ON THE ROBUSTNESS OF MAJORITY RULE

PARTHA DASGUPTA AND ERIC MASKIN

ABSTRACT

We show that simple majority rule satisfies five standard and attractive axioms—the Pareto property, anonymity, neutrality, independence of irrelevant alternatives, and (generic) decisiveness—over a larger class of preference domains than (essentially) any other voting rule. Hence, in this sense, it is the most robust voting rule. This characterization of majority rule provides an alternative to that of May (1952). (JEL: D71)

1. INTRODUCTION

How should a society select a president? How should a legislature decide which version of a bill to enact?

Reprinted from *Journal of the European Economic Association*, John Wiley and Sons, vol. 6(5), 949–973.

The casual response to these questions is probably to recommend that a vote be taken. But there are many possible voting rules—majority rule, plurality rule, rank-order voting, unanimity rule, runoff voting, and a host of others (a *voting rule*, in general, is any method for choosing a winner from a set of candidates on the basis of voters' reported preferences for those candidates[1])—and so this response, by itself, does not resolve the question. Accordingly, the theory of voting typically attempts to evaluate voting rules systematically by examining which fundamental properties or axioms they satisfy.

One generally accepted axiom is the *Pareto property*, the principle that if all voters prefer candidate *x* to candidate *y*, then *y* should not be chosen over *x*.[2] A second axiom with strong appeal is *anonymity*, the notion that no voter should have more influence on the outcome of an election than any other[3] (anonymity is sometimes called the "one person–one vote" principle). Just as anonymity demands that all voters be treated alike, a third principle, *neutrality*, requires the same thing for candidates: No candidate should get special treatment.[4]

Three particularly prominent voting rules that satisfy all three axioms—Pareto, anonymity, and neutrality—are (i) *simple majority rule*, according to which candidate *x* is chosen if, for all other candidates *y* in the feasible set, more voters prefer *x* to *y* than *y* to *x*; (ii) *rank-order voting* (also called the Borda count[5]), under which each candidate gets one point for every voter who ranks her first, two points for every voter who ranks her second, and so forth, with candidate *x* being chosen if *x*'s point total is lowest among those in the feasible set; and (iii) *plurality rule* (also called "first past the post"), according to which candidate *x* is chosen if more voters rank *x* first than they do any other feasible candidate.

But rank-order voting and plurality rule fail to satisfy a fourth standard principle, *independence of irrelevant alternatives* (IIA), which has attracted considerable attention since its emphasis by Nash (1950) and Arrow (1951).[6] IIA dictates that if candidate *x* is chosen from the feasible set, and now some other candidate *y* is removed from that set, then *x* is still chosen.[7] To see why rank-order voting and plurality rule violate IIA, consider an electorate consisting of 100 voters. Suppose that there are four feasible candidates *w*, *x*, *y*, *z*, and that the distribution of rankings is as follows:

Number of voters	**47**	**49**	**4**
Ranking (listed vertically from best to worst)	*x*	*y*	*w*
	y	*z*	*x*
	z	*x*	*y*
	w	*w*	*z*

Then, under rank-order voting, *y* will win the election for this profile (a *profile* is a specification of all voters' rankings) with a point total of 155 (49 points from 49 first-place votes, 94 points from 47 second-place votes, and 12 points from 4 third-place votes), compared with point totals of 202 for *x*, 388 for *w*, and 255 for *z*. Candidate *y* will also win under plurality rule: *y* gets 49 first-place votes whereas

x and z get only 47 and 4, respectively. Observe, however, that if z is eliminated from the feasible set, then x will win under rank-order voting with a point total of 153 (47 points from first-place votes and 106 points from second-place votes) compared with y's point total of 155. Moreover, if w is removed from the feasible set (either instead of or in addition to z), then x will also win under plurality rule: It is now top-ranked by 51 voters, whereas y has only 49 first-place votes. Thus, whether the candidates w and z are present or absent from the feasible set determines the outcome under both rank-order voting and plurality rule, contradicting IIA. Furthermore, this occurs even though neither w nor z comes close to winning under either voting rule (i.e., they are "irrelevant alternatives").

Under majority rule (we will henceforth omit the qualification "simple" when this does not cause confusion with other variants of majority rule), by contrast, the choice between x and y turns only on how many voters prefer x to y and how many prefer y to x—not on whether other candidates are also options. Thus, in our 100-voter example, x is the winner (she beats all other candidates in head-to-head comparisons) whether or not w or z is on the ballot. In other words, majority rule satisfies IIA.

But majority rule itself has a well-known flaw, discovered by Borda's archrival the Marquis de Condorcet (1785) and illustrated by the so-called paradox of voting (or Condorcet paradox): it may fail to generate any winner. Specifically, suppose there are three voters 1, 2, 3 and three candidates x, y, z, and suppose the profile of voters' preferences is

1	2	3
x	y	z
y	z	x
z	x	y

(i.e., voter 1 prefers x to y to z, voter 2 prefers y to z to x, and voter 3 prefers z to x to y). Then, as Condorcet noted, a (two-thirds) majority prefers x to y, so that y cannot be chosen; a majority prefers y to z, so that z cannot be chosen; and a majority prefers z to x, so that x cannot be chosen. That is, majority rule fails to select any alternative; it violates *decisiveness*, which requires that a voting rule pick a (unique) winner.

In view of the failure of these three prominent voting methods—rank-order voting, plurality rule, and majority rule—to satisfy all of the five axioms (Pareto, anonymity, neutrality, IIA, and decisiveness), it is natural to inquire whether there is some other voting rule that might succeed where they fail. Unfortunately, the answer is negative: *no* voting rule satisfies all five axioms when there are three or more candidates (see Theorem 1), a result closely related to Arrow's (1951) impossibility theorem.

Still, there is an important sense in which this conclusion is too pessimistic: It presumes that, in order to satisfy an axiom, a voting rule must conform to that axiom regardless of what voters' preferences turn out to be.[8] In practice, however, some preferences may be highly unlikely. One reason for this may be ideology. As Black (1948) notes, in many elections the typical voter's attitudes toward the leading candidates will be governed largely by how far away they are from his own position in left-right ideological space. In the 2000 U.S. presidential election—where the four major candidates from left to right were Ralph Nader, Al Gore, George W. Bush, and Pat Buchanan—a voter favoring Gore might thus have had the ranking

Gore		Gore
Nader	or even	Bush
Bush		Nader
Buchanan		Buchanan

but would most likely not have ranked the candidates as

Gore
Buchanan
Bush
Nader

because Bush is closer to Gore ideologically than Buchanan is. In other words, the graph of a voter's utility for candidates will be *single-peaked* when the candidates are arranged ideologically on the horizontal axis. Single-peakedness is of interest because, as Black shows, majority rule satisfies decisiveness generically[9] when voters' preferences conform to this restriction.

In fact, single-peakedness is by no means the only plausible restriction on preferences that ensures the decisiveness of majority rule. The 2002 French presidential election, where the three main candidates were Lionel Jospin (Socialist), Jacques Chirac (Conservative), and Jean-Marie Le Pen (National Front), offers another example. In that election, voters—regardless of their views on Jospin and Chirac—had strong views on Le Pen: polls suggested that, among the three candidates, he was ranked either first or third by nearly everybody; very few voters placed him second. Whether such polarization is good for France is open to debate, but it is definitely good for majority rule: As we will see in Section 4, such a restriction—in which one candidate is never ranked second—guarantees, like single-peakedness, that majority rule will be generically decisive.

Thus, majority rule *works well*—in the sense of satisfying our five axioms—for some domains of voters' preferences (e.g., a domain of single-peaked preferences), but not for others (e.g., the unrestricted domain). A natural issue to raise, therefore, is how its performance compares with that of other voting rules. As we have already noted, no voting rule can work well for all domains. So the obvious question to ask is: Which voting rules work well for the biggest class of domains?[10]

We show that majority rule is (essentially) the unique answer to this question. Specifically, we establish (see Theorem 2) that, if a given voting rule *F* works well on a domain of preferences, then majority rule works well on that domain, too. Conversely, if *F* differs from majority rule,[11] then there exists some other domain on which majority rule works well but *F* does not.

Thus, majority rule is essentially the unique voting rule that works well on the most domains; it is, in this sense, the most *robust* voting rule.[12] Indeed, this gives us a characterization of majority rule (see Theorem 3) that differs from the classic one derived by May (1952). For the case of two alternatives,[13] May shows that majority rule is the unique voting rule satisfying a weak version of decisiveness, anonymity, neutrality, and a fourth property, *positive responsiveness.*[14] Our Theorem 3 strengthens decisiveness, omits positive responsiveness, and imposes Pareto and IIA to obtain a different characterization.

Theorem 2 is also related to a result obtained in Maskin (1995).[15] Like May, Maskin imposes somewhat different axioms from ours. In particular, instead of decisiveness—which requires that there be a unique winner—he allows for the possibility of multiple winners but insists on *transitivity* (indeed, the same is true of earlier versions of this article; see Dasgupta and Maskin 1998): If x beats y and y beats z, then x should beat z. But more significantly, his proposition requires two strong and somewhat unpalatable assumptions. The first is that the number of voters be odd. This is needed to rule out exact ties: situations where exactly half the population prefers x to y and the other half prefers y to x (oddness is also needed for much of the early work on majority rule; see, e.g., Inada 1969). In fact, our own results also call for avoiding such ties. But rather than simply assuming an odd number of voters, we suppose that the number of voters is large, implying that an exact tie is unlikely even if the number is not odd. Hence, we suppose a large number of voters and ask only for *generic* decisiveness (i.e., decisiveness for "almost all" profiles). Formally, we work with a continuum of voters,[16] but it will become clear that we could alternatively assume a large but finite number by defining generic decisiveness to mean "decisive for a sufficiently high proportion of profiles." In this way we avoid "oddness," an unappealing assumption because it presumably holds only half the time.

Second, Maskin's (1995) proof invokes the restrictive assumption that the voting rule *F* being compared with majority rule satisfies Pareto, anonymity, IIA, and neutrality on any domain. This is quite restrictive because, although it accommodates certain methods (such as the supermajority rules and the Pareto-extension rule—the rule that chooses all Pareto optimal alternatives), it eliminates such voting rules as the Borda count, plurality voting, and run-off voting. These are the most common alternatives in practice to majority rule, yet they fail to satisfy IIA on the unrestricted domain. We show that this assumption can be dropped altogether.

We proceed as follows. In Section 2, we set up the model. In Section 3, we give formal definitions of our five properties: Pareto, anonymity, neutrality, independence of irrelevant alternatives, and generic decisiveness. We also show (Theorem 1) that no voting rule always satisfies these properties—that is, always works well. In Section 4 we establish three lemmas that characterize when rank-order voting, plurality rule, and majority rule work well. We use the third lemma in Section 5 to establish our main result, Theorem 2. We obtain our alternative to May's (1952) characterization as Theorem 3. Finally, in Section 6 we discuss two extensions.

2. THE MODEL

Our model in most respects falls within a standard social choice framework. Let *X* be the set of social alternatives (including alternatives that may turn out to be infeasible), which, in a political context, is the set of candidates.

For technical convenience, we take X to be finite with cardinality $m \geq 3$. The possibility of individual indifference often makes technical arguments in the social choice literature a great deal messier (see, e.g., Sen and Pattanaik 1969). We shall simply rule it out by assuming that voters' preferences over X can be represented by *strict orderings*[17] (with only a finite number of alternatives, the assumption that a voter is not exactly indifferent between any two alternatives does not seem very strong). If R is a strict ordering of X, then, for any two alternatives $x, y \in X$ with $x \neq y$, the notation xRy denotes "x is strictly preferred to y in ordering R." For any subset $Y \subseteq X$ and any strict ordering R, let $R|_Y$ be the restriction of R to Y.

Let $\mathfrak{R}_X$ be the set of all logically possible strict orderings of X. We shall typically suppose that voters' preferences are drawn from some subset $\mathfrak{R} \subseteq \mathfrak{R}_X$. For example, for some sequential arrangement $(x_1, x_2, \ldots, x_m)$ of the social alternatives, $\mathfrak{R}$ consists of *single-peaked preferences* (relative to this arrangement) if, for all $R \in \mathfrak{R}$ whenever $x_i R x_{i+1}$ for some i we have $x_j R x_{j+1}$ for all $j > i$ (i.e., if x lies between x_i and x_j in the arrangement, then a voter cannot prefer both x_i and x_j to x).

For the reason mentioned in the Introduction (and elaborated on hereafter), we shall suppose that there is a continuum of voters, indexed by points in the unit interval $[0, 1]$. A *profile* $\boldsymbol{R}$ on $\mathfrak{R}$ is a mapping $\boldsymbol{R}:[0, 1] \to \mathfrak{R}$, where, for any $i \in [0, 1]$, $\boldsymbol{R}(i)$ is voter i's preference ordering. Hence, profile $\boldsymbol{R}$ is a specification of the preferences of all voters. For any $Y \subseteq X$, $\boldsymbol{R}|_Y$ is the profile $\boldsymbol{R}$ restricted to Y.

We shall use Lebesgue measure μ as our measure of the size of voting blocs. Given alternatives x and y with $x \neq y$ and profile $\boldsymbol{R}$, let $q_{\mathrm{R}}(x, y) = \mu\{i | x\boldsymbol{R}(i)y\}$.[18] Then $q_R(x, y)$ is the fraction of the population preferring x to y in profile $\boldsymbol{R}$.

A *voting rule F* is a mapping that, for each profile[19] $\boldsymbol{R}$ on $\mathfrak{R}_X$ and each subset $Y \subseteq X$, assigns a (possibly empty) subset $F(\boldsymbol{R}, Y) \subseteq X$, where if $\boldsymbol{R}|_Y = \boldsymbol{R}'|_Y$, then $F(\boldsymbol{R}, Y) = F(\boldsymbol{R}', Y)$.[20] As suggested in the Introduction, Y can be interpreted as the *feasible set* of alternatives and $F(\boldsymbol{R}, Y)$ as the *winning candidate(s)*.

For example, suppose that F^M is *simple majority rule*. Then, for all $\boldsymbol{R}$ and Y,

$$F^M(\boldsymbol{R}, Y) = \{x \in Y \mid q_{\mathrm{R}}(x, y) \geq q_{\mathrm{R}}(y, x) \text{ for all } y \in Y - \{x\}\};$$

in other words, x is a winner in Y provided that, for any other alternative $y \in Y$, the proportion of voters preferring x to y is no less than the proportion preferring y to x. Such an alternative x is called a *Condorcet winner*. Note that there may not always be a Condorcet winner—that is, $F^M(\boldsymbol{R}, Y)$ need not be nonempty (as when the profile corresponds to that in the Condorcet paradox).

The supermajority rules provide a second example. For instance, *two-thirds majority* rule $F^{2/3}$ can be defined so that, for all $\boldsymbol{R}$ and Y,

$$F^{2/3}(\boldsymbol{R}, Y) = Y'.$$

Here Y' is a nonempty subset of Y such that, for all $x, y \in Y'$ with $x \neq y$ and all $z \in Y - Y'$, we have $q_R(y, x) < 2/3$ and $q_R(x, z) \geq 2/3$, provided that such a subset exists (if Y' exists, then it is clearly unique). That is, x is a winner if it beats all nonwinners by at least a two-thirds majority and if no other winner beats it by a two-thirds majority or more.

As a third example, consider rank-order voting. Given $R \in \Re_X$ and Y, let $v_R^Y(x)$ be 1 if x is the top-ranked alternative of R in Y, let it be 2 if x is second-ranked in Y, and so on. Then, given profile $\boldsymbol{R}$, it follows that $\int_0^1 v_{\boldsymbol{R}(i)}^Y(x)d\mu(i)$ is the total number of points assigned to x—that is, alternative x's rank-order score or *Borda count*. If F^{RO} is *rank-order voting*, then, for all $\boldsymbol{R}$ and Y,

$$F^{RO}(\boldsymbol{R},Y)=\left\{x\in Y\middle|\int_0^1 v_{\boldsymbol{R}(i)}^Y(x)d\mu(i)\leq\int_0^1 v_{\boldsymbol{R}(i)}^Y(y)d\mu(i),\text{ for all } y\in Y\right\}.$$

That is, x is a rank-order winner if no other alternative in Y has a lower rank-order score.

Finally, consider plurality rule F^P defined so that, for all $\boldsymbol{R}$ and Y,

$$\begin{aligned} F^P(\boldsymbol{R},Y)=\{x\in Y\,|\,\mu\{i\,|\,x\,\boldsymbol{R}(i)\,y \text{ for all } y\in Y-\{x\}\} \\ \geq\mu\{i\,|\,z\,\boldsymbol{R}(i)\,y \text{ for all } y\in Y-\{z\}\} \text{ for all } z\in Y\}. \end{aligned}$$

That is, x is a plurality winner if it is top-ranked in Y for at least as many voters as any other alternative in Y.

3. THE PROPERTIES

We consider five standard properties that one may wish a voting rule to satisfy.

Pareto Property on $\mathfrak{R}$. For all $\boldsymbol{R}$ on $\mathfrak{R}$ and all $x, y \in X$ with $x \neq y$, if $x\boldsymbol{R}(i)y$ for all i, then, for all Y, $x \in Y$ implies $y \notin F(\boldsymbol{R}, Y)$.

In words, the Pareto property requires that, if all voters prefer x to y, then the voting rule should not choose y if x is feasible. Probably all voting rules used in practice satisfy this property. In particular, majority rule, rank-order voting, and plurality rule (as well as the supermajority rules) satisfy it on the unrestricted domain $\mathfrak{R}_X$.

Anonymity on $\mathfrak{R}$. Suppose that $\pi: [0,1] \to [0,1]$ is a measure-preserving permutation of $[0, 1]$ (by "measure-preserving" we mean that, for all Borel sets $T \subset [0,1], \mu(T) = \mu(\pi(T))$). If, for all $\boldsymbol{R}$ on $\mathfrak{R}$, $\boldsymbol{R}^{\pi}$ is the profile such that $\boldsymbol{R}^{\pi}(i) = \boldsymbol{R}(\pi(i))$ for all i, then, for all Y, $F(\boldsymbol{R}^{\pi}, Y) = F(\boldsymbol{R}, Y)$.

In words, anonymity means that the winner(s) should not depend on which voter has which preference; only the preferences themselves matter. Thus, if we permute the assignment of voters' preferences by π, the winners should remain the same. (The reason for requiring that π be measure-preserving is purely technical: to ensure that, for all x and y, the fraction of voters preferring x to y is the same for $\boldsymbol{R}^{\pi}$ as it is for $\boldsymbol{R}$.) Anonymity embodies the principle that everyone's vote should count equally.[21] It is obviously satisfied on $\mathfrak{R}_X$ by majority rule, plurality rule, and rank-order voting, as well as by all other voting rules that we have discussed so far.

Neutrality on $\mathfrak{R}$. For any subset $Y \subseteq X$ and profile $\boldsymbol{R}$ on $\mathfrak{R}$, let $\rho: Y \to Y$ be a permutation of Y and let $\boldsymbol{R}^{\rho,Y}$ be a profile on $\mathfrak{R}$ such that, for all i and all $x,\ y \in Y$ with $x \neq y$, $x\boldsymbol{R}(i)y$ if and only if $\rho(x)\boldsymbol{R}^{\rho,Y}(i)\rho(y)$. Then $\rho(F(\boldsymbol{R},Y)) = F(\boldsymbol{R}^{\rho,Y},Y)$.

In words, neutrality requires that a voting rule treat all alternatives symmetrically: If the alternatives are relabeled via ρ, then the winner(s) are relabeled in the same way. Once again, all the voting rules we have talked about satisfy neutrality, including majority rule, rank-order voting, and plurality rule.

As noted in the Introduction, we will invoke the Nash (1950) version of IIA as follows.

Independence of Irrelevant Alternatives on $\mathfrak{R}$. For all profiles $\boldsymbol{R}$ on $\mathfrak{R}$ and all Y, if $x \in F(\boldsymbol{R},Y)$ and if Y' is a subset of Y such that $x \in Y'$, then $x \in F(\boldsymbol{R},Y')$.

In words, IIA says that if x is a winner for some feasible set Y and we now remove some of the other alternatives from Y, then x will remain a winner. Clearly, majority rule satisfies IIA on the unrestricted domain $\mathfrak{R}_X$: if x beats each other alternative by a majority, then it continues to do so when any of those other alternatives are removed. However, rank-order voting and plurality rule violate IIA on $\mathfrak{R}_X$, as we already showed by example.

Finally, we require that voting rules select a single winner.

Decisiveness. For all $\boldsymbol{R}$ and Y, $F(\boldsymbol{R}, Y)$ is a *singleton*—that is, it consists of a unique element.

Decisiveness formalizes the reasonably uncontroversial goal that an election should result in a clear-cut winner.[22] However, it is somewhat too strong because it rules out ties, even if these occur only rarely. Suppose, say, that $Y = \{x, y\}$

and that exactly half the population prefers x to y and the other half prefers y to x. Then no neutral voting rule will be able to choose between x and y; they are perfectly symmetric in this profile. Nevertheless, this indecisiveness is a knife-edge phenomenon—it requires that the population be split precisely 50–50. Thus, there is good reason for us to disregard it as pathological or irregular. And, because we are working with a continuum of voters, there is a simple formal way to do so.

Specifically, let S be a subset of $\mathbb{R}_+$. A profile $\boldsymbol{R}$ on $\mathfrak{R}$ is *regular* with respect to S (which we call an *exceptional set*) if, for all alternatives x and y with $x \neq y$,

$$q_R(x, y)/q_R(y, x) \notin S.$$

In words, a regular profile is one for which the proportions of voters preferring one alternative to another all fall outside the specified exceptional set. We can now state the version of decisiveness that we will use.

Generic Decisiveness on $\mathfrak{R}$. There exists a finite exceptional set S such that, for all Y and all profiles $\boldsymbol{R}$ on $\mathfrak{R}$ that are regular with respect to S, $F(\boldsymbol{R}, Y)$ is a singleton.

Generic decisiveness requires that a voting rule be decisive for regular profiles, where the preference proportions do not fall into some finite exceptional set. For example, as Lemma 3 implies, majority rule is generically decisive on a domain of single-peaked preferences because there exists a unique winner for all regular profiles if the exceptional set consists of the single point 1 (i.e., $S = \{1\}$). It is this decisiveness requirement that works against such supermajority methods as two-thirds majority rule, which selects a unique winner x only if x beats all other alternatives by at least a two-thirds majority. In fact, in view of the Condorcet paradox, simple majority rule itself is not generically decisive on the domain $\mathfrak{R}_X$. By contrast, rank-order voting and plurality rule are generically decisive on all domains, including $\mathfrak{R}_X$.[23]

We shall say that a voting rule *works well* on a domain $\mathfrak{R}$ if it satisfies the Pareto property, anonymity, neutrality, IIA, and generic decisiveness on that domain. Thus, given our previous discussion, majority rule works well, for example, on a domain of single-peaked preferences. In Section 4 we provide general characterizations of when majority rule, plurality rule, and rank-order voting work well.

Although decisiveness is the only axiom for which we are considering a "generic" version, we could easily accommodate generic relaxations of the other conditions, too. However, this seems pointless, because, to our knowledge, no commonly used voting rule has nongeneric failures except with respect to decisiveness.

We can now establish the impossibility result that motivates our examination of restricted domains $\mathfrak{R}$.

Theorem 1. *No voting rule works well on* $\mathfrak{R}_X$.

Proof. Suppose, contrary to the claim, that F works well on $\mathfrak{R}_X$. We will use F to construct a social welfare function satisfying the Pareto property, anonymity, and IIA (the Arrow 1951 version), contradicting the Arrow impossibility theorem.

Let S be the exceptional set for F on $\Re$. Because S is finite (by definition of generic decisiveness), we can find an integer $n \geq 2$ such that, if we divide the population into n groups of equal size $[0, 1/n], (1/n, 2/n], (2/n, 3/n], \ldots, (n - 1/n, 1]$, then any profile for which all voters within a given group have the same ranking must be regular with respect to S. Given profile $\boldsymbol{R}$ for which all voters within a given group have the same ranking and $X' \subseteq X$, let $\boldsymbol{R}^{X'}$ be the same profile as $\boldsymbol{R}$ except that the elements of X' have been moved to the top of all voters' rankings: for all i and for all $x, y \in X$ with $x \neq y$, $x\boldsymbol{R}^{X'}(i)y$ if and only if

(a) $x\boldsymbol{R}(i)y$ and $x, y \in X'$; or
(b) $x\boldsymbol{R}(i)y$ and $x, y \notin X'$; or
(c) $x \in X'$ and $y \notin X'$.

Construct an n-person social welfare function $f : \Re_X^n \rightarrow \Re_X$ such that, for all n-tuples $(R_1, \ldots, R_n) \in \Re_X^n$ and $x, y \in X$ with $x \neq y$,

$$xf(R_1, \ldots, R_n)y \text{ if and only if } x \in F(\boldsymbol{R}^{\{x,y\}}, X). \quad (1)$$

Here $\boldsymbol{R}$ corresponds to $(R_1, \ldots, R_n)$; it is the profile such that, for all i and j, $\boldsymbol{R}(i) = R_j$ if and only if $i \in (j/n, j + 1/n]$ (i.e., if voter i belongs to group j). To begin with, f is well defined because, since F satisfies the Pareto Principle and generic decisiveness, either $x \in F(\boldsymbol{R}^{\{x,y\}}, X)$ or $y \in F(\boldsymbol{R}^{\{x,y\}}, X)$. Similarly, f satisfies the Pareto Principle and anonymity.[24] To see that f satisfies Arrow-IIA, consider two n-tuples $(R_1, \ldots, R_n)$ and $(\hat{R}_1, \ldots, \hat{R}_n)$ such that

$$(R_1, \ldots, R_n)|_{\{x,y\}} = (\hat{R}_1, \ldots, \hat{R}_n)|_{\{x,y\}}, \tag{2}$$

and let $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ be the corresponding profiles. From generic decisiveness, Pareto, and IIA, we obtain

$$\begin{aligned} F(\boldsymbol{R}^{\{x,y\}}, X) &= F(\boldsymbol{R}^{\{x,y\}}, \{x,y\} \in \{x,y\}, \\ F(\hat{\boldsymbol{R}}^{\{x,y\}}, X) &= F(\hat{\boldsymbol{R}}^{\{x,y\}}, \{x,y\} \in \{x,y\}. \end{aligned} \tag{3}$$

But from equation (2) and the definition of a voting rule it follows that $F(\boldsymbol{R}^{\{x,y\}}, \{x,y\}) = F(\hat{\boldsymbol{R}}^{\{x,y\}}, \{x,y\})$. Hence, by equations (1) and (3),

$$xf(R_1, \ldots, R_n)y \text{ if and only if } xf(\hat{R}_1, \ldots, \hat{R}_n)y$$

establishing Arrow-IIA.

Finally, we must show that f is transitive. That is, for any n-tuple $(R_1, \ldots, R_n)$ and distinct alternatives x, y, z for which $xf(R_1, \ldots, R_n)y$ and $yf(R_1, \ldots, R_n)z$, we must establish that $xf(R_1, \ldots, R_n)z$. Consider $F(\boldsymbol{R}^{\{x,y,z\}}, X)$, where $\boldsymbol{R}$ is the profile corresponding to $(R_1, \ldots, R_n)$. Because $\boldsymbol{R}^{\{x,y,z\}}$ is regular, generic decisiveness implies that $F(\boldsymbol{R}^{\{x,y,z\}}, X)$ is a singleton, and the Pareto property implies that $F(\boldsymbol{R}^{\{x,y,z\}}, X) \in \{x, y, \mathrm{z}\}$. If $F(\boldsymbol{R}^{\{x,y,z\}}, X) = y$ then, by IIA, $F(\boldsymbol{R}^{\{x,y,z\}}, \{x,y\}) = y$. But from $xf(R_1, \ldots, R_n)y$ and IIA we obtain $F(\boldsymbol{R}^{\{x,y\}}, X) = F(\boldsymbol{R}^{\{x,y\}}, \{x,y\}) = x$; this is a contradiction because, by definition of a voting rule, $F(\boldsymbol{R}^{\{x,y,z\}}, \{x,y\}) = F(\boldsymbol{R}^{\{x,y\}}\ \{x,y\})$. If $F(\boldsymbol{R}^{\{x,y,z\}}, X) = z$ then we can derive a similar contradiction from $yF(R_1, \ldots ,R_n)z$. Hence $F(\boldsymbol{R}^{\{x,y,z\}}, X) = x$, and so by definition we have $F(\boldsymbol{R}^{\{x,z\}}, X) = x$, implying that $xF(R_1, \ldots, R_n)\ z$. Thus, transitivity obtains and so f is a social welfare function satisfying Pareto, anonymity, and IIA. The Arrow impossibility theorem now applies to obtain the theorem (anonymity implies that Arrow's nondictatorship requirement is satisfied).

That F satisfies neutrality is a fact not used in the proof, so Theorem 1 remains true if we drop that desideratum from the definition of "working well."

4. CHARACTERIZATION RESULTS

We have seen that rank-order voting and plurality rule violate IIA on $\mathfrak{R}_X$. We now characterize the domains for which they do satisfy this property. For rank-order voting, "quasi-agreement" is the key.

Quasi-Agreement (QA) on $\mathfrak{R}$. Within each triple of distinct alternatives $\{x, y, z\} \subseteq X$, there exists an alternative, say x, that satisfies one of the following three conditions:

(i) for all $R \in \mathfrak{R}$, xR_y and xR_z;
(ii) for all $R \in \mathfrak{R}$, yR_x and zR_x;
(iii) for all $R \in \mathfrak{R}$, either yR_xR_z or zR_xR_y.

In other words, QA holds on domain $\mathfrak{R}$ if, for any triple of alternatives, all voters with preferences in $\mathfrak{R}$ agree on the relative ranking of one of these alternatives: either it is best within the triple, or it is worst, or it is in the middle.

LEMMA 1. *For any domain* $\mathfrak{R}$, *rank-order voting* F^{RO} *satisfies IIA on* $\mathfrak{R}$ *if and only if QA holds on* $\mathfrak{R}$.

Remark. Of our five principal axioms, rank-order voting violates only IIA on $\mathfrak{R}_X$. Hence, Lemma 1 establishes that rank-order voting works well on $\mathfrak{R}$ if and only if $\mathfrak{R}$ satisfies QA.

See the Appendix for the proof of Lemma 1.

We turn next to plurality rule, for which a condition called *limited favoritism* is needed for IIA.

Limited Favoritism (LF) on $\mathfrak{R}$. Within each triple of distinct alternatives $\{x, y, z\} \subseteq X$ there exists an alternative, say x, such that for all $R \in \mathfrak{R}$ either yR_x or zR_x.

That is, LF holds on domain $\mathfrak{R}$ if, for any triple of alternatives, there exists one alternative that is never the favorite (i.e., is never top-ranked) for preferences in $\mathfrak{R}$.

Lemma 2. *For any domain* $\mathfrak{R}$, *plurality rule* F^P *satisfies IIA on* $\mathfrak{R}$ *if and only if LF holds on* $\mathfrak{R}$.

Remark. Just as QA characterizes when rank-order voting works well, so Lemma 2 shows that LF characterizes when plurality rule works well, because the other four axioms are always satisfied. Indeed, LF also characterizes when a number of other prominent voting rules, such as *runoff voting*,[25] work well.

Proof of Lemma 2. Suppose first that $\mathfrak{R}$ satisfies LF. Consider profile $\boldsymbol{R}$ on $\mathfrak{R}$ and subset Y such that $x \in F^P(\boldsymbol{R}, Y)$ for some $x \in Y$. Then the proportion of voters in $\boldsymbol{R}$ who rank x first among alternatives in Y is at least as big as that for any other alternative. Furthermore, given LF, there can be at most one other alternative that is top-ranked by anyone. That is, x must get a majority of the first-place rankings among alternatives in Y. But clearly x will only increase its majority if some other alternative y is removed from Y. Hence, $F^P(\boldsymbol{R}, Y - \{y\}) = x$, and so F^P satisfies IIA on $\mathfrak{R}$.

Next suppose that domain $\hat{\mathfrak{R}}$ violates LF. Then there exist $\{x, y, z\}$ and $R_x, R_y, R_z \in \mathfrak{R}$ such that, within $\{x, y, z\}$, x is top-ranked for R_x, z is top-ranked for R_z, and $yR_y zR_y x$. Consider profile $\hat{\boldsymbol{R}}$ on $\mathfrak{R}$ such that 40% of voters have ordering R_x, 30% have R_y, and 30% have R_z. Then $x \in F^P(\hat{\boldsymbol{R}}, \{x, y, z\})$, because x is top-ranked by 40% of the population whereas y and z are top-ranked by only 30% each. Suppose now that y is removed from $\{x, y, z\}$. Note that $z \in F^P(\hat{\boldsymbol{R}}, \{x, z\})$, because z is now top-ranked by 60% of voters. Hence, F^P violates IIA on $\hat{\mathfrak{R}}$.

We turn finally to majority rule. We suggested in Section 3 that a single-peaked domain ensures generic decisiveness, and we noted in the Introduction that the same is true when the domain satisfies the property that, for every triple of alternatives, there is one that is never ranked second. But these are only sufficient conditions for generic transitivity; what we want is a condition that is both sufficient and necessary.

To obtain that condition we first note that, for any three alternatives x, y, z, there are six logically possible strict orderings, which can be sorted into two Condorcet "cycles":[26]

x	y	z	x	z	y
y	z	x	z	y	x
z	x	y	y	x	z
	cycle 1			cycle 2	

We shall say that a domain $\mathfrak{R}$ satisfies the *no-Condorcet-cycle* property[27] if it contains no Condorcet cycles. That is, for every triple of alternatives, at least one ordering is missing from each of cycles 1 and 2. (More precisely, for each triple $\{x, y, z\}$, there exist no orderings R, R', R'' in $\mathfrak{R}$ that, when restricted to $\{x, y, z\}$, generate cycle 1 or cycle 2.)

LEMMA 3. *Majority rule is generically decisive on domain $\mathfrak{R}$ if and only if $\mathfrak{R}$ satisfies the no-Condorcet-cycle property.*[28]

Proof. If there existed a Condorcet cycle for alternatives $\{x, y, z\}$ in $\mathfrak{R}$, then we could reproduce the Condorcet paradox by taking $Y = \{x, y, z\}$. Hence, the no-Condorcet-cycle property is clearly necessary.

To show that it is also sufficient, we must demonstrate, in effect, that the Condorcet paradox is the only thing that can interfere with majority rule's generic decisiveness. Toward this end, we suppose that F^M is not generically decisive on domain $\mathfrak{R}$. Then, in particular, if $S = \{1\}$ then there must exist Y and a profile $\boldsymbol{R}$ on $\mathfrak{R}$ that is regular with respect to $\{1\}$ but for which $F^M(\boldsymbol{R}, Y)$ is either empty or contains multiple alternatives. If there exist $x, y \in F^M(\boldsymbol{R}, Y)$ with $x \neq y$, then $q_{\boldsymbol{R}}(x, y) = q_{\boldsymbol{R}}(y, x) = 1/2$ and so

$$q_{\boldsymbol{R}}(x, y)/q_{\boldsymbol{R}}(y, x) = 1,$$

contradicting $\boldsymbol{R}$'s regularity with respect to $\{1\}$. Hence $F^M(\boldsymbol{R}, Y)$ must be empty. Choose $x_1 \in Y$. Then, because $x_1 \notin F^M(\boldsymbol{R}, Y)$, there exists an $x_2 \in Y$ such that

$$q_{\boldsymbol{R}}(x_2, x_1) > \frac{1}{2}.$$

Similarly, because $x_2 \notin F^M(\boldsymbol{R}, Y)$, there exists an $x_3 \in Y$ such that

$$q_R(x_3, x_2) > \frac{1}{2}.$$

Continuing in this way, we must eventually (because there are only finitely many alternatives in X) reach $x_t \in Y$ such that

$$q_R(x_t, x_{t-1}) > \frac{1}{2} \tag{4}$$

but with some $t < t$ for which

$$q_R(x_\tau, x_t) > \frac{1}{2}. \tag{5}$$

If t is the smallest index for which (5) holds, then

$$q_R(x_{t-1}, x_\tau) > \frac{1}{2}. \tag{6}$$

Combining (4) and (6), we conclude that there must be a positive fraction of voters in $\boldsymbol{R}$ who prefer x_t to x_{t-1} and x_{t-1} to x_τ; that is,

$$\begin{matrix} x_t \\ x_{t-1} \\ x_\tau \end{matrix} \in \Re. \tag{7}$$ [29]

Similarly, (5) and (6) yield

$$\begin{matrix} x_{t-1} \\ x_\tau \\ x_t \end{matrix} \in \Re,$$

and from (4) and (5) we obtain

$$\begin{matrix} x_\tau \\ x_t \\ x_{t-1} \end{matrix} \in \Re.$$

Hence, $\mathfrak{R}$ violates the no-Condorcet-cycle property, as was to be shown. □

It is easy to see that a domain of single-peaked preferences satisfies the no-Condorcet-cycle property. Hence, Lemma 3 implies that majority rule is generically decisive on such a domain. The same is true of the domain we considered in the Introduction in connection with the 2002 French presidential election.

The results of this section give us an indication of the stringency of the requirement of "working well" across our three voting rules. Lemma 1 establishes that, for any triple of alternatives, four of the six possible strict orderings must be absent from a domain $\mathfrak{R}$ in order for rank-order voting to work well on $\mathfrak{R}$. By contrast, Lemmas 2 and 3 show that only two orderings must be absent if we instead consider plurality rule or majority rule (although LF is strictly a more demanding condition than the no-Condorcet-cycle property[30]).

5. THE ROBUSTNESS OF MAJORITY RULE

We can now state our main finding as follows.

THEOREM 2. *Suppose that voting rule F works well on domain $\mathfrak{R}$. Then majority rule F^M works well on $\mathfrak{R}$ too. Conversely, suppose that F^M works well on domain $\mathfrak{R}^M$. Then, if there exists a profile $\boldsymbol{R}^\circ$ on $\mathfrak{R}^M$, regular with respect to F's exceptional set, such that*

$$F(\boldsymbol{R}^\circ, Y) \neq F^M(\boldsymbol{R}^\circ, Y) \text{ for some } Y, \tag{8}$$

then there exists a domain $\mathfrak{R}'$ on which F^M works well but F does not.

Remark 1. Without the requirement that the profile $\boldsymbol{R}^\circ$ for which F and F^M differ belong to a domain on which majority rule works well, the second assertion of Theorem 2 would be false. In particular, consider a voting rule that coincides with majority rule except for profiles that violate the no-Condorcet-cycle property. It is easy to see that such a rule works well on any domain for which majority rule does because it coincides with majority rule on such a domain.

Remark 2. Theorem 2 allows for the possibility that $\mathfrak{R}' = \mathfrak{R}^M$, and indeed this equality holds in the example we consider after the proof. However, more generally, F may work well on $\mathfrak{R}^M$ even though equation (8) holds, in which case $\mathfrak{R}'$ and $\mathfrak{R}^M$ must differ.

Proof of Theorem 2. Suppose first that F works well on $\mathfrak{R}$. If, contrary to the theorem, F^M does not work well on $\mathfrak{R}$, then by Lemma 3 there exists a Condorcet cycle in $\mathfrak{R}$:

$$\begin{matrix} x & y & z & \\ y & z & x & \in \mathfrak{R} \\ z & x & y & \end{matrix} \tag{9}$$

for some $x, y, z \in X$. Let S be the exceptional set for F on $\mathfrak{R}$. Because S is finite (by definition of generic transitivity), we can find an integer n such that, if we divide the population into n equal groups, then any profile for which all voters in each particular group have the same ordering in $\mathfrak{R}$ must be regular with respect to S.

Let $[0, 1/n]$ be group 1, let $(1/n, 2/n]$ be group 2, . . . , and let $(n - 1/n, 1]$ be group n. Consider a profile $R1$ on $\Re$ such that all voters in group 1 prefer x to z and all voters in the other groups prefer z to x. That is, the profile is

$$\begin{array}{cccc} \underline{1} & \underline{2} & \cdots & \underline{n} \\ x & z & & z. \\ z & x & & x \end{array} \tag{10}$$

Because F is generically decisive on $\Re$ and because $\boldsymbol{R}_1$ is regular, there are two cases: $F(\boldsymbol{R}_1, \{x, z\}) = z$ or x.

Case (i)*:* $F(\boldsymbol{R}_1, \{x, z\}) = z$. Consider a profile $\boldsymbol{R}_1^*$ on $\Re$ in which all voters in group 1 prefer x to y to z, all voters in group 2 prefer y to z to x, and all voters in the remaining groups prefer z to x to y. That is,[31]

$$\boldsymbol{R}_1^* = \begin{array}{ccccc} \underline{1} & \underline{2} & \underline{3} & \cdots & \underline{n}. \\ x & y & z & & z \\ y & z & x & & x \\ z & x & y & & y \end{array} \tag{11}$$

By (9), such a profile exists on $\Re$. Notice that, in profile $\boldsymbol{R}_1^*$, voters in group 1 prefer x to z and that all other voters prefer z to x. Hence, the case (i) hypothesis implies that

$$F(\boldsymbol{R}_1^*, \{x, z\}) = z. \tag{12}$$

From equation (12) and IIA, $F(\boldsymbol{R}_1^*, \{x, y, z\}) \neq x$. If $F(\boldsymbol{R}_1^*, \{x, y, z\}) = y$, then neutrality implies that

$$F(\hat{\boldsymbol{R}}_1^*, \{x, y, z\}) = x \tag{13}$$

where

$$\hat{R}_1^* = \begin{array}{ccccc} \underline{1} & \underline{2} & \underline{3} & \cdots & \underline{n} \\ z & x & y & & y \\ x & y & z & & z \\ y & z & x & & x \end{array}.$$

and $\hat{R}_1^*$ is on $\Re$. Then IIA yields $F(\hat{R}_1^*, \{x, z\}) = x$, which by anonymity, contradicts the case (i) hypothesis. Hence, $F(R_1^*, \{x, y, z\}) = z$ and so, by IIA,

$$F(R_1^*, \{y, z\}) = z. \tag{14}$$

Applying neutrality, we obtain from equation (14) that

$$F(R_2, \{y, z\}) = y,$$

where

$$R_2 = \begin{array}{ccccc} \underline{1} & \underline{2} & \underline{3} & \cdots & \underline{n} \\ z & z & y & & y \\ x & x & z & & z \\ y & y & x & & x \end{array} \tag{15}$$

and R_2 is on $\Re$. Applying neutrality once again then gives

$$F(\hat{R}_2, \{x, z\}) = z, \tag{16}$$

where

$$\hat{R}_2 = \begin{array}{ccccc} \underline{1} & \underline{2} & \underline{3} & \cdots & \underline{n} \\ x & x & z & & z \\ y & y & x & & x \\ z & z & y & & y \end{array} \tag{17}$$

and $\hat{R}_2$ is on $\mathfrak{R}$. Formulas (16) and (17) establish that if z is chosen over x when just one of n groups prefers x to z (case (i) hypothesis), then z is again chosen over x when two of n groups prefer x to z as in (16).

Now, choose $\boldsymbol{R}_2^*$ on $\mathfrak{R}$ so that

$$\boldsymbol{R}_2^* = \begin{array}{cccccc} \underline{1} & \underline{2} & \underline{3} & \underline{4} & \cdots & \underline{n} \\ x & y & y & z & & z \\ y & z & z & x & & x \\ z & x & x & y & & y \end{array}.$$

Arguing as we did for $\boldsymbol{R}_1^*$, we can show that $F(\boldsymbol{R}_2^*, \{y, z\}) = z$ and then apply neutrality twice to conclude that z is chosen over x if three groups out of n prefer x to z. Continuing iteratively, we conclude that z is chosen over x even if $n-1$ groups out of n prefer x to z—which, in view of neutrality, violates the case (i) hypothesis. Hence, this case is impossible.

Case (ii): $F(\boldsymbol{R}_1, \{x, z\}) = x$. From the case (i) argument, case (ii) leads to the same contradiction as before. We conclude that F^M must work well on $\mathfrak{R}$ after all, as claimed.

For the converse, suppose that there exist (a) domain $\mathfrak{R}^M$ on which F^M works well and (b) Y and $x, y \in Y$ and regular profile $\boldsymbol{R}^\circ$ on $\mathfrak{R}^M$ such that

$$y = F(\boldsymbol{R}^\circ, Y) \neq F^M(\boldsymbol{R}^\circ, Y) = x. \qquad (18)$$

If F does not work well on $\mathfrak{R}^M$, then we can take $\mathfrak{R}' = \mathfrak{R}^M$ to complete the proof. Hence, assume that F works well on $\mathfrak{R}^M$ with exceptional set S. From IIA and equation (17) (and because $\boldsymbol{R}^{\circ}$ is regular), there exists an $\alpha \in (0,1)$ with $\alpha/(1-\alpha) \notin S, (1-\alpha)/\alpha \notin S$,

$$1-\alpha > \alpha, \tag{19}$$

and $q_{R^{\circ}}(x,y) = 1-\alpha$ such that $F^M(\boldsymbol{R}^{\circ}, \{x,y\}) = x$ and

$$F(\boldsymbol{R}^{\circ}, \{x,y\}) = y. \tag{20}$$

Consider $z \notin \{x, y\}$ and profile $\boldsymbol{R}^{\circ\circ}$ such that

$$\boldsymbol{R}^{\circ\circ} = \begin{array}{ccc} \underline{[0,\alpha)} & \underline{[\alpha, 1-\alpha)} & \underline{[1-\alpha, 1]} \\ z & z & x \\ y & x & z \\ x & y & y \end{array}. \tag{21}$$

Observe that in equation (21) we have left out the alternatives other than x, y, and z. To make matters simple, assume that the orderings of $\boldsymbol{R}^{\circ\circ}$ are all the same for those other alternatives. Suppose, furthermore, that in these orderings x, y, and z are each preferred to any alternative not in $\{x, y, z\}$. Then, because $\alpha/(1-\alpha) \notin / S$ and $(1-\alpha)/\alpha \notin S$, it follows that $\boldsymbol{R}^{\circ\circ}$ is regular.

Let $\hat{\mathfrak{R}}'$ consist of the orderings in $\boldsymbol{R}^{\circ\circ}$ together with ordering $\begin{matrix} x \\ y \\ z \end{matrix}$ (where $\{x, y, z\}$ are ranked at the top and the other alternatives are ranked as in the other three orderings). By Lemma 3, F^M works well on $\hat{\mathfrak{R}}'$ so we can assume that F does, too (otherwise, we are done). Given generic decisiveness and that $\boldsymbol{R}^{\circ\circ}$ is regular, $F(\boldsymbol{R}^{\circ\circ}, \{x,y,z\})$ is a singleton. We cannot have $F(\boldsymbol{R}^{\circ\circ}, \{x,y,z\}) = y$, because z Pareto dominates y.

If $F(\boldsymbol{R}^{\circ\circ}, \{x, y, z\}) = x$, then $F(\boldsymbol{R}^{\circ\circ}, \{x, y\}) = x$ by IIA. But anonymity and (21) yield,

$$F(\boldsymbol{R}^{\circ\circ}, \{x, y\}) = F(\boldsymbol{R}^{\circ}, \{x, y\}) = y, \tag{22}$$

a contradiction. Thus, we must have $F(\boldsymbol{R}^{\circ\circ}, \{x, y, z\}) = z$, implying from IIA that $F(\boldsymbol{R}^{\circ\circ}, \{x, z\}) = z$. Then neutrality in turn implies that

$$F(\hat{\boldsymbol{R}}^{\circ\circ}, \{x, z\}) = x, \tag{23}$$

where $\hat{\boldsymbol{R}}^{\circ\circ}$ is a profile on $\hat{\mathfrak{R}}'$ such that

$$\hat{\boldsymbol{R}}^{\circ\circ} = \begin{array}{ccc} \underline{[0, \alpha)} & \underline{[\alpha, 1-\alpha)} & \underline{[1-\alpha, 1]} \\ x & x & z \\ y & z & x \\ z & y & y \end{array} . \tag{24}$$

Next, take $\mathfrak{R}'$ to consist of the orderings in equation (24) together with $\begin{array}{c} x \\ y \\ z \end{array}$. Again, F^M works well on $\mathfrak{R}'$ and so we can assume that F does, too. By equation (23) by we can deduce, using neutrality, that

$$F(\boldsymbol{R}^{\circ\circ\circ}, \{x, y\}) = x$$

for profile $\boldsymbol{R}^{\circ\circ\circ}$ on $\mathfrak{R}'$ such that

$$\boldsymbol{R}^{\circ\circ\circ} = \begin{array}{ccc} \underline{[0, \alpha)} & \underline{[\alpha, 1-\alpha)} & \underline{[1-\alpha, 1]} \\ x & x & y \\ z & y & x \\ y & z & z \end{array} ,$$

which, from anonymity, contradicts equation (22). Hence, F does not work well on $\mathfrak{R}'$ after all.