Strategy-Proofness, IIA, and Majority Rule[*]

Partha Dasgupta[†] and Eric Maskin[††]

March 2020

[†] Faculty of Economics and Politics, University of Cambridge, Cambridge UK. partha.dasgupta@econ.cam.ac.uk
[††] Department of Economics, Harvard University, Cambridge MA 02138 USA. emaskin@fas.harvard.edu
   Higher School of Economics, Moscow

Abstract

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.

1. Introduction

Current election methods leave much 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.

Of course, there are many possible election methods, called *voting rules,* to choose from. Here are a few examples.

In *plurality rule* (used to elect MPs in the U.K. and members of Congress in the U.S., and to award a state's electoral votes in U.S. presidential elections), each citizen votes for a candidate, and the winner is the candidate with the most votes,[2] even if short of majority.[3]

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 that determines the winner.[4]

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

---

[2] 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 3.
[3] Plurality rule was the method the Republican Party adopted for many of its 2016 primaries.
[4] Runoff voting is used for presidential elections in France, Russia, Brazil, and many other countries.

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 3). [5] Next are two voting rules that are cardinal (i.e., a citizen's vote depends 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. [6]

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. [7]

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

We use the axiomatic approach here (section 3 gives precise definitions of our axioms, which are all familiar from the literature). Specifically, we suppose that there is a large number

---

[5] More accurately, the way the citizen votes can be deduced if he is voting non-strategically. We consider *strategic voting* – a major theme of this paper – later in this introduction.

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

[7] 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.

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, for short) – 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 [8]; and *ordinality* – the winner should depend only on citizens' *ordinal* rankings and not on preference intensities or other cardinal information.[9]

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, 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 and Figure 2 for a violation by the Borda count.

The Arrow Impossibility Theorem establishes that there is *no* voting rule satisfying all of Pareto, anonymity, neutrality, decisiveness, IIA, and ordinality with at least three candidates and unrestricted voter preferences (see Theorem A in section 3).[10] In particular, majority rule –

---

[8] 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).

[9] 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.

[10] 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.

although it satisfies the other axioms – fails to be decisive, as Condorcet himself showed in a famous example of a "Condorcet cycle" (see Figure 3).

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 that domain does not contain a Condorcet cycle.[11] 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 voting rule we started with is itself majority rule, there exists another domain on which majority rule satisfies all six axioms and the original voting rule does not (Theorem C).

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 work out his own preferences: he has to study the candidates' characters, backgrounds, and positions on the issues. If, on top of this, he must know *other* citizens' preferences and strategies in order to react strategically to them, his decision

---

[11] A sufficient condition ruling out Condorcet cycles is that each voter is *ideological*: he ranks candidates according to how far away they are from him on a left-right continuum (this is an example of *single-peaked preferences*; for another example, see the discussion of public goods in section 5). Another sufficient condition is that among each group of 3 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 from places where voters have been ranking candidates for many years (e.g., Australia and Ireland) suggests that, in actual elections, Condorcet winners nearly always exist.

problem becomes several orders of magnitude 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 4) 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 5).

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 5, 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.[12] 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[13] 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, 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.

2. Model

There is a finite set *X* of potential candidates for a given office[14]. 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 3).

---

[12] We could alternatively work with a refinement of the Smith set, e.g., the *uncovered set*, but this wouldn't change our analysis or results significantly.

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

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

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* $u.$ 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 correspondence that, for each profile $u.$ and each ballot Y, selects a winner $F(u., Y) \in \Delta Y$, 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. [15] 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 $u.$ and $u.'$, if $h_i(u_i) = h_i(u_i')$ for all $i \in [0,1]$, then for all $Y \subseteq X, F(u., Y) = F(u.', Y)$. 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

---

[15] This focus makes it easier to define the concept of a generic profile: see the discussion of decisiveness in section 3.

use Lebesgue measure μ on [0,1]. Thus, for profile $u_.$, $\mu\left(\left\{i \mid u_i(x) > u_i(y)\right\}\right)$ is the proportion of voters who prefer candidate x to candidate y.

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:

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

$$F^P(u_.,Y) = q\left(\left\{x \in Y \mid \mu\left(\left\{i \mid u_i(x) > u_i(y) \text{ for all } y \neq x, y \in Y\right\}\right)\right.\right.$$

$$\geq \mu\left(\left\{i \mid u_i(x') > u_i(y) \text{ for all } y \neq x', y \in Y\right\}\right) \text{ for all } x' \in Y\right\}\right)$$

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:

Majority Rule (Condorcet's method):

$$F^C(u_.,Y) = q(\left\{x \in Y \mid \mu(\left\{i \mid u_i(x) > u_i(y)\right\} \geq \frac{1}{2}, \text{ for all } y \neq x, y \in Y\right\})$$

That is, candidate x is a Condorcet winner if, for any other candidate y, a majority prefer x to y; ties are resolved randomly.[16]

---

[16] Strictly speaking, majority rule is not a voting rule on a domain for which a Condorcet winner may not exist.

3. Axioms

We now define our axioms, which with two small exceptions are standard in the voting literature.[17] We say that a voting rule satisfies a given axiom on domain U if the axiom holds for all profiles u. drawn from U.

Pareto on U: For all u. on U, $Y \subseteq X$ and $x, y \in Y$, if $u_i(x) > u_i(y)$ for all i, then $y \notin F(u., Y)$.

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[18] permutation of the electorate $\pi : [0,1] \to [0,1]$.

For any u. on U, let $u.^{\pi}$ be the profile such that, for all i, $u_i^{\pi} = u_{\pi(i)}$. Then, for all Y, if

$x = F(u., Y)$, we have $x = F(u.^{\pi}, Y)$. 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.

Neutrality on U: Fix any ballot Y and any permutation $\rho : Y \to Y$ of Y. For any profile u. on U,

suppose $u.^{\rho,Y}$ 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

$x = F(u., Y)$, we have $F(u.^{\rho,Y}, Y) = \rho(x)$. That is, suppose we start with a profile in 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 in 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.

---

[17] 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).

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

11

IIA on *U*: For any *u.* 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).

We next have:

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 – determines the winner.

We next turn to decisiveness, the principle that there should always be a nonstochastic winner. In fact, *none* of the voting rules from the introduction is fully decisive in this sense; for each, ties may occur. 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 $h_i : U_X \to S$ for all $i \in [0,1]$. Given profile $u_{\cdot}$, let

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

i.e., 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 unique deterministic winner for generic $(m_{s_1}, \ldots, m_{s_{|S|}})$ on *U*, i.e., 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 $|X| \geq 3$ [19] and $U = U_X$, there exists no voting rule on U satisfying all of Pareto, anonymity, neutrality, IIA, decisiveness, and ordinality.

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 3) 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): F^C 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 F^C also satisfies these axioms on U. Furthermore, if $F(u,.Y) \neq F^C(u_., Y)$ for some profile $u_.$ 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:

---

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

(*Group*) *Strategy-Proofness* on *U*: For a generic profile $u_.$ on *U,* all coalitions $C \subseteq [0,1]$, all

profiles $u'_.$ (with $u'_i = u_i$ for all $i \notin C$) on *U* and all ballots *Y*, suppose $x = F(u_.,Y)$. Then, there

exist $i \in C$ and $y \in$ supp $F(u'_.,Y)^{20}$ 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. [21]

4. 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 $u_.^*$ and $u_.^{**}$ 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

$x^* = F(u_.^*,Y)$ and $x^{**} = F(u_.^{**},Y)$. We will show that transforming $u_.^*$ to $u_.^{**}$ one ordering at a

time leads to contradiction.

  Let $\succ^1$ be an ordering of *Y* and let $u_.^1$ be the profile such that, for all *i*,

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

---

[20] supp $F(u'_.,Y)$ is the set of candidates from which the winner is selected randomly.

[21] If supp $F(u'_.,Y)$ is multi-valued, then implicitly we are assuming that voter *i* is deterred from deviating by the positive probability of a realization *y* worse than *x*.

14

Take $x^1 = F(u_{\cdot}^1, Y)$. If $x^* \succ^1 x^1$, then voters with ordering $\succ^1$ in profile $u_{\cdot}^1$ (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 $u_{\cdot}^1$. Thus, from strategy-proofness, we must have $x^1 = x^*$.

Next choose $\succ^2 \neq \succ^1$ and let $u_{\cdot}^2$ 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 $x^2 = F(u_{\cdot}^2, Y)$. Continuing iteratively, we eventually obtain

$u_{\cdot}^n = u_{\cdot}^{**}$ (since there are only finitely many orderings of Y) and thus $x^n = x^*$, a contradiction of

$x^{**} = F(u_{\cdot}^{**}, Y)$ and $x^* \neq x^{**}$. Q.E.D.

5. Results for Majority Rule

In view of Theorems A and 1, we immediately obtain

Theorem D: If $|X| \geq 3$ there exists no voting rule satisfying strategy-proofness, Pareto,

anonymity, neutrality, IIA, and decisiveness on U_X.[22]

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

axioms:

---

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

***Theorem* 2:** F^C satisfies strategy-proofness, Pareto, anonymity, neutrality, IIA, and decisiveness

on U if and only if U does not contain Condorcet cycles.[23]

***Proof*:** If U contains a Condorcet cycle, then from Figure 2, F^C violates decisiveness. 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 and profile u' on U, and coalition C such that

(1) $x = F^C(u_., Y)$, where x is a strict Condorcet winner

and

(2) $u_i(y) > u_i(x)$ for all $i \in C$,

for all

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

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

(4) $\mu\left(\left\{i \big| u'_i(y) > u'_i(x)\right\}\right) \geq \dfrac{1}{2}$.

Hence, from (2) and (4),

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

which contradicts (1), the fact that x is a strict Condorcet winner for u. Q.E.D

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

---

[23] 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.

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:

Strategy-proofness* on U : For each $C \subseteq [0,1]$ and generic profile $u_{\cdot}$ on U and profile $u'_{\cdot}$ on U_X^{24}

(with $u'_j = u_j$ for all $j \notin C$), suppose $x = F(u_{\cdot}, Y)$. Then there exist $i \in C$ and $y \in \text{supp } F(u'_{\cdot}, Y)$

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 concept. Indeed, Fishburn (1977) argues that the following extension of majority voting

rule best preserves the spirit of Condorcet:

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

---

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

Theorem 3: 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 $u.$ is a generic profile on U, then no coalition C gains by manipulating. Suppose, to the contrary, that $x = F^{C^*}(u., Y)$ is a strict Condorcet winner and coalition C gains from manipulation $u'.$ (where $u'_i = u_i$ for all $i \notin C$).

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

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

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

Because C gains from the manipulation

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

And so, from (5) and (6)

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

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

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

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 $F = F^C$ on U.[25]

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, 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*,[26] 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 $|Y| = 2$, i.e.,

---

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

[26] 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.

$Y = \{x, y\}$. If U contains only the ranking $\frac{x}{y}$, then the result follows from Pareto. Hence, assume

that U contains both $\frac{x}{y}$ and $\frac{y}{x}$. If, contrary to the theorem,

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

then from neutrality and anonymity

$$(8)\quad F\left(\begin{matrix} a & 1-a \\ y & x \\ x & y \end{matrix}, \{x, y\}\right) = x$$

But, in profile (8), if a coalition of voters with ranking $\frac{y}{x}$ pretends to have ranking $\frac{x}{y}$ 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

$$F\left(\begin{matrix} a & 1-a \\ | & | \\ x & y \\ | & | \\ y & x \\ | & | \end{matrix}, Y\right) = y, \text{where } a > \frac{1}{2}$$

Then, from IIA

$$F\left(\begin{matrix} a & 1-a \\ x & y \\ y & x \end{matrix}, \{x, y\}\right) = y, \text{ contradicting the previous paragraph. Q.E.D.}$$

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 $\varepsilon-manipulable$ on U if, for all $\varepsilon > 0,$ there exist coalition C with $\mu(C) < \varepsilon,$ profiles $u_{\boldsymbol{\cdot}}$ and $u'_{\boldsymbol{\cdot}}$ on U (with $u'_i = u_i$ for all $i \notin C$), ballot $Y,$ and $x, y \in Y$ such that $x = F(u_{\boldsymbol{\cdot}}, Y),$ $y = F(u'_{\boldsymbol{\cdot}}, Y),$ and $u_i(y) > u_i(x)$ for all $i \in C.$ That is, a voting rule is $\varepsilon-$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 $\varepsilon-$manipulable on U.[27]

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.[28] 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 $F = F^C$ on $U.$[29]

***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 $\begin{matrix} a & 1-a \\ x & y \\ y & x \end{matrix}$ where

(9) $a > \dfrac{1}{2} > 1 - a$ with $2a - 1 < \varepsilon$

and

---

[28] 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 $\varepsilon.$ Then rather than changing all the u_i^*s to $u_i^{**}s$ *at the same time* (as in the current proof), we can change the subsets *sequentially* (so no coalition manipulating is bigger than ε).

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

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

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

$$(11) \quad F \begin{pmatrix} a & 1-a \\ x & y \\ y & x \end{pmatrix} = x$$

Next consider a' with

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

If

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

then, from (11) and (12), a coalition of voters with ranking $\frac{x}{y}$ and size smaller than ε can pretend to have ranking $\frac{y}{x}$ 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.

Q.E.D.

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

For the rankings above, Trump wins (with 40%) under plurality rule. However, both Kasich and 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 3-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.

Figure 1 Vote Splitting

45%	55%	
x	y	Under rank-order voting,
z	x	x gets $3 \times 45 + 2 \times 55 = 245$ points
y	z	y gets $3 \times 55 + 1 \times 45 = 210$ points
		z gets $2 \times 45 + 1 \times 55 = 145$ points
		so x wins

But if z doesn't run, we have

45%	55%	
x	y	Under rank-order voting,
y	x	x gets $2 \times 45 + 1 \times 55 = 145$ points
		y gets $2 \times 55 + 1 \times 45 = 155$ points
		so y wins

Figure 2 Rank-order Voting (Borda count)

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

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

Figure 3 Condorcet cycles

References

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

Balinski, Michel and Rida Laraki. 2010. *Majority Judgment: Measuring, Ranking, and Electing*. 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.* 657-665.

Condorcet, Marie Jean Antoine Nicolas Caritat. 1785. *Essai sur l'application de l'analyse à la pluralité des voix*. Imprimerie Royale.

Dasgupta, Partha and Eric Maskin. 2008. "On the Robustness of Majority Rule." *Journal of the European Economic Association.* 6(5): 949-973.

Dummett, Michael and Robin Farquharson. 1961. "Stability in voting." *Econometrica*, 33-43.

Fishburn, Peter C.1977. "Condorcet Social Choice Functions." *SIAM Journal on Applied Mathematics*. 33(3): 469-489.

Gibbard, Allan. 1973. "Manipulation of Voting Schemes: A General Result." *Econometrica.* 41: 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 Kaushik Basu, Prasanta K. Pattanaik, and Kōtarō Suzumura, eds., *Choice, Welfare and Development*. Clarendon Press.

May, Kenneth. 1952. "Set of Necessary and Sufficient Conditions for Simple Majority Decisions." *Econometrica*. 20(4): 680-684.

Nash, John F.1950. "The Bargaining Problem." *Econometrica*. 18(2): 155-162.

Satterthwaite, Mark A. 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.1966. "A Possibility Theorem on Majority Decisions." *Econometrica.* 34(2): 491-499.

Smith, John H.1973. "Aggregation of Preferences with Variable Electorate." *Econometrica* 41(6):1027-1041.