that voter j gets i's preferences, k gets j's preferences, etc., the winner remains the same.

[15] Strictly speaking, majority rule as defined here is not a voting rule on domains for which a Condorcet winner may not exist, since, by definition, a voting rule produces a winner. In Section IV, we extend the definition of majority rule to ensure that a winner always exists.

[16] Decisiveness and strategy-proofness are slightly nonstandard because they explicitly deal with ties (ties are usually ruled out by assumption; for example, in the literature on majority rule the number of the voters is typically assumed to be odd).

[17] "Measure-preserving" means that, for any $C \subseteq [0,1]$, $\mu(C) = \mu(\pi(C))$.

Neutrality on U: Fix any ballot Y and any permutation $\rho : Y \rightarrow Y$ of Y. For any profile $u_{(\cdot)}$ on U, suppose $u_{(\cdot)}^{\rho,Y}$ is a profile on U such that, for all i, $u_i^{\rho,Y}(\rho(x)) = u_i(x)$ for all $x \in Y$. Then, if $x = F(u_{(\cdot)},Y)$, we have $F(u_{(\cdot)}^{\rho,Y},Y) = \rho(x)$. That is, suppose we start with a profile on U and we (i) permute the candidates so that candidate x becomes y, y becomes z, etc., and (ii) permute voters' utilities for the candidates correspondingly. Assume that the resulting profile is on U. Then if x won originally, now y wins.

Pareto, anonymity, and neutrality are so "natural" in political elections that few voting rules used in practice or studied theoretically violate any of them. The same is not true of the next axiom.

IIA on U: For any $u_{(\cdot)}$ on U and any ballot Y, if $x = F(u_{(\cdot)},Y)$ and $x \in Y' \subseteq Y$, then $x = F(u_{(\cdot)},Y')$.

As mentioned before, IIA implies that voting rules shouldn't be vulnerable to vote splitting. It rules out plurality rule, runoff voting, and rank-order voting (leaving only majority rule, approval voting, and range voting from the introduction).

Ordinality on U: For all $u_{(\cdot)}$ and $u'_{(\cdot)}$ on U and all $Y \subseteq X$, if $u_i(x) > u_i(y) \Leftrightarrow u'_i(x) > u'_i(y)$ for all $i \in [0,1]$ and all $x,y \in Y$, then $F(u_{(\cdot)},Y) = F(u'_{(\cdot)},Y)$. That is, only voters' rankings—and not cardinal information about preferences—determine the winner.

We next turn to decisiveness, the principle that the winner should be deterministic. In fact, *none* of the voting rules from the introduction is fully decisive in this sense; for each, ties may occur (and so, from neutrality, must be broken stochastically). For example, with plurality rule, two (or more) candidates might be ranked first by a maximal proportion of voters. Nevertheless, if the number of voters is large, the likelihood of a tie under plurality rule is small. That is why we assume a continuum of voters: the probability of a tie under plurality rule is zero, or, more precisely, ties are *nongeneric*. To express this formally, fix a (finitely based) voting rule F with base set S and mappings $h_i : U_X \rightarrow S$ for all $i \in [0,1]$. Given profile $u_{(\cdot)}$, let

$$m_s = \mu\big(\{i | h_i(u_i) = s\}\big) \quad \text{for each } s \in S,$$

that is, m_s is the proportion of voters whose utility functions correspond to s in $u_{(\cdot)}$.

Decisiveness on U: For any Y, F results in a deterministic winner for generic $(m_{s_1}, \ldots, m_{s_{|S|}})$ on U, that is, the Lebesgue measure of the set of $|S|$-tuples for which there are ties is zero when profiles are drawn from U.

It is easy to verify that all the voting rules in the introduction (except majority rule) satisfy decisiveness on any domain U.

We can now state a version of the Arrow impossibility theorem.

THEOREM A (Arrow 1951): *If $|X| \geq 3$*[18] *and $U = U_X$, there exists no voting rule on U satisfying all of Pareto, anonymity, neutrality, IIA, decisiveness, and ordinality.*

[18] For any set T, $|T|$ denotes the number of elements of T.

In view of this negative result, Dasgupta-Maskin (2008) considers restricted domains U. Specifically, although majority rule F^C fails to determine a winner on U_X, as Condorcet's own example (Figure 2) illustrates, this problem cannot arise on a domain U that does not contain *Condorcet cycles* (i.e., U doesn't contain utility functions u, u', u'' and candidates x, y, z such that $u(x) > u(y) > u(z)$, $u'(y) > u'(z) > u'(x)$, and $u''(z) > u''(x) > u''(y)$).

THEOREM B (Dasgupta-Maskin 2008; see also Sen 1966 and Inada 1969)*: Majority rule satisfies Pareto, anonymity, neutrality, IIA, ordinality, and decisiveness on U if and only if U does not contain Condorcet cycles. Moreover, when U doesn't have Condorcet cycles, Condorcet winners are generically strict* (*i.e., the winner beats all other candidates by a strict majority*).

Furthermore, majority rule dominates all other voting rules in the sense that it satisfies the axioms on a wider class of domains than any other.

THEOREM C (Dasgupta-Maskin 2008): *If F satisfies Pareto, anonymity, neutrality, IIA, decisiveness, and ordinality on domain U, then majority rule F^C also satisfies these axioms on U. Furthermore, if $F(u_{(\cdot)},Y) \neq F^C(u_{(\cdot)},Y)$ for some profile $u_{(\cdot)}$ on U, then there exists domain U' on which F^C satisfies all the axioms but F does not.*

The current paper's contribution is to add *strategy-proofness* to the mix.

(**Group**) **Strategy-Proofness on *U*:** For a generic profile $u_{(\cdot)}$ on U, all coalitions $C \subseteq [0,1]$, all profiles $u'_{(\cdot)}$ (with $u'_i = u_i$ for all $i \notin C$) on U, and all ballots Y, suppose $x = F(u_{(\cdot)},Y)$ and $x \neq F(u'_{(\cdot)},Y)$. Then, there exist $i \in C$ and $y \in \operatorname{supp} F(u'_{(\cdot)},Y)$[19] such that $u_i(x) > u_i(y)$. That is, if coalition C causes the winner to change from x to random variable $\tilde{y}$ by manipulating preferences from u_C to u'_C, someone in the coalition doesn't gain from the manipulation.[20]

III. Strategy-Proofness and Ordinality

Together with decisiveness, strategy-proofness implies that a voting rule is generically ordinal.

THEOREM 1: *Suppose that F satisfies strategy-proofness and decisiveness on U. Then, F satisfies ordinality for generic profiles on U.*

PROOF:

Suppose, to the contrary, that there exist generic profiles $u^*_{(\cdot)}$ and $u^{**}_{(\cdot)}$ and ballot $Y \subseteq X$ such that $u^*_i(x) > u^*_i(y) \Leftrightarrow u^{**}_i(x) > u^{**}_i(y)$ for all $i \in [0,1]$ and $x, y \in Y$ and yet $x^* \neq x^{**}$, where $x^* = F(u^*_{(\cdot)},Y)$ and $x^{**} = F(u^{**}_{(\cdot)},Y)$. We will show that transforming $u^*_{(\cdot)}$ to $u^{**}_{(\cdot)}$ one ordering at a time leads to contradiction.

[19] supp $F(u'_{(\cdot)},Y)$ is the set of candidates from which the winner is selected randomly.

[20] If supp $F(u'_{(\cdot)},Y)$ is multivalued, then implicitly we are assuming that voter i is deterred from deviating by the positive probability of a realization y worse than x.

Let $\succ^1$ be an ordering of Y, and let $u^1_{(\cdot)}$ be the profile such that, for all i,

$$u_i^1 = \begin{cases} u_i^{**}, & \text{if } \succ^1 \text{ is the ordering corresponding to } u_i^{**}; \\ u_i^*, & \text{otherwise.} \end{cases}$$

Take $x^1 = F(u^1_{(\cdot)}, Y)$. If $x^* \succ^1 x^1$, then voters with ordering $\succ^1$ in profile $u^1_{(\cdot)}$ (each such voter i has utility function u_i^{**}) are better off manipulating to make the profile $u^*_{(\cdot)}$ (i.e., voter i will pretend to have utility function u_i^*). If $x^1 \succ^1 x^*$, then voters with ordering $\succ^1$ in profile $u^*_{(\cdot)}$ are better off manipulating to make the profile $u^1_{(\cdot)}$. Thus, from strategy-proofness, we must have $x^1 = x^*$.

Next choose $\succ^2 \neq \succ^1$, and let $u^2_{(\cdot)}$ be the profile such that, for all i,

$$u_i^2 = \begin{cases} u_i^{**}, & \text{if} \succ^2 \text{is the ordering corresponding to } u_i^{**}; \\ u_i^1, & \text{otherwise.} \end{cases}$$

By similar argument, $x^2 = x^*$ where $x^2 = F(u^2_{(\cdot)}, Y)$. Continuing iteratively, we eventually obtain $u^n_{(\cdot)} = u^{**}_{(\cdot)}$ (since there are only finitely many orderings of Y), and thus $x^n = x^*$, a contradiction of $x^{**} = F(u^{**}_{(\cdot)}, Y)$ and $x^* \neq x^{**}$. ∎

IV. Results for Majority Rule

In view of Theorems A and 1, we immediately obtain the following result.

THEOREM D: *If* $|X| \geq 3$ *there exists no voting rule satisfying strategy-proofness, Pareto, anonymity, neutrality, IIA, and decisiveness on* U_X.[21]

Hence, we show that Theorem B continues to hold if we add strategy-proofness to the list of axioms.

THEOREM 2: *Majority rule* F^C *satisfies strategy-proofness, Pareto, anonymity, neutrality, IIA, and decisiveness on U if and only if U does not contain Condorcet cycles.*[22]

PROOF:

If U contains a Condorcet cycle, then from Figure 2, F^C isn't even a voting rule. For the converse, it suffices—in view of Theorem B—to show that F^C satisfies strategy-proofness on U. Suppose, to the contrary, there exist generic $u_{(\cdot)}$ and profile $u'_{(\cdot)}$ on U and coalition C such that

(1) $$x = F^C(u_{(\cdot)}, Y), \qquad \text{where } x \text{ is a strict Condorcet winner,}$$

[21] This result also follows directly from Gibbard (1973) and Satterthwaite (1975), except that they also impose ordinality.

[22] In view of Theorem 1, ordinality is redundant. For the case of single-peaked preferences, this result is implicit in Dummett and Farquharson (1961). Theorem 2 extends the result to *any* domain without Condorcet cycles.

and

(2) $$u_i(y) > u_i(x) \quad \text{for all } i \in C,$$

for all

(3) $$y \in \operatorname{supp} F^C(u'_{(\cdot)}, Y),$$

where $u'_j = u_j$ for all $j \notin C$. From (3),

(4) $$\mu\big(\{i \mid u'_i(y) > u'_i(x)\}\big) \geq \frac{1}{2}.$$

Hence, from (2) and (4),

$$\mu\big(\{i \mid u_i(y) > u_i(x)\}\big) \geq \frac{1}{2},$$

which contradicts (1), the fact that x is a strict Condorcet winner for $u_{(\cdot)}$. ∎

Our definition of strategy-proofness presumes that voters can manipulate preferences only within U. This presumption makes sense in some circumstances. For example, suppose there are two goods—one public and one private—and that a "candidate" x consists of a level p of the public good together with a tax cp levied on each citizen, where c is the per capita cost of the public good in terms of the private good. Consider the mechanism in which each citizen i chooses p_i, the *median choice* p^* is implemented, and each citizen pays cp^*. If citizens' preferences are convex and increasing in the two goods, then preferences for candidates x are *single-peaked*. Hence, the mechanism results in a Condorcet winner (see Black 1948).

Implicitly, the mechanism constrains a citizen to submit single-peaked preference and thus presumes that the planner knows in advance that preferences *are* single peaked. While this may be plausible in the public good context, knowing how preferences are restricted for presidential elections seems less likely. In such settings, constraining manipulations to U seems unrealistic. Thus, the definition of strategy-proofness becomes the following.

Strategy-Proofness* on U: For each $C \subseteq [0,1]$ and generic profile $u_{(\cdot)}$ on U and profile $u'_{(\cdot)}$ on U_X[23] (with $u'_j = u_j$ for all $j \notin C$), suppose $x = F(u_{(\cdot)}, Y)$ and $x \neq F(u'_{(\cdot)}, Y)$. Then there exist $i \in C$ and $y \in \operatorname{supp} F(u'_{(\cdot)}, Y)$ such that $u_i(x) > u_i(y)$.

Since coalitions now can manipulate outside U, a Condorcet winner may not exist.

Following Smith (1973) and Fishburn (1977), define the *Smith set* $Z(u_{(\cdot)}, Y)$ ($\subseteq Y$) for profile $u_{(\cdot)}$ and ballot Y to be the set of all Condorcet winners (if there are any) or else a minimal set of candidates such that, for each $x \in Z(u_{(\cdot)}, Y)$ and each $y \notin Y - Z(u_{(\cdot)}, Y)$, a majority of voters prefer x to y. The *Smith set* is unique (as Fishburn shows); it is a natural generalization of the majority winner

[23] The only change in going from strategy-proofness to strategy-proofness* is that $u'_{(\cdot)}$ is no longer restricted to U.

concept. Indeed, Fishburn (1977) argues that the following extension of majority voting rule best preserves the spirit of Condorcet:

$$F^{C^*}(u_{(\cdot)},Y) = \begin{cases} x, & \text{if } x \text{ is the unique Condorcet winner for } u_{(\cdot)} \text{ and } Y; \\ q(Z(u_{(\cdot)},Y)), & \text{if a Condorcet winner doesn't exist or is multiple.} \end{cases}$$

THEOREM 3: *Extended majority rule F^{C^*} satisfies strategy-proofness*, Pareto, anonymity, neutrality, IIA, and decisiveness on U if and only if U contains no Condorcet cycles.*

PROOF:

In view of the proof of Theorem 2, we need show only that if U contains no Condorcet cycles and $u_{(\cdot)}$ is a generic profile on U, then no coalition C gains by manipulating. Suppose, to the contrary, that $x = F^{C^*}(u_{(\cdot)},Y)$ is a strict Condorcet winner and coalition C gains from manipulation $u'_{(\cdot)}$ (where $u'_i = u_i$ for all $i \notin C$).

If $Z(u'_{(\cdot)},Y)$ consists of Condorcet winners, then we obtain the same contradiction as in the proof of Theorem 2. Hence, suppose that $Z(u'_{(\cdot)},Y) = \{x^1, \ldots, x^m\}$, where no x^i is a Condorcet winner for $u'_{(\cdot)}$.

Assume first that $x \in \{x^1, \ldots, x^m\}$. By definition of the Smith set, there exists $k \in \{1, \ldots, m\}$ such that

$$(5) \qquad \mu\left(\left\{i \,|\, u'_i(x^k) > u'_i(x)\right\}\right) \geq \frac{1}{2} \quad \text{(otherwise, } x \text{ is a Condorcet winner).}$$

Because C gains from the manipulation,

$$(6) \qquad u_i(x^k) > u_i(x) \quad \text{for all } i \in C.$$

And so, from (5) and (6),

$$\mu\left(\left\{i \,|\, u_i(x^k) > u_i(x)\right\}\right) \geq \frac{1}{2},$$

which contradicts the fact that x is a strict Condorcet winner for u.

If $x \notin \{x^1, \ldots, x^m\}$, then (5) holds for all $k \in \{1, \ldots, m\}$, and the rest of the argument is the same. ∎

Strikingly, the axioms under discussion *uniquely* characterize majority rule on any domain that admits a voting rule satisfying these axioms.

THEOREM 4: *If F satisfies strategy-proofness, Pareto, anonymity, neutrality, IIA, and decisiveness on U, then F is majority rule on U.*[24]

Remark 1: This result generalizes the classic axiomatization of majority rule by May (1952) to the case of three or more alternatives. May's characterization focuses on the case $|Y| = 2$, which is of only limited interest because then plurality rule,

[24] This result still holds if strategy-proofness is replaced by strategy-proofness*.

runoff voting, rank-order voting, and many other rules all coincide with majority rule. At first glance, May's axioms may look somewhat different from ours: only neutrality and anonymity overlap. In particular, he imposes (in addition to neutrality and anonymity) *positive responsiveness*,[25] and we don't. However, in the case $|Y| = 2$, this axiom immediately implies strategy-proofness and decisiveness. Also, our list of axioms, unlike May's, includes the Pareto property. But we require this only because, in contrast to May, we allow for restricted preference domains. Of course, our IIA axiom has no bite when $|Y| = 2$.

Remark 2: An important difference between Theorem 4 and Theorem C is that, in the latter, Pareto, anonymity, neutrality, IIA, decisiveness, and ordinality don't uniquely characterize F^C.

PROOF:

The proof is remarkably simple. Suppose F satisfies the axioms on U. From Theorem 1, we can confine attention to ordinal preferences (rankings). Assume first that $|Y| = 2$, that is, $Y = \{x, y\}$. If U contains only the ranking $\begin{smallmatrix} x \\ y \end{smallmatrix}$, then the result follows from Pareto. Hence, assume that U contains both $\begin{smallmatrix} x \\ y \end{smallmatrix}$ and $\begin{smallmatrix} y \\ x \end{smallmatrix}$. If, contrary to the theorem,

$$(7)\quad F\begin{pmatrix} \underline{a} & \underline{1-a} & \\ x & y & ,\{x,y\} \\ y & x & \end{pmatrix} = y, \quad \text{where } a > \frac{1}{2} \text{ for some generic profile } \begin{matrix} \underline{a} & \underline{1-a} \\ x & y \\ y & x \end{matrix},$$

then from neutrality and anonymity,

$$(8)\qquad F\begin{pmatrix} \underline{a} & \underline{1-a} & \\ y & x & ,\{x,y\} \\ x & y & \end{pmatrix} = x.$$

But, in profile (8), if a coalition of voters with ranking $\begin{smallmatrix} y \\ x \end{smallmatrix}$ pretends to have ranking $\begin{smallmatrix} x \\ y \end{smallmatrix}$ so as to attain profile (7), then they attain outcome y, which they prefer to x. Hence strategy-proofness is violated, and the theorem is established for $|Y| = 2$.

Assume next that $|Y| > 2$. If, for some profile, everyone prefers x to y but F chooses y, then we have a contradiction of Pareto. Hence, assume that there exist x and y and a generic profile such that

$$F\begin{pmatrix} \underline{a} & \underline{1-a} & \\ x & y & ,Y \\ y & x & \end{pmatrix} = y, \quad \text{where } a > \frac{1}{2}.$$

[25] Positive responsiveness says that if we alter voters' preferences so that all voters like x at least as much vis à vis y as they did before and some now like x strictly more (and no other changes to preferences are made), then (i) if x and y were both chosen by F before, now x is uniquely chosen and (ii) if x was uniquely chosen before, it still is.

Then, from IIA,

$$F\begin{pmatrix} \underline{a} & \underline{1-a} & \\ x & y & ,\{x,y\} \\ y & x & \end{pmatrix} = y,$$

contradicting the previous paragraph. ∎

Strategy-proofness is demanding in the sense that a voting rule must be unmanipulable by coalitions of any size. Let us relax this axiom. Call a voting rule F *ε-manipulable* on U if, for all $\varepsilon > 0$, there exist coalition C with $\mu(C) < \varepsilon$, profiles $u_{(\cdot)}$ and $u'_{(\cdot)}$ on U (with $u'_i = u_i$ for all $i \notin C$), ballot Y, and $x, y \in Y$ such that $x = F(u_{(\cdot)}, Y)$, $y = F(u'_{(\cdot)}, Y)$, and $u_i(y) > u_i(x)$ for all $i \in C$. That is, a voting rule is ε-manipulable if there exists a coalition of arbitrarily small size that can benefit from misrepresenting. Because imposing a limit on a coalition's size makes profitable manipulations harder, the following is a relaxation of strategy-proofness.

Weak Strategy-Proofness on U**:** F is not ε-manipulable on U.[26]

Theorems 2 and 3 clearly continue to hold when weak strategy-proofness replaces strategy-proofness because the latter implies the former. Showing that Theorem 1 still holds with this substitution is also quite easy.[27] We now establish a stronger version of Theorem 4.

THEOREM 5: *If F satisfies weak strategy-proofness, Pareto, anonymity, neutrality, IIA, and decisiveness on U, then F is majority rule on U.*[28]

PROOF:

Suppose F satisfies the axioms on U. As in the proof of Theorem 4, we start with the case $Y = \{x, y\}$. Fix $\varepsilon > 0$, and suppose there exists generic profile $\begin{matrix} \underline{a} & \underline{1-a} \\ x & y \\ y & x \end{matrix}$ where

(9) $$a > \frac{1}{2} > 1 - a \quad \text{with } 2a - 1 < \varepsilon$$

and

(10) $$F\begin{pmatrix} \underline{a} & \underline{1-a} \\ x & y \\ y & x \end{pmatrix} = y.$$

[26] We thank Shengwu Li, who suggested that we consider this version of strategy-proofness.

[27] Specifically, consider the proof of Theorem 1 and fix $\varepsilon > 0$. For each ordering $\succ^j$, we can partition the sets of utility functions u_i^* and u_i^{**} corresponding to $\succ^j$ into subsets of measure no greater than ε. Then rather than changing all the u_i^* terms to u_i^{**} at the same time (as in the current proof), we can change the subsets *sequentially* (so no coalition manipulating is bigger than ε).

[28] As with Theorem 4, we can replace weak strategy-proofness with weak strategy-proofness* (the analog of strategy-proofness*).

Then we can apply neutrality and anonymity to obtain a contradiction of weak strategy-proofness just like the contradiction of strategy-proofness in the proof of Theorem 4, since the manipulation entailed is by a coalition of size less than ε. Hence, for any a satisfying (9) we have

$$F\begin{pmatrix} \underline{a} & \underline{1-a} \\ x & y \\ y & x \end{pmatrix} = x. \tag{11}$$

Next consider a' with

$$a' > \frac{1}{2} > 1 - a' \quad \text{with } 2a' - 1 < 2\varepsilon. \tag{12}$$

If

$$F\begin{pmatrix} \underline{a'} & \underline{1-a'} \\ x & y \\ y & x \end{pmatrix} = y,$$

then, from (11) and (12), a coalition of voters with ranking $\begin{smallmatrix} x \\ y \end{smallmatrix}$ and size smaller than ε can pretend to have ranking $\begin{smallmatrix} y \\ x \end{smallmatrix}$ and thereby change the outcome to x, contradicting weak strategy-proofness.

Proceeding iteratively, we can show that (11) holds for all a satisfying $a > 1 - a$. Hence $F = F^C$ on U. The rest of the proof is the same as that for Theorem 4.[29] ∎

V. Concluding Remarks

We have shown that, as long as a domain is free of Condorcet cycles, majority rule satisfies all the standard axioms for voting rules—Pareto, anonymity, neutrality, decisiveness, IIA, and strategy-proofness—and that no other voting rule satisfies these axioms, irrespective of the domain. A sufficient condition ruling out Condorcet cycles is that each voter be *ideological*: he ranks candidates according to how far away from him they are on a left-right continuum (this is an example of the more general sufficient condition of *single-peaked preferences*; for another example, see the discussion of public goods in Section IV). Another sufficient condition is that, among each group of three candidates, there is one whom voters feel "strongly" about: he might be ranked first or third, but never second (Donald Trump seems to have been such a candidate in 2016). A wealth of evidence suggests that, in actual elections, Condorcet winners nearly always exist. For example, Popov, Popova, and Regenwetter (2014) study elections for the presidency of the American Psychological Association and find "virtually no trace of a Condorcet paradox" (see Gehrlein 2006, chapter 2 for other evidence on the existence of Condorcet winners).

Indeed, by highlighting the theoretical shortcoming of majority rule, Condorcet may have done himself—and the world—an injustice.

[29] The proof of Theorem 5 indicates that if we worked with a finite number of voters instead of a continuum, we could establish the uniqueness result using ordinary strategy-proofness (rather than group strategy-proofness).

REFERENCES

Arrow, Kenneth J. 1951. *Social Choice and Individual Values*. Hoboken: John Wiley and Sons.

Balinski, Michel, and Rida Laraki. 2010. *Majority Judgment: Measuring, Ranking, and Electing*. Cambridge, MA: MIT Press.

Black, Duncan. 1948. "On the Rationale of Group Decision-Making." *Journal of Political Economy* 56 (1): 23–34.

Borda, Jean-Charles de. 1781. "Mémoire sur les Élections au Scrutin." *Hiswire de I'Academie Royale des Sciences*: 657–65.

Condorcet, Marie Jean Antoine Nicolas Caritat de. 1785. *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*. Paris: Imprimerie Royale.

Dasgupta, Partha, and Eric Maskin. 2008. "On the Robustness of Majority Rule." *Journal of the European Economic Association* 6 (5): 949–73.

Dummett, Michael, and Robin Farquharson. 1961. "Stability in Voting." *Econometrica* 29 (1) 33–43.

Fishburn, Peter C.1977. "Condorcet Social Choice Functions." *SIAM Journal on Applied Mathematics* 33 (3): 469–89.

Gehrlein, William V. 2006. *Condorcet's Paradox*. Berlin: Springer.

Gibbard, Allan. 1973. "Manipulation of Voting Schemes: A General Result." *Econometrica* 41 (4): 587–601.

Inada, Ken-ichi. 1969. "The Simple Majority Decision Rule." *Econometrica* 37 (3): 490–506.

Maskin, Eric. 1995. "Majority Rule, Social Welfare Functions, and Game Form." In *Choice, Welfare and Development: A Festschrift in Honour of Amartya K. Sen*, edited by Kaushik Basu, Prasanta K. Pattanaik, and Kōtarō Suzumura, chapter 6. Oxford: Clarendon Press.

May, Kenneth O. 1952. "Set of Necessary and Sufficient Conditions for Simple Majority Decisions." *Econometrica* 20 (4): 680–84.

Miller, Nicholas R. 1977. "Graph-Theoretical Approaches to the Theory of Voting." *American Journal of Political Sciences* 21 (4): 769–803.

Nash, John F., Jr. 1950. "The Bargaining Problem." *Econometrica* 18 (2): 155–62.

Popov, Sergey V., Anna Popova, and Michael Regenwetter. 2014. "Consensus in Organizations: Hunting for the Social Choice Conundrum in APA Elections." *Decision* 1 (2): 123–46.

Satterthwaite, Mark Allen. 1975. "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions." *Journal of Economic Theory* 10: 187–217.

Sen, Amartya K. 1966. "A Possibility Theorem on Majority Decisions." *Econometrica* 34 (2): 491–99.

Smith, John H. 1973. "Aggregation of Preferences with Variable Electorate." *Econometrica* 41 (6): 1027–41.