

# Majority Judgment vs. Majority Rule \*

Michel Balinski

CNRS, CREST, École Polytechnique, France, mlbalinski@gmail.com

and

Rida Laraki\*

CNRS, University of Paris Dauphine-PSL, France, rida.laraki@dauphine.fr<sup>†</sup>  
and University of Liverpool, United Kingdom

June 6, 2019

## Abstract

The validity of majority rule in an election with but two candidates—and so also of Condorcet consistency—is challenged. Axioms based on evaluating candidates—paralleling those of K. O. May characterizing majority rule for two candidates based on comparing candidates—lead to another method, majority judgment, that is unique in agreeing with the majority rule on pairs of “polarized” candidates. It is a practical method that accommodates any number of candidates, avoids both the Condorcet and Arrow paradoxes, and best resists strategic manipulation.

Key words: majority, measuring, ranking, electing, Condorcet consistency, domination paradox, tyranny of majority, intensity problem, majority-gauge, strategy-proofness, polarization.

## 1 Introduction

Elections *measure*. Voters express themselves, a rule amalgamates them, the candidates’ scores—measures of support—determine their order of finish, and the winner.

Traditional methods ask voters to express themselves by *comparing* them:

- ticking one candidate at most (majority rule, first-past-the-post) or several candidates (approval voting), candidates’ total ticks ranking them;
- rank-ordering the candidates (Borda rule, alternative vote or preferential voting, Lull’s rule, Dasgupta-Maskin method, ...), differently derived numerical scores ranking them.

However, no traditional method measures the support of *one* candidate.<sup>1</sup> Instead, the theory of voting (or of social choice) elevates to a basic distinguishing axiom the faith that in an election between *two* candidates majority rule is the only proper rule. Every traditional method of voting tries to generalize majority rule to more candidates and reduces to it when there are only two candidates. Yet majority rule decides unambiguously only when there are two candidates: it says nothing when there is only one and its generalizations to three or more are incoherent.

Using the majority rule to choose one of two candidates is widely accepted as infallible: since infancy who in this world has not participated in raising their hand to reach a collective decision on two alternatives? A. de Tocqueville believed “It is the very essence of democratic governments that the dominance of the majority be absolute; for other than the majority, in democracies, there is nothing that resists” ([52], p. 379)<sup>2</sup>. Judging

---

\*The authors are deeply indebted to Jack Nagel for very helpful comments and suggestions, particularly with regard to Dahl’s intensity problem.

<sup>†</sup>On February 4, 2019, my co-author Michel Balinski, passed away at 85 years old. He was not only a great researcher, but also a kind, considerate gentleman and an attentive and faithful friend. I will miss him.

<sup>1</sup>A significant number of elections, including U.S. congressional elections, have but one candidate.

<sup>2</sup>Our translation of: “Il est de l’essence même des gouvernements démocratiques que l’empire de la majorité y soit absolu; car en dehors de la majorité, dans les démocraties, il n’y a rien qui résiste.”

from W. Sadurski’s assertion—“The legitimating force of the majority rule is so pervasive that we often do not notice it and rarely do we question it: We usually take it for granted” ([49], p. 39)—that conviction seems ever firmer today. Students of social choice will high universally accept majority rule for choosing between two alternatives, and much of the literature takes *Condorcet consistency*—that a candidate who defeats each of the others separately in majority votes must be the winner (the *Condorcet-winner*)—to be either axiomatic or at least a most desirable property.

Why this universal acceptance of majority rule for *two* candidates? First, the habit of centuries; second, K.O. May’s [39] axiomatic characterization; third, the fact that it is strategy-proof or incentive compatible; fourth, the Condorcet jury theorem [20].

Regrettably, as will be shown, the majority rule can easily go wrong when voting on but *two* candidates. Moreover, asking voters to compare candidates when there are three or more inevitably invites the Condorcet and Arrow paradoxes. The first shows that it is possible for a method to yield a non-transitive order of the candidates, so no winner [20]. It occurs, e.g., in elections [35], in figure skating ([4] pp. 139-146, [8]), in wine-tasting [6]. The second shows that the presence or absence of a (often minor) candidate can change the final outcome among the others. It occurs frequently, sometimes with dramatic global consequences, e.g., the election of George W. Bush in 2000 because of the candidacy of Ralph Nader in Florida; the election of Nicolas Sarkozy in 2007 although all the evidence shows François Bayrou, eliminated in the first-round, was the Condorcet-winner.

A recent econometric study of French parliamentary (1978-2012) and local (2011, 2015) elections [46] shows that in 19.2% of the 577 elections the presence of a third minor candidate in run-offs causes a change in the winner and concludes “that a large fraction of voters are ... ‘expressive’ and vote for their favorite candidate at the cost of causing the defeat of their second-best choice” ([46], p. 42).

How are these paradoxes to be avoided? Some implicitly accept one or the other (e.g., Borda’s method, Condorcet’s method). Others believe voters’ preferences are governed by some inherent restrictive property. For example, Dasgupta and Maskin [23, 22] appeal to the idea that voters’ preferences exhibit regularities—they are “single-peaked,” meaning candidates may be ordered on (say) a left/right political spectrum so that any voter’s preference peaks on some candidate and declines in both directions from her favorite; or they satisfy “limited agreement,” meaning that for every three candidates there is one that no voter ranks in the middle. Another possibility is the “single crossing” restriction (e.g., [12, 47]): it posits that both candidates and voters may be aligned on a (say) left/right spectrum and the more a voter is to the right the more she will prefer a candidate to the right.

Such restrictions could in theory reflect sincere patterns of preference in some situations; nevertheless in *all* such cases voters could well cast strategic ballots of a very different stripe. Moreover, all the sets of ballots that we have studied show that *actual* ballots are in no sense restricted. There are many examples. (1) An approval voting experiment was conducted in parallel with the first-round of the 2002 French presidential election that had 16 candidates. 2,587 voters participated and cast 813 different ballots: had the preferences been single-peaked there could have been at most 137 sincere ballots ([4] p.117). (2) A Social Choice and Welfare Society presidential election included an experiment that asked voters to give their preferences among the three candidates: they failed to satisfy any of the above restrictions [17, 48]. (3) A recent study of voting in Switzerland uses principal component analysis to show that at least three dimensions are necessary to map voters’ preferences, so the one-dimensional single-peaked condition cannot be met [26]. Real ballots—true opinions or strategic choices—eschew ideological divides.

All of this, we conclude, shows that the domain of voters’ preferences is in real life unrestricted. An entirely different approach is needed. This has motivated the development of majority judgment [3, 4, 8] based on a different paradigm: instead of comparing candidates, voters are explicitly charged with a solemn task of expressing their opinions precisely by evaluating the merit of every candidate in an ordinal scale of measurement or language of grades.

Thus, for example, the ballot in a presidential election could ask:

“Having taken into account all relevant considerations, I judge, in conscience, that as President of France, each of the following candidates would be:”

The language of grades constituting the possible answers may contain any number of grades<sup>3</sup> though in elections with many voters five to seven have proven to be good choices.

A scale that has been used is:

*Outstanding, Excellent, Very Good, Good, Fair, Poor, To Reject;*

<sup>3</sup>Majority judgment theory [4] put no restrictions on the size or the nature of the scale, which can be numerical or verbal, finite or infinite. Methods based on a numerical scale (e.g. point-summing methods) are characterized in Chapter 17.

The method then specifies that majorities determine the electorate’s evaluation of each candidate and the ranking between every pair of candidates—necessarily transitive—with the first-placed among them the winner.

Majority judgment has been criticized by the proponents of the traditional paradigm in articles, reviews, and talks [13, 15, 24, 27, 28].. Most objections have been answered in chapter 16 of [4], see also [2]. New arguments to the no-show paradox charge are detailed in a companion paper [9]. The present article gives a completely new argument for majority judgment that responds to the objection that it does not agree with the majority rule on two candidates, i.e., that it is not Condorcet-consistent.

This paper first gives a new description of majority judgment that shows it naturally emerges from a majority principle when candidates are assigned grades. Then, it shows that majority rule on two candidates is not incontestable, it is open to serious error because it admits the “domination paradox”: it may elect a candidate whose grades are dominated by another candidate’s grades.

The infallibility of majority rule on two candidates has been challenged before. R. A. Dahl charged: “By making ‘most preferred’ equivalent to ‘preferred by most’ we deliberately bypassed a crucial problem: What if the minority prefers its alternative much more passionately than the majority prefers a contrary alternative? Does the majority principle still make sense? This is the *problem of intensity*, ... [W]ould it be possible to construct rules so that an apathetic majority only slightly preferring its alternative could not override a minority strongly preferring its alternative?” ([21], 90-92).

Finally, starting from the basic principles that justify majority rule (May’s axioms), majority judgment is characterized as the unique method based on evaluations that satisfies those same principles, avoids the Condorcet and Arrow paradoxes, and coincides with majority rule on two candidates when the electorate is “polarized,” that is, when the higher (the lower) a voter evaluates one candidate the lower (the higher) she evaluates the other, so there can be no consensus. Majority rule does not violate domination on polarized candidates, and since it is incentive compatible on that domain, so is majority judgment, and that is when voters are most tempted to manipulate.

## 2 Evaluating vs Comparing

Except for political elections, the practice in virtually every instance that ranks entities is to evaluate each of them (see [4], chapters 7 and 8). The *Guide Michelin* uses stars to rate restaurants and hotels. Competitive diving, figure skating, and gymnastics use carefully defined number scales. Wine competitions use words: *Excellent*, *Very Good*, *Good*, *Passable*, *Inadequate*, *Mediocre*, *Bad*. Students are graded by letters, numbers, or phrases. Pain uses sentences to describe each element of a scale that is numbered from 0 (“Pain free”) to 10 (“Unconscious. Pain makes you pass out.”), a 7 defined by “Makes it difficult to concentrate, interferes with sleep. You can still function with effort. Strong painkillers are only partially effective.”

In the political sphere polls, seeking more probing information about voter opinion, also ask more. Thus a Harris poll: “... [H]ow would you rate the overall job that President Barack Obama is doing on the economy?” Among the answers spanning 2009 to 2014 were those given in Table 1a.

	<i>Excellent</i>	<i>Pretty good</i>	<i>Only fair</i>	<i>Poor</i>
March 2009	13%	34%	30%	23%
March 2011	5%	28%	29%	38%
March 2013	6%	27%	26%	41%

Table 1a. Measures evaluating the performance of Obama on the economy [33].

Recent Pew Research Center national political surveys ([44], pp. 11-12) show that polling professionals believe asking registered voters to evaluate candidates is a natural way of eliciting opinions. Pew conducted four such surveys in mid-January, late March, mid-August, and late October 2016 (just before the election on November 8). One of the questions asked in each survey was:

“Regardless of who you currently support, I’d like to know what kind of president you think each of the following would be if elected in November 2016. ... [D]o you think (he/she) would be a great, good, average, poor or terrible president?”

The results are given in Table 1b. Majority rule, used in the primaries of both parties, led to the nominations of Clinton evaluated *Poor* or worse by 47% and Trump evaluated *Poor* or worse by 62%: the least respected major candidates in more than a half century according to many assessments.

	<i>Great</i>	<i>Good</i>	<i>Average</i>	<i>Poor</i>	<i>Terrible</i>	<i>Never heard of</i>
John Kasich	5%	28%	39%	13%	7%	9%
Bernie Sanders	10%	26%	26%	15%	21%	3%
Ted Cruz	7%	22%	21%	17%	19%	4%
Hillary Clinton	11%	22%	20%	16%	30%	1%
Donald Trump	10%	16%	12%	15%	44%	3%

Table 1b. Pew Research Center poll results, March 17-27, 2016 [44].

The distributions of Clinton’s and Trump’s grades in answer to the four 2016 Pew polls are given in Table 1c. In contrast with many majority rule polling results, the evaluations of Clinton and Trump are stable throughout the entire year despite the ups and downs—the many revelations and accusations—of the campaign. In this election, majority rule fails because many voters were forced by the system to choose between two candidates they disliked and so some did not participate, some decided for whom to vote depending on the last news, others wished to reject politics and majority rule left them with only one way to do so: vote for Trump.

		<i>Great</i>	<i>Good</i>	<i>Average</i>	<i>Poor</i>	<i>Terrible</i>	<i>Never heard of</i>
Clinton:	January	11%	24%	18%	16%	28%	2%
	March	11%	22%	20%	16%	30%	1%
	August	11%	20%	22%	12%	33%	2%
	October	8%	27%	20%	11%	34%	1%
Trump:	January	11%	20%	12%	14%	38%	5%
	March	10%	16%	12%	15%	44%	3%
	August	9%	18%	15%	12%	43%	3%
	October	9%	17%	16%	11%	44%	2%

Table 1c. Pew Research Center poll results, 2016 [45].

A similar situation was observed in the second round of the French presidential election of 2017 when majority rule asked voters to choose between Emmanuel Macron (En Marche, a centrist newly formed party) and Marine Le Pen (Front National, extreme right party). Compared to the participation of the first round—see Table 1D— we observe 1.5 million fewer voters, 5 times as many blank ballots, 4 times as many invalid ballots and so almost 5 million fewer valid votes. Why? Many voters refused to be counted as supporting either candidate, yet they may have wished to express a difference between Macron and Le Pen (two radically opposed candidates).<sup>4</sup>

	1st ROUND			2nd ROUND		
	Number	% Regis.	% Voters	Number	% Regis.	% Voters
Regis.	47 582 183			47 568 693		
Absten.	10 578 455	22.23%		12 101 366	25.44%	
Voters	37 003 728	77.77%		35 467 327	74.56%	
Blank	659 997	1.39%	1.78%	3 021 499	6.35%	8.52%
Inval.	289 337	0.61%	0.78%	1 064 225	2.24%	3.00%
Votes	36 054 394	75.77%	97.43%	31 381 603	65.97%	88.48%

Table 1D. 2nd vs 1st round results in the 2017 French presidential election [45].

One of the reasons that things go so wrong in those real examples is that majority rule on two candidates measures badly. Voters are charged with nothing other than to tick the name of at most one candidate or to abstain. They are not given the means to express their opinions.

All this implies, we believe, that it is not only natural to use grades to evaluate one, two, or more alternatives but necessary. To be able to measure the support a candidate enjoys, a voter must be given the means to express her opinions or “feelings.” To assure that voters are treated equally, voters must be confined to a set of expressions that is shared by all. To allow for meaningful gradations—different shades ranging from very positive, through mediocre, to very negative—the gradations must faithfully represent the possible likes and dislikes. Such finite, ordered sets of evaluations are common and accepted in every day life. Call it a *scale* or *common language of grades*  $\Lambda$  linearly ordered by  $\succ$ .

<sup>4</sup>Except this election, participation is much higher in a second round presidential election in France compared to the first round. For example, in 2002, there were 2.5 million more valid votes in the second round compared to the first, while in 2017 there were 5 million less valid votes!

An electorate’s *opinion profile* on a candidate is the set of her, his, or its grades  $\alpha = (\alpha_1, \dots, \alpha_n)$ , where  $\alpha_j \in \Lambda$  is voter  $j$ ’s evaluation of the candidate.

Since voters must have equal voices, only the grades can count: which voter gave what grade should have no impact on the electorate’s global measure of a candidate. The number of times each grade occurs or their percentages (as in Tables 1a, 1b and 1c) is called the candidates’s *merit profile* (to distinguish it from an *opinion profile* that specifies the grade given the candidate by each of the judges). A candidate’s merit profile will always be written from the highest grades on the left down to the lowest on the right.

## 2.1 A new description of majority judgment

In this section, we show why majority judgment naturally emerges from the majority principle applied to grades instead of preference orders (leading to majority rule).

What is the electorate’s *majority* opinion of a candidate with grades  $\alpha = (\alpha_1, \dots, \alpha_n)$ ? An example best conveys the basic idea. In the spring of 2015 majority judgment was used by a jury of six ( $J_1$  to  $J_6$ ) at LAMSADE, Université Paris-Dauphine to rank six students ( $A$  to  $F$ ) seeking fellowships to prepare Ph.D dissertations. The jury agreed their solemn task was to evaluate the students and chose the language of grades

*Excellent, Very Good, Good, Passable, Insufficient.*

The opinion profile of candidate  $C$  was

	$J_1$	$J_2$	$J_3$	$J_4$	$J_5$	$J_6$
$C$ :	<i>Passable</i>	<i>Excellent</i>	<i>Good</i>	<i>V. Good</i>	<i>V. Good</i>	<i>Excellent</i>

Since voters or judges must have equal voices, only the grades count, not which voter or judge gave what grade. Accordingly the number of times each grade occurs or their percentages is called the candidate’s *merit profile*, always written from the highest grades on the left down to the lowest on the right.  $C$ ’s merit profile was

$C$ :	<i>Excellent</i>	<i>Excellent</i>	<i>V. Good</i>	$\updownarrow$	<i>V. Good</i>	<i>Good</i>	<i>Passable</i>
-------	------------------	------------------	----------------	----------------	----------------	-------------	-----------------

The middle of  $C$ ’s grades in the merit profile is indicated by a two-sided arrow.

There is a majority of  $\frac{6}{6}$ —unanimity—for  $C$ ’s grade to be at most *Excellent* and at least *Passable*, or for [*Excellent*, *Passable*]; a majority of  $\frac{5}{6}$  for  $C$ ’s grade to be at most *Excellent* and at least *Good*, or for [*Excellent*, *Good*]; and a majority of  $\frac{4}{6}$  for  $C$ ’s grade to be at most *Very Good* and at least *Very Good*, or for [*Very Good*, *Very Good*]. The closer the two—equally distant from the middle—are to the middle, the closer are their values, so the more accurate is the majority decision. When the two are equal it is the “majority-grade” and it suffices to specify one grade. If  $n$  is odd there is certain to be an absolute majority for a single grade; if  $n$  is even and the middlemost grades are different (very rare in a large electorate) there is a majority consensus for two grades.

In general, letting  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  be a candidate  $A$ ’s set of  $n$  grades written from highest to lowest,  $\alpha_i \succeq \alpha_{i+1}$  for all  $i$ , there is a majority of (at least)  $\frac{n-k+1}{n}$  for  $A$ ’s grade to be at most  $\alpha_k$  and at least  $\alpha_{n-k+1}$ , for all  $1 \leq k \leq (n+1)/2$ . Call this the  $(\frac{n-k+1}{n})$ -majority for  $[\alpha_k, \alpha_{n-k+1}]$ . When  $k > h$  the two grades of the  $(\frac{n-k+1}{n})$ -majority are closer together than (or the same as) those of the  $(\frac{n-h+1}{n})$ -majority: they are *more accurate*.

A measure of  $A$ ’s global merit is the most accurate possible majority decision on  $A$ ’s grades.  $C$ ’s majority-grade is *Very Good*, Obama’s majority-grade in each of the evaluations of his performance on the economy (Table 1a) is *Only fair*.

How does a majority of an electorate rank candidates having sets of grades? For example, how does it rank two distributions of grades of Obama’s performance at different dates (Table 1a)? Had the March 2011 or 2013 distributions been identical to that on March 2009 the electorate would have judged the performances to be the same. In fact, Obama’s March 2009 evaluations “dominate” those of the same month in 2011 and 2013 and so the electorate clearly ranks it highest; but it is not clear how to compare the evaluations in 2011 and 2013.

In general, a candidate  $A$ ’s merit profile  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  *dominates*  $B$ ’s merit profile  $\beta = (\beta_1, \beta_2, \dots, \beta_n)$  (both written from highest to lowest) when  $\alpha_i \succeq \beta_i$  for all  $i$  and  $\alpha_k \succ \beta_k$  for at least one  $k$  (equivalently, when  $A$  has at least as many of the highest grade as  $B$ , at least as many of the two highest grades, ..., at least as many of the  $k$  highest grades for all  $k$ , and at least one “at least” is “more”).<sup>5</sup> Any reasonable method of ranking should *respect domination*: namely, evaluate one candidate above another when that candidate’s grades dominate the other’s. Surprisingly, some methods do not (as will be seen).

<sup>5</sup>In other contexts domination is called first-order stochastic dominance.

With  $m$  candidates the basic input is an electorate's *opinion profile*: it gives the grades assigned to every candidate by each voter and may be represented as a matrix  $\alpha = (\alpha_{ij})$  of  $m$  rows (one for each candidate) and  $n$  columns (one for each voter),  $\alpha_{ij}$  the grade assigned to candidate  $i$  by voter  $j$ . Table 2a gives the LAMSADE Jury's opinion profile. The *preference profile* of the traditional theory—voters' rank-orderings of the candidates—may be deduced from the opinion profile whenever the language of grades is sufficient for a voter to distinguish between any two candidates when he evaluates their merit differently. Thus, for example,  $J_1$ 's preferences are  $A \approx B \succ D \approx F \succ E \succ C$ . Note that no judge used all five grades even though there were six candidates.

	$J_1$	$J_2$	$J_3$	$J_4$	$J_5$	$J_6$
A:	<i>Excellent</i>	<i>Excellent</i>	<i>V. Good</i>	<i>Excellent</i>	<i>Excellent</i>	<i>Excellent</i>
B:	<i>Excellent</i>	<i>V. Good</i>	<i>V. Good</i>	<i>V. Good</i>	<i>Good</i>	<i>V. Good</i>
C:	<i>Passable</i>	<i>Excellent</i>	<i>Good</i>	<i>V. Good</i>	<i>V. Good</i>	<i>Excellent</i>
D:	<i>V. Good</i>	<i>Good</i>	<i>Passable</i>	<i>Good</i>	<i>Good</i>	<i>Good</i>
E:	<i>Good</i>	<i>Passable</i>	<i>V. Good</i>	<i>Good</i>	<i>Good</i>	<i>Good</i>
F:	<i>V. Good</i>	<i>Passable</i>	<i>Insufficient</i>	<i>Passable</i>	<i>Passable</i>	<i>Good</i>

Table 2a. Opinion profile, LAMSADE Jury.<sup>6</sup>

To see how majority judgment (MJ) ranks the candidates of the LAMSADE Jury consider the corresponding merit profile given in two equivalent forms: extensively (Table 2b) and by counts of grades (Table 2c).

A:	<i>Excellent</i>	<i>Excellent</i>	<b><i>Excellent</i></b>	<b><i>Excellent</i></b>	<i>Excellent</i>	<i>V. Good</i>
B:	<i>Excellent</i>	<i>V. Good</i>	<b><i>V. Good</i></b>	<b><i>V. Good</i></b>	<i>V. Good</i>	<i>Good</i>
C:	<i>Excellent</i>	<i>Excellent</i>	<b><i>V. Good</i></b>	<b><i>V. Good</i></b>	<i>Good</i>	<i>Passable</i>
D:	<i>V. Good</i>	<i>Good</i>	<b><i>Good</i></b>	<b><i>Good</i></b>	<i>Good</i>	<i>Passable</i>
E:	<i>V. Good</i>	<i>Good</i>	<b><i>Good</i></b>	<b><i>Good</i></b>	<i>Good</i>	<i>Passable</i>
F:	<i>V. Good</i>	<i>Good</i>	<b><i>Passable</i></b>	<b><i>Passable</i></b>	<i>Passable</i>	<i>Insufficient</i>

Table 2b. Merit profile (extensive), LAMSADE Jury.

The  $\frac{4}{6}$ -majorities are indicated in bold in Tables 2b,c.  $A$ 's  $\frac{4}{6}$ -majority dominates all the others, so  $A$  is the MJ-winner;  $B$ 's and  $C$ 's dominate the remaining candidates, so are next in the MJ-ranking; and  $D$  and  $E$ —tied in the MJ-ranking since their sets of grades are identical—follow in the MJ-ranking since their  $\frac{4}{6}$ -majorities dominate  $F$ 's. How are  $B$  and  $C$  to be compared (Table 2d)?

	<i>Excellent</i>	<i>Very Good</i>	<i>Good</i>	<i>Passable</i>	<i>Insufficient</i>
A:	<b>5</b>	1			
B:	1	<b>4</b>	1		
C:	2	<b>2</b>	1	1	
D:		1	<b>4</b>	1	
E:		1	<b>4</b>	1	
F:		1	1	<b>3</b>	1

Table 2c. Merit profile (counts), LAMSADE Jury.

Their  $\frac{4}{6}$ -majorities are identical but their  $\frac{5}{6}$ -majorities (indicated by square brackets in Table 2d) differ:  $B$ 's is for [*Very Good*, *Very Good*] and  $C$ 's for [*Excellent*, *Good*]. Since neither pair of grades dominates the other and there is more consensus<sup>7</sup> for  $B$ 's grades than for  $C$ 's, MJ ranks  $B$  above  $C$ . Thus the MJ ranking is  $A \succ_{MJ} B \succ_{MJ} C \succ_{MJ} D \approx_{MJ} E \succ_{MJ} F$ .

B:	<i>Excellent</i>	[ <b><i>V. Good</i></b>	<b><i>V. Good</i></b>	↕	<b><i>V. Good</i></b>	<b><i>V. Good</i></b> ]	<i>Good</i>
C:	<i>Excellent</i>	[ <b><i>Excellent</i></b>	<b><i>V. Good</i></b>	↕	<b><i>V. Good</i></b>	<b><i>Good</i></b> ]	<i>Passable</i>

Table 2d. Merit profile,  $B$  and  $C$ , LAMSADE Jury.

In general, if  $A$ 's grades are  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  and  $B$ 's  $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ , both written from highest to lowest, suppose the most accurate majority where the candidates differ is the  $(\frac{n-k+1}{n})$ -majority for  $[\bar{\alpha}, \underline{\alpha}] \neq [\bar{\beta}, \underline{\beta}]$ . Thus for  $A$ ,  $\bar{\alpha} = \alpha_k$  and  $\underline{\alpha} = \alpha_{n-k+1}$ , similarly for  $B$ , and  $\alpha_j = \beta_j$  for  $k < j < n - k + 1$ . Comparing them,  $A$ 's *middlemost block* of grades is  $[\alpha_k, \alpha_{k+1}, \dots, \alpha_{n-k}, \alpha_{n-k+1}]$  and  $B$ 's is  $[\beta_k, \beta_{k+1}, \dots, \beta_{n-k}, \beta_{n-k+1}]$ .

<sup>6</sup>The order of the students is chosen to coincide with the MJ-ranking for clarity.

<sup>7</sup>In other contexts this is called second-order stochastic dominance (favoring less "risk.") It is related to Hammond's equity principle [32] and implied by the Rawlsian criterion,  $B$ 's minimum being above  $C$ 's.

The *majority-ranking*  $\succeq_{MJ}$  ranks  $A$  above  $B$  when (a)  $A$ 's middlemost block dominates  $B$ 's or (b)  $A$ 's middlemost block is more consensual than  $B$ 's:

$$A \succ_{MJ} B \text{ when } \begin{cases} \text{(a) } \bar{\alpha} \succeq \bar{\beta} \text{ and } \underline{\alpha} \succeq \underline{\beta}, \text{ with at least one } \succeq \text{ strict, or} \\ \text{(b) } \bar{\beta} \succ \bar{\alpha} \succeq \underline{\alpha} \succ \underline{\beta}; \end{cases} \quad (1)$$

otherwise, their sets of grades are identical and  $A \approx_{MJ} B$ . The most accurate majority is either for a single grade called the *majority-grade*, or, when  $n$  is even, it may be for two grades (very rare in a large electorate).

## 2.2 Majority judgment with many voters

When there are many voters simpler arithmetic is almost always sufficient to determine the MJ ranking. This is due to two facts: the most accurate majority decision concerning a candidate is generically a single grade—the *majority-grade*—and (1b) almost never occurs, so it suffices to detect when a difference in grades first occurs according to (1a). An example shows how it works.

Terra Nova, a Parisian think-tank, sponsored a national presidential poll carried out by OpinionWay April 12-16, 2012 (just before the first-round of the election on April 22) to compare MJ with other methods. 993 participants voted with MJ and also according to usual practice—first-past-the-post (FPP) among all ten candidates, followed by a MR run-off between every pair of the expected five leaders in the first-round. Since the results of FPP varied slightly from the actual national percentages on election day (up to 5%) a set of 737 ballots was found for which those tallies are closely matched and are presented here.

To begin consider one candidate's merit profile ("H" for Hollande). 50% of the grades are to the left of the middle, 50% are to the right:

	<i>Outstanding</i>	<i>Excellent</i>	<i>Very Good</i>	<i>Good</i>	$\updownarrow$	<i>Good</i>	<i>Fair</i>	<i>Poor</i>	<i>To Reject</i>
H	12.48%	16.15%	16.42%	<b>4.95%</b>	$\updownarrow$	<b>6.72%</b>	14.79%	14.25%	14.24%

Hollande's majority-grade is *Good* because there is a  $(50 + \epsilon)\%$ -majority<sup>8</sup> for  $[Good, Good]$  for every  $\epsilon$ ,  $0 < \epsilon \leq 4.95$ . Similarly, there is a  $(54.95 + \epsilon)\%$ -majority for  $[Very Good, Good]$  for every  $\epsilon$ ,  $0 < \epsilon \leq 1.77$ , and a  $(56.72 + \epsilon)\%$ -majority for  $[Very Good, Fair]$  for every  $\epsilon$ ,  $0 < \epsilon \leq 8.07$ . The  $x\%$ -majority decision—in general, a pair of grades—may be found for any  $x\% > 50\%$ .

The MJ-ranking with many voters is determined in exactly the same manner as when there are few: the most accurate majority where two candidates differ decides. Compare, for example, Hollande (H) and Bayrou (B):

	<i>Outstanding</i>	<i>Excellent</i>	<i>Very Good</i>	<i>Good</i>	$\updownarrow$	<i>Good</i>	<i>Fair</i>	<i>Poor</i>	<i>To Reject</i>
H	12.48%	16.15%	16.42%	<b>4.95%</b>	$\updownarrow$	<b>6.72%</b>	14.79%	14.25%	14.24%
B	2.58%	9.77%	21.71%	<b>15.94%</b>	$\updownarrow$	<b>9.30%</b>	20.08%	11.94%	8.69%

Both have  $(50 + \epsilon)\%$ -majorities for  $[Good, Good]$  for every  $\epsilon$ ,  $0 < \epsilon \leq 4.95$ , so both have the majority-grade *Good*. But for  $0 < \epsilon \leq 1.77$  Hollande has a  $(54.95 + \epsilon)\%$ -majority for  $[Very Good, Good]$  whereas Bayrou has a  $(54.95 + \epsilon)\%$ -majority for  $[Good, Good]$ . Since Hollande's middlemost block dominates Bayrou's, MJ ranks Hollande above Bayrou. This happens because  $4.95 < \min\{6.72, 15.94, 9.30\}$ . Had the smallest of these four numbers 4.95 been Hollande's but to the right of the middle, his  $(54.95 + \epsilon)\%$ -majority would have been for  $[Good, Fair]$  whereas Bayrou's  $(54.95 + \epsilon)\%$ -majority would have remained  $[Good, Good]$ , putting Bayrou ahead of Hollande. Finding the smallest of these four numbers is the same as finding the highest percentage of each candidate's grades strictly above and strictly below their majority-grades: if that highest is the percentage above the majority-grade it puts its candidate ahead, if that highest is the percentage below it puts its candidate behind. The rule has another natural interpretation: of the four sets of voters who disagree with the majority-grades the largest tips the scales.

The general rule for ranking two candidates with many voters makes the generic assumption that there is a  $(50 + \epsilon)\%$ -majority,  $\epsilon > 0$ , for each candidate's majority-grade. Let  $p_A$  be the percentage of  $A$ 's grades strictly above her majority-grade  $\alpha_A$  and  $q_A$  the percentage of  $A$ 's grades strictly below  $\alpha_A$ .  $A$ 's *majority-gauge* (MG) is  $(p_A, \alpha_A, q_A)$ . The *majority-gauge rule*  $\succ_{MG}$  ranks  $A$  above  $B$  when

$$A \succ_{MG} B \text{ when } \begin{cases} \alpha_A \succ \alpha_B \text{ or,} \\ \alpha_A = \alpha_B \text{ and } p_A > \max\{q_A, p_B, q_B\} \text{ or,} \\ \alpha_A = \alpha_B \text{ and } q_B > \max\{p_A, q_A, p_B\}. \end{cases} \quad (2)$$

<sup>8</sup> $\epsilon$  may be thought of as one grade.

The MG-rule is generically decisive, i.e., there are (unique) majority-grades  $\alpha_A$  and  $\alpha_B$ , and a unique maximum among the four numbers  $p_A, q_A, p_B$  and  $q_B$ . This is as certain as for the majority rule to be decisive. The majority-grade may be endowed with a sign,  $\alpha_A+$  when  $p_A > q_A$ , otherwise  $\alpha_A-$ . When the MG-rule is decisive its ranking is identical to that of the MJ-rule (specified in (1)) by construction.

The full merit profile of the French 2012 presidential poll is given in Table 3a. The MJ(=MG)-ranking is given in Table 3b together with the first-past-the-post (FPP) ranking to show the marked differences between them.

	<i>Outstanding</i>	<i>Excellent</i>	<i>Very Good</i>	<i>Good</i>	<i>Fair</i>	<i>Poor</i>	<i>To Reject</i>
F. Hollande	12.48%	16.15%	16.42%	11.67%	14.79%	14.25%	14.24%
F. Bayrou	2.58%	9.77%	21.71%	25.24%	20.08%	11.94%	08.69%
N. Sarkozy	9.63%	12.35%	16.28%	10.99%	11.13%	7.87%	31.75%
J.-L. Mélenchon	5.43%	9.50%	12.89%	14.65%	17.10%	15.06%	25.37%
N. Dupont-Aignan	0.54%	2.58%	5.97%	11.26%	20.22%	25.51%	33.92%
E. Joly	0.81%	2.99%	6.51%	11.80%	14.65%	24.69%	38.53%
P. Poutou	0.14%	1.36%	4.48%	7.73%	12.48%	28.09%	45.73%
M. Le Pen	5.97%	7.33%	9.50%	9.36%	13.98%	6.24%	47.63%
N. Arthaud	0.00%	1.36%	3.80%	6.51%	13.16%	25.24%	49.93%
J. Cheminade	0.41%	0.81%	2.44%	5.83%	11.67%	26.87%	51.97%

Table 3a. Merit profile,<sup>9</sup> 2012 French presidential poll (737 ballots) [7].

	Above Maj-grade $p$	Maj-grade $\alpha \pm$	Below Maj-grade $q$	FPP
F. Hollande	<b>45.05%</b>	<i>Good+</i>	43.28%	28.7%
F. Bayrou	34.06%	<i>Good-</i>	<b>40.71%</b>	9.1%
N. Sarkozy	<b>49.25%</b>	<i>Fair+</i>	39.62%	27.3%
J.-L. Mélenchon	<b>42.47%</b>	<i>Fair+</i>	40.43%	11.0%
N. Dupont-Aignan	<b>40.57%</b>	<i>Poor+</i>	33.92%	1.5%
E. Joly	36.77%	<i>Poor-</i>	<b>38.53%</b>	2.3%
P. Poutou	26.19%	<i>Poor-</i>	<b>45.73%</b>	1.2%
M. Le Pen	46.13%	<i>Poor-</i>	<b>47.63%</b>	17.9%
N. Arthaud	24.83%	<i>Poor-</i>	<b>49.93%</b>	0.7%
J. Cheminade	48.03%	<i>To Reject</i>	—	0.4%

Table 3b. MJ and first-past-the-post rankings, 2012 French presidential poll (737 ballots) [7].

## 2.3 Point-summing methods

A *point-summing method*<sup>10</sup> chooses (ideally) an ordinal scale—words or descriptive phrases (but often undefined numbers)—and assigns to each a numerical grade, the better the evaluation the higher the number. There are, of course, infinitely many ways to assign such numbers to ordinal grades. Every voter evaluates each candidate in that scale and the candidates are ranked according to the sums or averages of their grades. A point-summing method clearly respects domination, but it harbors two major drawbacks.

The Danish educational system uses six grades with numbers attached to each: *Outstanding* 12, *Excellent* 10, *Good*, 7, *Fair* 4, *Adequate* 2 and *Inadequate* 0 [54]. Its numbers address a key issue of measurement theory [34]: the scale of grades must constitute an *interval scale* for sums or averages to be *meaningful*. An interval scale is one in which equal intervals have the same meaning; equivalently, for which an additional point anywhere in the scale—going from 3 to 4 or from 10 to 11—has the same significance. In practice—in grading divers, students, figure skaters, wines or pianists—when (say) the scale is multiples of  $\frac{1}{2}$  from a high of 10 to a low of 0, it is much more difficult and much rarer to go from 9 to  $9\frac{1}{2}$  than from  $4\frac{1}{2}$  to 5, so adding or averaging such scores is—in the language of measurement theory—meaningless. The Danes specified a scale that they believed constitutes an interval scale (for an extended discussion of these points see [4], pp. 171-174, or [8]). “Range voting” is a point-summing method advocated on the web where voters assign a number between 0 and a 100 to each candidate, but the numbers are given no meaning other than that they contribute to a candidate’s total number of points and they do not constitute an interval scale.

A second major drawback of point-summing methods is their manipulability. Any voter who has not given the highest (respectively, the lowest) grade to a candidate can increase (can decrease) the candidate’s average

<sup>9</sup>The row sums may differ from 100% due to round-off errors.

<sup>10</sup>Point-summing methods are characterized in [4], chapter 17. Combining with Theorem 3 of this article, one concludes that they are the unique ranking methods that satisfy axiom 1\* and axioms 2-8 of this article, and avoids the no-show paradox.



grade, so it pays voters or judges to exaggerate up and down. A detailed analysis of an actual figure skating competition [8] shows that with point-summing every one of the nine judges could alone manipulate to achieve precisely the order-of-finish he prefers by changing his scores. A companion analysis of the same competition shows that with majority judgment the possibilities for manipulation are drastically curtailed.

Long time users of point-summing methods such as the Fédération Internationale de Natation (FINA) have recently tried to combat strategic manipulation. Divers must specify the dives they will perform, each of which has a known degree of difficulty expressed as a number. Judges assign a number grade to each dive from 10 to 0 in multiples of  $\frac{1}{2}$ : *Excellent* 10, *Very Good* 8.5-9.5, *Good* 7.0-8.0, *Satisfactory* 5.0-6.5, *Deficient* 2.5-4.5, *Unsatisfactory* 0.5-2.0, *Completely failed* 0 (the meanings of each are further elaborated). There are five or seven judges. The highest and lowest scores are eliminated when there are five judges and the two highest and two lowest are eliminated when there are seven judges. The sum of the remaining three scores is multiplied by the degree of difficulty to obtain the score of the dive. Competitions in skating and gymnastics have chosen similar methods. Had any of them gone a little further—eliminating only what must be to distinguish a difference—they would have used MJ: increasingly practical people choose methods approaching MJ.

## 2.4 Approval voting

Analyses, experiments, and uses of approval voting [16] have deliberately eschewed ascribing any meaning to *Approve* and *Disapprove*—except that *Approve* means giving one vote to a candidate and *Disapprove* means giving none—leaving it entirely to voters to decide how to try to express their opinions [53, 16]. Thus, for example, the Social Choice and Welfare Society’s ballot for electing its president had small boxes next to candidates’ names with the instructions: “You can vote for any number of candidates by ticking the appropriate boxes,” the number of ticks determining the candidates’ order of finish. In this description AV may be seen as a point-summing method where voters assign a 0 or 1 to each candidate and the electorate’s rank-order is determined by the candidates’ total sums of points.

Recently, however, that view has changed: “the idea of judging each and every candidate as acceptable or not is fundamentally different” from either voting for one candidate or ranking them ([14], pp.vii-viii). This implies a belief that voters are able to judge candidates in an ordinal scale of merit with two grades. With this paradigm approval voting becomes MJ with a language of two grades: “approval judgment.” For example, if *Approve* meant *Good* or better the AV results of the LAMSADE Jury would be those given in Table 2f.

Student:	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>
AV-score:	6	6	5	5	5	2
AV-ranking:	1 <sup>st</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	2 <sup>nd</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>

Table 2f. AV-scores and -ranking, *Approve* means *Good* or better, LAMSADE Jury.

When there are few voters AV’s two grades are not sufficient to distinguish the competitors. Further evidence shows two grades are too few even when there are many voters [4, 8, 9].

## 3 Majority rule characterized for two candidates

*Majority rule* (MR) in a field of two elects that candidate preferred to the other by a majority of the electorate. May proved [39] that the majority rule is the one rule that satisfies the following six simple properties in an election with two candidates. This theorem is considered to be a major argument in its favor [25].

**Axiom 1 (Based on comparing)** *A voter’s opinion is a preference for one candidate or indifference between them.*<sup>11</sup>

Thus the input is a *preference profile* that specifies the preference or indifference of each voter.

**Axiom 2 (Unrestricted domain)** *Voters’ opinions are unrestricted.*

**Axiom 3 (Anonymity)** *Interchanging the names of voters does not change the outcome.*

**Axiom 4 (Neutrality)** *Interchanging the names of candidates does not change the outcome.*

Anonymity stipulates equity among voters, neutrality the equitable treatment of candidates.

<sup>11</sup>In a footnote May had the wisdom to admit, “The realism of this condition may be questioned.”

**Axiom 5 (Monotonicity)** *If candidate  $A$  wins or is tied with his opponent and one or more voters change their preferences in favor of  $A$  then  $A$  wins.*

A voter's change in favor of  $A$  means changing from a preference for  $B$  to either indifference or a preference for  $A$ , or from indifference to a preference for  $A$ .

**Axiom 6 (Completeness)** *The rule guarantees an outcome: one of the two candidates wins or they are tied.*

**Theorem 1 (May [39])** *For  $n = 2$  candidates majority rule  $\succeq_{MR}$  is the unique method that satisfies Axioms 1 through 6.*

**Proof.** The argument is simple. That MR satisfies the axioms is obvious.

So suppose the method  $\succeq_M$  satisfies the axioms. Anonymity implies that only the numbers count: the number of voters  $n_A$  who prefer  $A$  to  $B$ , the number  $n_B$  that prefer  $B$  to  $A$ , and the number  $n_{AB}$  that are indifferent between  $A$  and  $B$ . Completeness guarantees there must be an outcome. (1) Suppose  $n_A = n_B$  and  $A \succ_M B$ . By neutrality switching the names results in  $B \succ_M A$ : but the new profile is identical to the original, a contradiction that shows  $A \approx_M B$  when  $n_A = n_B$ . (2) If  $n_A > n_B$  change the preferences of  $n_A - n_B$  voters who prefer  $A$  to  $B$  to indifferences to obtain a valid profile (by Axiom 2). With this profile  $A \approx_M B$ . Changing them back one at a time to the original profile proves  $A \succ_M B$  by monotonicity. ■

There are three further arguments in favor of MR for *two* candidates. First, its simplicity and familiarity. Second, its incentive compatibility: the optimal strategy of a voter who prefers one of the two candidates is to vote for that candidate [11]. Third, the Condorcet jury theorem.

In its simplest form, the jury theorem supposes that one of the two outcomes is correct and that each voter has an independent probability  $p > 50\%$  of voting for it, concludes that the greater the number of voters the more likely the majority rule makes the correct choice, and that furthermore, in the limit, it is certain to do so. In most votes between two alternatives, however, there is no “correct” choice or “correct” candidate: all opinions are valid judgments, disagreement is inherent to any democracy, and must be accepted. A mechanism that chooses the consensus is what is needed.

## 4 The domination paradox

The dangers of a “tyranny of the majority” have been discussed for centuries: the phrase was used by John Adams in 1788, Alexis de Tocqueville in 1835 and John Stuart Mill in 1859. Robert A. Dahl's now classic treatise [21] considers it at length in the context of what he calls “Madisonian Democracy” and poses the problem of intensity: “[J]ust as Madison believed that government should be constructed so as to prevent majorities from invading the natural rights of minorities, so a modern Madison [38] might argue that government should be designed to inhibit a relatively apathetic majority from cramming its policy down the throats of a relatively intense minority” ([21], p. 90).

Given voters' evaluations of candidates, MR for two candidates harbors a major drawback heretofore unrecognized. It admits the *domination paradox*: a candidate  $A$ 's evaluations dominate  $B$ 's evaluations but MR elects  $B$  not  $A$ .

Contrast the merit profiles of Hollande and Sarkozy in the national poll of the 2012 French presidential election (Table 4a). Hollande's grades very generously dominate Sarkozy's. But this merit profile could come from the opinion profile of Table 4b where Sarkozy is the MR-winner with 59.57% of the votes to Hollande's 26.19%, 14.24% rejecting both.

	<i>Outstanding</i>	<i>Excellent</i>	<i>Very Good</i>	<i>Good</i>	<i>Fair</i>	<i>Poor</i>	<i>To Reject</i>
Hollande	12.48%	16.15%	16.42%	11.67%	14.79%	14.25%	14.24%
Sarkozy	9.63%	12.35%	16.28%	10.99%	11.13%	7.87%	31.75%

Table 4a. Merit profile, Hollande-Sarkozy, 2012 French presidential election poll.

	9.63%	12.35%	11.67%	4.61%	10.18%	11.13%	14.24%
Hollande:	<i>Exc.</i>	<i>V.Good</i>	<i>Good</i>	<i>Fair</i>	<i>Fair</i>	<i>Poor</i>	<i>Rej.</i>
Sarkozy:	<i>Outs.</i>	<i>Exc.</i>	<i>V.Good</i>	<i>V.Good</i>	<i>Good</i>	<i>Fair</i>	<i>Rej.</i>
	0.81%	7.87%	3.80%	6.52%	4.07%	3.12%	
Hollande:	<i>Outs.</i>	<i>Outs.</i>	<i>Outs.</i>	<i>Exc.</i>	<i>V.Good</i>	<i>Poor</i>	
Sarkozy:	<i>Good</i>	<i>Poor</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	

Table 4b. Possible opinion profile, Hollande-Sarkozy (giving the merit profile of table 4a), national poll, 2012 French presidential election.

Those voters who rate Sarkozy above Hollande do so mildly (with small differences in grades, top of profile), whereas Holland is rated above Sarkozy intensely (with large differences in grades, bottom of the profile): this is a situation Dahl elicited in questioning the validity of MR. In the actual national vote—and in the poll—Hollande won by a bare majority of 51.6% to 48.4%, suggesting that the possibility for MR to err on two candidates is important and real. To elect a candidate whose evaluations are dominated by another is clearly unacceptable.

## 5 May's axioms for more than two candidates

Given a fixed scale of linearly ordered grades  $\Lambda$ , a *ranking problem* is defined by an electorate's opinion profile  $\Phi$ , an  $m$  by  $n$  matrix of grades when there are  $m$  candidates and  $n$  voters. A *method of ranking*  $\succeq_M$  is an asymmetric binary relation between all pairs of candidates.

With the grading model voters' inputs are grades. With the traditional model voters' inputs are comparisons. Individual rationality implies that a voter's preference is a rank order over all the candidates. It may be deduced from a voter's grades when the scale is *sufficient* to distinguish between any two candidates whenever the voter believes their merit to be different.

The following are May's axioms extended to any number of candidates.<sup>12</sup>

**Axiom 1 (Based on comparing)** *A voter's input is a rank-order of the candidates.*

**Axiom 2 (Unrestricted domain)** *Voters' opinions are unrestricted.*

**Axiom 3 (Anonymity)** *Interchanging the names of voters does not change the outcome.*

**Axiom 4 (Neutrality)** *Interchanging the names of candidates does not change the outcome.*

In the traditional model a voter's input *becomes better* for a candidate  $A$  if  $A$  rises in his rank-order. In the grading model a voter's input *becomes better* for  $A$  if  $A$  is given a higher grade.

**Axiom 5 (Monotonicity)** *If  $A \succeq_M B$  and one or more voters' inputs become better for  $A$  then  $A \succ_M B$ .*

**Axiom 6 (Completeness)** *For any two candidates either  $A \succeq_M B$  or  $A \preceq_M B$  (or both, implying  $A \approx_M B$ ).*

With more than two candidates the Condorcet and Arrow paradoxes must be excluded.

**Axiom 7 (Transitivity)** *If  $A \succeq_M B$  and  $B \succeq_M C$  then  $A \succeq_M C$ .*

**Axiom 8 (Independence of irrelevant alternatives (IIA))** *If  $A \succeq_M B$  then whatever other candidates are either dropped or adjoined  $A \succeq_M B$ .*

This is the Nash-Chernoff [19, 43] formulation of IIA defined for a variable number of candidates [43, 19], not Arrow's definition for a fixed number of candidates, and it implies Arrow's. It is this conception that is often violated in practice (e.g., elections, figure skating, wines).

**Theorem 2 (Arrow's impossibility [1])** *For  $n \geq 3$  candidates there is no method of ranking  $\succeq_M$  that satisfies Axioms 1 through 8.*

This is a much watered-down form of Arrow's theorem, based on more axioms—though all necessary in a democracy—and is very easily proven. It is stated to contrast the two models: comparing candidates versus evaluating them.

**Proof.** Take any two candidates  $A$  and  $B$ . By IIA it suffices to deal with them alone to decide who leads. Axioms 1 through 6 imply that the method  $\succeq_M$  must be MR. Since the domain is unrestricted Condorcet's paradox now shows MR violates transitivity. So there can be no method satisfying all the axioms. ■

Now replace Axiom 1 above by:

<sup>12</sup>The paper of Goodin and List [30] offers another possible extension of May's theorem to more than two candidates that characterizes plurality-rule and provides an argument for approval voting.

**Axiom 1\* (Based on measuring)** *A voter’s input is the grades given the candidates.*

Dahl, when he challenged the validity of MR on two candidates, proposed as a solution to use an ordinal “intensity scale” obtained “simply by reference to some observable response, such as a statement of one’s feelings ...” ([21], p. 101), and argued that it is meaningful to do so: “I think that the core of meaning is to be found in the assumption that the uniformities we observe in human beings must carry over, in part, to the unobservables like feeling and sensation” ([21], p. 100). This is precisely the role of our new axiom 1\* which also relates to a key problem raised by measurement theorists, the *faithful representation problem*: when measuring some attribute of a class of objects or events how to associate a scale “in such a way that the properties of the attribute are faithfully represented ...” ([34], p.1). Practice—in figure skating, wine tasting, diving, gymnastics, assessing pain, etc.—has spontaneously and naturally resolved it.

However, some social choice theorists continue to express the opinion that a scale of intensities or grades is inappropriate in elections. Why should intensities be valid—indeed, be necessary—and practical in judging competitions but not in elections? The validity of using intensities as inputs in voting as versus using rankings as inputs is not a matter of opinion or of mathematics: it is a practical, experimental issue. Repeated experiments and real uses of MJ ([4] pp. 9-16 and chapter 15; [31]; [8] pp. 496-497 and 504-509; and others referred to in this article) show that nuances in evaluations are as valid for candidates in elections as they are for figure skaters, divers or wines in competitions, though the criteria and scales of measures must be crafted for each individually. Participants have uniformly found it easy to assign grades and have done so quickly. Some practical-minded people are also persuaded that MJ should be used in elections. Terra Nova—“an independent progressive think tank whose goal is to produce and diffuse innovative political solutions in France and Europe”—has included majority judgment in its recommendations for reforming the presidential election system of France [51]. A non-profit association<sup>13</sup> has been created in France to promote majority judgment. Finally, a real use of MJ was initiated by several concerned French citizens in 2016—an engineer, a lawyer and a mathematician. Called LaPrimaire.org the idea was for citizens independent of any political parties to freely nominate a candidate for the 2017 French presidential election. The entire effort was carried out on the web, including campaigning and voting. The organizers considered the various known voting methods and chose to use MJ. An initial slate of some 200 candidates was whittled down to twelve. Then in a first vote each voter was asked to evaluate five prescribed candidates of the twelve on the scale *Excellent, Very Good, Good, Passable, Insufficient*. The assignment of five (of the twelve) candidates to voters was done randomly but in such manner that each candidate was evaluated by approximately the same number of voters. The twelve were then ranked by MJ and the five leaders designated to participate in the final campaign and vote [36]. The reason for not directly voting on all twelve was to incite voters to invest the time to make a careful comparative study of the proposals and approaches of each candidate: evaluating twelve candidates was seen as asking for too great an investment of time and effort. The final election among the five was conducted using MJ with the same scale; 32,685 persons voted; each candidate’s grades dominated those of the candidates lower in the MJ-ranking. The winner’s majority-grade was *Excellent*, the runner up’s was *Very Good*, all of the last three’s were *Good* (see [37] for a complete description). The results were well accepted and MJ was included in the electoral reforms proposed by the winner and others.

**Theorem 3** *For  $n \geq 1$  candidates there are an infinite number of methods of ranking  $\succeq_M$  that satisfy Axioms 1\* and 2 through 8. Their rankings depend only on candidates’ merit profiles and they respect domination.*

**Proof.** Majority judgment and any point-summing method clearly satisfy all axioms, so there are as many methods as one wishes that satisfy the axioms.

All methods satisfying the axioms depend only on their merit profiles. To see this compare two competitors. By Axiom 8 (IIA) it suffices to compare them alone. If two candidates  $A$  and  $B$  have the same set of grades, so that  $B$ ’s list of  $n$  grades is a permutation  $\sigma$  of  $A$ ’s list, it is shown that they must be tied. Consider, first, the opinion profile  $\phi^1$  of  $A$  and a candidate  $A'$  who may be added to the set of candidates by IIA,

$$\phi^1 : \begin{array}{c|cccccc} & v_1 & \cdots & v_{\sigma 1} & \cdots & v_n \\ \hline A : & \alpha_1 & \cdots & \alpha_{\sigma 1} & \cdots & \alpha_n \\ A' : & \alpha_{\sigma 1} & \cdots & \alpha_1 & \cdots & \alpha_n \end{array}$$

where  $A'$ ’s list is the same as  $A$ ’s except that the grades given by voters  $v_1$  and  $v_{\sigma 1}$  have been interchanged.  $\phi^1$  is possible since the domain is unrestricted. Suppose  $A \succeq_M A'$ . Interchanging the votes of the voters  $v_1$  and  $v_{\sigma 1}$  yields the profile  $\phi^2$

---

<sup>13</sup>Mieux Voter: [www.mieuxvoter.fr](http://www.mieuxvoter.fr)

$$\phi^2 : \begin{array}{c|cccccc} & v_{\sigma 1} & \cdots & v_1 & \cdots & v_n \\ \hline A : & \alpha_{\sigma 1} & \cdots & \alpha_1 & \cdots & \alpha_n \\ A' : & \alpha_1 & \cdots & \alpha_{\sigma 1} & \cdots & \alpha_n \end{array}$$

Nothing has changed by Axiom 3 (anonymity), so the first row of  $\phi^2$  ranks at least as high as the second. But by Axiom 4 (neutrality)  $A' \succeq_M A$ , implying  $A \approx_M A'$ . Thus  $A \approx_M A'$  where  $A'$ 's first grade agrees with  $B$ 's first grade.

Compare, now,  $A'$  with another added candidate  $A''$ , with profile  $\phi^3$

$$\phi^3 : \begin{array}{c|cccccc} & v_{\sigma 1} & v_2 & \cdots & v_{\sigma 2} & \cdots & v_n \\ \hline A' : & \alpha_{\sigma 1} & \alpha_2 & \cdots & \alpha_{\sigma 2} & \cdots & \alpha_n \\ A'' : & \alpha_{\sigma 1} & \alpha_{\sigma 2} & \cdots & \alpha_2 & \cdots & \alpha_n \end{array}$$

where  $A''$ 's list is the same as  $A'$ 's except that the grades of voters  $v_2$  and  $v_{\sigma 2}$  have been interchanged. Suppose  $A' \succeq_M A''$ . Interchanging the votes of the voters  $v_2$  and  $v_{\sigma 2}$  yields the profile  $\phi^4$

$$\phi^4 : \begin{array}{c|cccccc} & v_{\sigma 1} & v_{\sigma 2} & \cdots & v_2 & \cdots & v_n \\ \hline A' : & \alpha_{\sigma 1} & \alpha_{\sigma 2} & \cdots & \alpha_2 & \cdots & \alpha_n \\ A'' : & \alpha_{\sigma 1} & \alpha_2 & \cdots & \alpha_{\sigma 2} & \cdots & \alpha_n \end{array}$$

so as before conclude that  $A' \approx_M A''$ . Axiom 7 (transitivity) now implies  $A \approx_M A''$  where  $A''$ 's first two grades agree with  $B$ 's first two grades.

Repeating this reasoning shows  $A \approx B$ , so which voter gave which grade has no significance. Therefore a candidate's distribution of grades—his merit profile—is what determines his place in the ranking with any method that satisfies the Axioms. It has a unique representation when the grades are listed from the highest to the lowest.

Suppose  $A$ 's grades  $\alpha$  dominates  $B$ 's grades  $\beta$ , both given in order of decreasing grades. Domination means  $\alpha_j \succeq \beta_j$  for all  $j$ , with at least one strictly above the other. If  $\alpha_k \succ \beta_k$  replace  $\beta_k$  in  $\beta$  by  $\alpha_k$  to obtain  $\beta^1 \succ_M \beta$  by monotonicity (Axiom 5). Either  $\beta^1 \approx_M \alpha$  proving that  $\alpha \succ_M \beta$ , or else  $\alpha \succ_M \beta^1$ . In the second case, do as before to obtain  $\beta^2 \succ_M \beta^1$ , and either  $\beta^2 \approx_M \alpha$ , or else  $\alpha \succ_M \beta^2$ . If  $\beta^2 \approx_M \alpha$  then  $\beta \prec_M \beta^1 \prec_M \beta^2 \approx_M \alpha$  and transitivity implies  $\beta \prec_M \alpha$ . Otherwise, repeating the same argument shows that  $\alpha \succ_M \beta$ .

Monotonicity implies domination is respected. ■

A reasonable method should certainly avoid the domination paradox. This theorem shows that any method based on evaluations (as Dahl suggested) that satisfies May's axioms and avoids the Condorcet and Arrow paradoxes does so. When majority rule fails to respect domination it differs from all such methods: why then the persistent insistence on agreeing with the majority rule?

## 6 Polarization

Are there reasons to choose one method among all that meet the demands of Theorem 3? All ranking methods that satisfy IIA (Axiom 8) are determined by how they rank pairs of candidates. So what makes sense when two candidates are to be ranked? In particular, are there circumstances when majority rule for two candidates *is* acceptable?

One instance leaps to mind: jury decisions. The goal is to arrive at the truth, the correct decision, either the defendant is guilty or is not guilty. A juror may be more or less confident in his judgment: the higher his belief that one decision is correct the lower his belief that the opposite decision is correct. In this context Condorcet's jury theorem strongly supports MR. But this context is very different from that of most elections between two candidates where gradations of opinion are inherent and an excellent opinion of one does not necessarily imply a low opinion of the other.

"Political polarization" is a phenomenon that is more and more observed and discussed in the United States and other countries (see e.g., [10, 18]). It refers to a partisan cleavage in political attitudes in support of ideological extremes—pro-abortion vs. anti-abortion, pro-evolution vs. anti-evolution, left vs. right—and applies to voters, elites, candidates, and parties. The concept necessarily concerns an opposition between two. The word is used when (say) large majorities of Democratic and Republican voters are vehemently on opposite sides in their evaluations of issues or candidates. The notion evokes the idea that most voters are at once intensely for one side and intensely against the other, so the situation approaches that of a jury decision where there is no question of (in Dahl's words) pitting a passionate minority against an apathetic majority.

Consider the two major opponents of the 2012 French presidential election poll (Table 3a), Hollande (moderate left) and Sarkozy (traditional right). Table 5a gives the electorate’s opinion profile concerning them (where, e.g., 1.63% in the first column give Sarkozy *Fair* and Hollande *Outstanding*). Tables 5b and 5c contain respectively the distributions and cumulative distributions of the grades given Hollande for each of the grades given Sarkozy (Table 5b is obtained from Table 5a by normalizing each line to sum to 100%).

		Hollande							Total
		<i>Outs.</i>	<i>Exc.</i>	<i>V.G.</i>	<i>Good</i>	<i>Fair</i>	<i>Poor</i>	<i>Rej.</i>	
S	<i>Outs.</i>	0.14%	0.00%	0.41%	1.09%	2.04%	2.99%	2.99%	09.63%
a	<i>Exc.</i>	0.27%	1.09%	0.95%	2.17%	2.71%	2.71%	2.44%	12.35%
r	<i>V.G.</i>	0.27%	1.22%	2.04%	3.12%	2.99%	3.93%	2.71%	16.28%
k	<i>Good</i>	1.22%	1.09%	1.76%	1.76%	2.85%	1.63%	0.68%	10.99%
o	<i>Fair</i>	1.63%	2.44%	2.58%	1.09%	2.31%	0.68%	0.41%	11.13%
z	<i>Poor</i>	1.75%	2.58%	1.09%	0.27%	0.54%	0.81%	0.81%	07.87%
y	<i>Rej.</i>	7.19%	7.73%	7.60%	2.17%	1.36%	1.49%	4.21%	31.75%
Total		12.48%	16.15%	16.42%	11.67%	14.79%	14.25%	14.25%	

Table 5a. Opinion profile, Hollande-Sarkozy, 2012 French presidential poll.

		Hollande						
		<i>Outs.</i>	<i>Exc.</i>	<i>V.Good</i>	<i>Good</i>	<i>Fair</i>	<i>Poor</i>	<i>Rej.</i>
S	<i>Outs.</i>	01.41%	00.00%	04.23%	11.27%	21.13%	30.99%	30.99%
a	<i>Exc.</i>	02.20%	08.79%	07.69%	17.58%	21.98%	21.98%	19.78%
r	<i>V.Good</i>	01.67%	07.50%	12.50%	19.17%	18.33%	24.17%	16.67%
k	<i>Good</i>	11.11%	09.88%	16.05%	16.05%	25.93%	14.81%	06.17%
o	<i>Fair</i>	14.63%	21.95%	23.17%	09.76%	20.73%	06.10%	03.66%
z	<i>Poor</i>	22.41%	32.76%	13.79%	03.45%	06.90%	10.34%	10.34%
y	<i>Rej.</i>	22.65%	24.36%	23.93%	06.84%	04.27%	04.70%	13.25%

Table 5b. Distributions of Hollande’s grades for each of Sarkozy’s grades, 2012 French presidential poll.

		Hollande						
		<i>Outs.</i>	$\succeq$ <i>Exc.</i>	$\succeq$ <i>V.Good</i>	$\succeq$ <i>Good</i>	$\succeq$ <i>Fair</i>	$\succeq$ <i>Poor</i>	$\succeq$ <i>Rej.</i>
S	<i>Outs.</i>	01.41%	01.41%	05.64%	16.91%	38.04%	69.03%	100%
a	<i>Exc.</i>	02.20%	10.99%	18.68%	36.26%	58.24%	80.23%	100%
r	<i>V.Good</i>	01.67%	09.17%	21.67%	40.84%	59.17%	83.34%	100%
k	<i>Good</i>	11.11%	20.99%	37.04%	53.09%	79.02%	93.83%	100%
o	<i>Fair</i>	14.63%	36.58%	59.75%	69.51%	90.24%	96.34%	100%
z	<i>Poor</i>	22.41%	55.17%	68.96%	72.41%	79.31%	89.65%	100%
y	<i>Rej.</i>	22.65%	47.01%	70.94%	77.78%	82.05%	86.75%	100%

Table 5c. Cumulative distributions of Hollande’s grades for each of Sarkozy’s grades, 2012 French presidential poll.

Comparing any two lines of Table 5c, the lower (almost) always dominates the upper. Thus the lower the grade given Sarkozy the higher the distribution of grades given Hollande: statistically there is diametrical opposition in their grades. The same occurs with the distributions of Sarkozy’s grades for each of Hollande’s grades. Moreover, as may be seen in Table 5d, 47.36% of the voters are polarized on the two candidates.

	7.19%	7.73%	7.60%	2.17%	1.36%	0.54%	2.31%
Hollande:	<i>Outs.</i>	<i>Exc.</i>	<i>V.Good</i>	<i>Good</i>	<i>Fair</i>	<i>Fair</i>	<i>Fair</i>
Sarkozy:	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Poor</i>	<i>Fair</i>
	2.85%	2.99%	3.93%	2.71%	2.44%	2.99%	
Hollande:	<i>Fair</i>	<i>Fair</i>	<i>Poor</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	
Sarkozy:	<i>Good</i>	<i>V.Good</i>	<i>V.Good</i>	<i>V.Good</i>	<i>Exc.</i>	<i>Outs.</i>	

Table 5d. Polarized part (47.36%) of true opinion profile, Hollande-Sarkozy, 2012 French presidential election poll (for merit profile of Table 4a).

On the other hand, analyzing in the same manner the grades given to two extreme left candidates, Mélenchon and Poutou, the higher the grade given one the higher the distribution of grades given the other: statistically there is a concordance of grades. All of this confirms common sense.

Given an electorate’s merit profile take any two candidates. Different opinion profiles on the two candidates have this same merit profile. An opinion profile on a pair of candidates will be said to be “polarized” when the higher a voter evaluates one candidate the lower the voter evaluates the other. Specifically:

Two candidates  $A$  and  $B$  are *polarized* if, for any two voters,  $v_i$  evaluates  $A$  higher (respectively, lower) than  $v_j$  then  $v_i$  evaluates  $B$  no higher (respectively, no lower) than  $v_j$ .

This definition includes and extends the usual (loosely defined) notion of a polarized electorate and formulates a “pure” form of a type of distribution of grades found in practice. It brings to mind the single crossing opinion profile of the traditional theory [12, 47].

The merit profile of Hollande and Sarkozy (Table 4a or 5a) showed that Hollande’s grades largely dominate Sarkozy’s, and yet there are opinion profiles (e.g., Table 4b) that make Sarkozy the overwhelming MR-winner. The polarized opinion profile that has the same merit profile is given in Table 5e.

	12.48%	16.15%	3.12%	7.87%	5.43%	5.70%	5.97%
Hollande:	<i>Outs.</i>	<i>Exc.</i>	<i>V.Good</i>	<i>V.Good</i>	<i>V.Good</i>	<i>Good</i>	<i>Good</i>
Sarkozy:	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Poor</i>	<i>Fair</i>	<i>Fair</i>	<i>Good</i>

  

	5.02%	9.77%	6.51%	7.74%	4.61%	9.63%
Hollande:	<i>Fair</i>	<i>Fair</i>	<i>Poor.</i>	<i>Poor</i>	<i>Rej.</i>	<i>Rej.</i>
Sarkozy:	<i>Good</i>	<i>V.Good</i>	<i>V.Good</i>	<i>Exc.</i>	<i>Exc.</i>	<i>Outs.</i>

Table 5e. Polarized opinion profile, Hollande-Sarkozy, 2012 French presidential election poll (for merit profile of Table 4a).

The polarized opinion profile makes Hollande the MR-winner with a score of 50.75% to Sarkozy’s 43.28%, 5.97% giving both the grade *Good*, so agrees with Hollande’s domination in grades. With this profile MR makes the right decision.

It would seem that it is precisely when an electorate is polarized—or when a jury seeks the correct answer between two opposites—and there can be no consensus that the “strongly for or strongly against” characteristic of MR should render the acceptable result. The electorate is not polarized on Hollande vs. Sarkozy (Table 5a): it is only statistically polarized. Completely polarized pairs are rare; however, significant parts may be polarized (Tables 5d,h). Consider, for example, the breakdown of the grades given to Poutou (extreme left) and Le Pen (extreme right) in the 2012 French presidential poll (see Table 5f).

		Poutou						
		<i>Outs.</i>	<i>Exc.</i>	<i>V.G.</i>	<i>Good</i>	<i>Fair</i>	<i>Poor</i>	<i>Rej.</i>
L	<i>Outs.</i>	0.00%	0.14%	0.14%	0.00%	0.27%	0.95%	4.48%
	<i>Exc.</i>	0.00%	0.14%	0.41%	0.41%	1.36%	1.90%	3.12%
e	<i>V.G.</i>	0.00%	0.00%	0.81%	0.14%	1.09%	3.39%	4.07%
	<i>Good</i>	0.00%	0.00%	0.41%	1.09%	1.22%	1.36%	5.29%
P	<i>Fair</i>	0.00%	0.14%	0.41%	0.68%	2.58%	4.75%	5.43%
e	<i>Poor</i>	0.00%	0.27%	0.14%	0.27%	1.09%	1.90%	2.58%
n	<i>Rej.</i>	0.14%	0.68%	2.17%	5.16%	4.88%	13.84%	20.76%
Total		0.14%	1.36%	4.48%	7.73%	12.48%	28.09%	45.73%

Table 5f. Opinion profile, Le Pen-Poutou, 2012 French presidential poll.

The polarized opinion profile that agrees with the merit profile is given in Table 5g (MR places Poutou with 47.63% ahead of Le Pen with 46.14%). It does not agree with the true opinion profile.

	0.14%	1.36%	4.48%	7.73%	12.48%	21.44%	6.24%
Poutou:	<i>Outs.</i>	<i>Exc.</i>	<i>V.Good</i>	<i>Good</i>	<i>Fair</i>	<i>Poor</i>	<i>Poor</i>
Le Pen:	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Poor</i>

  

	0.41%	13.57%	9.36%	9.50%	7.33%	5.97%
Poutou:	<i>Poor</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>
Le Pen:	<i>Fair</i>	<i>Fair</i>	<i>Good</i>	<i>V.Good</i>	<i>Exc.</i>	<i>Outs.</i>

Table 5g. Polarized opinion profile, Le Pen-Poutou, 2012 French presidential election poll (for merit profile of Table 4a).

	0.14%	0.68%	2.17%	5.16%	4.88%	13.84%	20.76%
Poutou:	<i>Outs.</i>	<i>Exc.</i>	<i>V.Good</i>	<i>Good</i>	<i>Fair</i>	<i>Poor</i>	<i>Rej.</i>
Le Pen:	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>

  

	2.58%	5.43%	5.29%	4.07%	3.12%	4.48%
Poutou:	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>	<i>Rej.</i>
Le Pen:	<i>Poor</i>	<i>Fair</i>	<i>Good</i>	<i>V.Good</i>	<i>Exc.</i>	<i>Outs.</i>

Table 5h. Polarized part (72.60%) of true opinion profile, Le Pen-Poutou, 2012 French presidential election poll (for merit profile of Table 4a).

However, a large part of the true opinion profile—almost three-quarters—is polarized on Le Pen and Poutou as may be seen in Table 5h, and every opinion profile will have parts that are polarized.

Polarization, as defined here, is a significant phenomenon in practice. It is also of significance in theory. It is proven in the next section—and was observed in the example of Table 5e—that when two candidates are polarized and one’s grades dominate the other’s MR (when decisive) makes the correct decision: it *is* acceptable.

## 7 A new chacterization of majority judgment

MR on two candidates has a particularly important property: it resists manipulation or is incentive compatible (see next section). Any method that agrees with it inherits the property. MR may err, but it does not when a pair of candidates is polarized (see corollary below), and that is precisely the situation when voters are most tempted to manipulate. Thus for a method to agree with MR on polarized pairs is very attractive.

**Definition:** A method of ranking  $\succeq_M$  is *consistent with the majority rule on polarized pairs of candidates* if both give the identical ranking between every pair of polarized candidates whenever both are decisive (meaning no ties). Generically, ties have a zero probability of occurring when the number of voters is large.

Not every method satisfying Axioms 1\* and 2 – 8 is consistent with MR on polarized electorates. Point-summing methods, for example, are not. Take the polarized opinion profile of Poutou and Le Pen (Table 5g). MR makes Poutou the winner with 47.63% of the votes to Le Pen’s 46.14%, 6.24% evaluating both to be *Poor*. But if *Outstanding* is worth 6 points, *Excellent* 5,  $\dots$ , down to *To Reject* 0, Le Pen’s 172.75 points easily defeats Poutou’s 101.80.

**Theorem 4** *A method of ranking  $\succeq_M$  that satisfies Axioms 1\* and 2-8 and is consistent with the majority rule on polarized pairs of candidates must coincide with the majority-gauge rule  $\succ_{MG}$  when the scale of grades  $\Lambda$  is sufficient.*<sup>14</sup>

**Proof.** Given an opinion profile with any number of candidates, IIA (Axiom 8) implies that the order between any two must be determined between the two alone. By Theorem 3, the method  $\succeq_M$  depends only on the merit profile—the distributions of the candidates’ grade—not on which voters gave which grades. Thus, the order between them determined by the method is the same as that determined by the polarized opinion profile.

	$x_1\%$		$x_j\%$		$x_{k-1}\%$		$x_k\%$		$x_{k+1}\%$		$x_s\%$
$A$	$\lambda_A^1 \succeq$	$\dots$	$\lambda_A^j$	$\dots$	$\succeq$	$\lambda_A^{k-1}$	$\succeq$	$\lambda_A^k$	$\succeq$	$\lambda_A^{k+1}$	$\succeq \dots \succeq \lambda_A^s$
	$\succ$	$\dots$	$\succ$	$\dots$	$\succ$		$=$		$\prec$	$\dots$	$\prec$
$B$	$\lambda_B^1 \preceq$	$\dots$	$\lambda_B^j$	$\dots$	$\preceq$	$\lambda_B^{k-1}$	$\preceq$	$\lambda_B^k$	$\preceq$	$\lambda_B^{k+1}$	$\preceq \dots \preceq \lambda_B^s$

Table 6. Polarized opinion profile.

So consider the polarized opinion profile of two candidates,  $A$ ’s grades going from highest on the left to lowest on the right and  $B$ ’s from lowest to highest, as displayed in Table 6.  $A$ ’s grades are non-increasing and  $B$ ’s non-decreasing and for any  $j$  either  $\lambda_A^j \succ \lambda_A^{j+1}$  or  $\lambda_B^j \prec \lambda_B^{j+1}$ , so the corresponding grades can be equal at most once (as indicated in the middle line of the profile).<sup>15</sup>

Suppose  $A$  is the MR-winner. Then  $x_A = \sum_1^{k-1} x_i > \sum_{k+1}^s x_i = x_B$ . By assumption  $x_A = 50\%$  is excluded (since  $\succ_{MG}$  must be decisive).

If  $x_A > 50\%$  then  $A$ ’s majority-grade is at least  $\lambda_A^{k-1}$  and  $B$ ’s at most  $\lambda_B^{k-1}$ , so  $A$  is ahead of  $B$  by the MG rule.

Otherwise  $x_B < x_A < 50\%$  implying  $x_A + x_k > 50\%$  and  $x_B + x_k > 50\%$ , so the candidates’ majority-grades are the same,  $\lambda_A^k = \lambda_B^k$ . Notice that  $p_A \leq x_A$  and  $q_B \leq x_A$  but one of the two must be an equality; similarly  $q_A \leq x_B$  and  $p_B \leq x_B$  but one of the two must be an equality as well. Thus  $x_A = \max\{p_A, q_B\}$  and  $x_B = \max\{p_B, q_A\} < x_A$ . If  $p_A = x_A$  then  $p_A$  is the largest of the  $p$ ’s and  $q$ ’s, and MG puts  $A$  above  $B$ ; if  $q_B = x_A$  then  $q_B$  is the largest of the  $p$ ’s and  $q$ ’s, and MG puts  $B$  below  $A$ , as was to be shown.

<sup>14</sup>The assumption of a sufficient scale of grades is necessary only to compare the rankings of a method based on evaluations with the rankings based on preferences; otherwise, several preference profiles may be associated with the same opinion profile (depending on how two candidates with the same grade are compared: indifferent or one of the two preferred).

<sup>15</sup>This brings to mind the single crossing property mentioned earlier.



Now assume the MG rule places  $A$  above  $B$ ,  $(p_A, \lambda_A, q_A) \succ_{MG} (p_B, \lambda_B, q_B)$ . Either  $A$ 's majority-grade is above  $B$ 's,  $\lambda_A \succ \lambda_B$ ; or they are equal and either  $p_A > \max\{q_A, p_B, q_B\}$  or  $q_B > \max\{p_A, q_A, p_B\}$ .

In the first case, at least  $(50 + \epsilon)\%$  (for some  $\epsilon > 0$ ) of the voters gave the grade  $\lambda_A$  or better to  $A$  and  $\lambda_B$  or worse to  $B$ . They constitute a majority of at least  $(50 + \epsilon)\%$  that makes  $A$  the MR-winner.

In the second case, they are equal and are the  $k$ th column of the polarized opinion profile:  $\lambda_A = \lambda_B = \lambda_A^k = \lambda_B^k$ . As before,  $x_A = \sum_{i=1}^{k-1} x_i = \max\{p_A, q_B\}$  and  $x_B = \sum_{i=k+1}^s x_i = \max\{p_B, q_A\}$ . Thus,  $p_A > \max\{q_A, p_B, q_B\}$  implies  $x_A = p_A > \max\{p_B, q_A\} = x_B$  and  $q_B > \max\{p_A, q_A, p_B\}$  implies  $x_A = q_B > \max\{p_B, q_A\} = x_B$ , so in both instances  $A$  is the MR-winner. ■

**Corollary** *Majority rule on a pair of polarized candidates elects (when decisive) a candidate whose grades dominate the other's.*

**Proof.** The majority-gauge rule elects the candidate whose grades dominate the other's. But as just shown the majority rule coincides with it when the candidates are polarized, proving the corollary. ■

If the scale of grades is not sufficient—i.e., a voter is forced to give a same grade to two candidates whereas she has a preference between them—the theorem does not apply since it is impossible to deduce the ranking from the grades (a same grade for two candidates could bear three meanings, preference for one of the two or indifference). For example, with two grades MJ becomes approval voting, a voter may *Approve* both candidates or *Disapprove* both without actually being indifferent between them, so that agreement with majority rule calculated on the basis of preferences is impossible even when the electorate is polarized, and there is no guarantee that AV agrees with MR. For example, if *Approve* means *Good* or better in the polarized profile of Poutou versus Le Pen (Table 5f), approval voting elects Le Pen with 32.16% not Poutou with 13.71% in disagreement with MR (and so MJ with the full complement of grades). Thus, to guarantee agreement with the majority rule on polarized pairs, the scale of grades must be rich enough to faithfully represent the preferences and indifferences. The optimal number of grades to use depends on the application.

In the most cited paper of the first hundred years of the *Psychological Review* G. A. Miller [40] concluded, “There is a clear and definite limit to the accuracy with which we can identify absolutely the magnitude of a unidimensional stimulus variable. I would propose to call this limit the *span of absolute judgment*, and I maintain that for unidimensional judgments this span is usually somewhere in the neighbourhood of seven.”

The evidence with MJ in voting concords with Miller's finding (Table 6). In the 2007 experiment where voters were offered a scale of six grades only 14% used all six although there were twelve candidates. In the 2012 poll voters had a scale of seven grades, yet the number of grades they used is remarkably similar to that of the participants in the 2007 experiment, only 1% using all seven grades to evaluate ten candidates. This suggests that at least five grades are necessary, six is a good number, seven is unnecessary.

No. of grades:	1	2	3	4	5	6	7	Total
2007 (12 candidates):	1%	2%	10%	31%	42%	14%	–	100%
2012 (10 candidates):	1%	6%	13%	31%	36%	13%	1%	100%

Table 6. Percentages of ballots with  $k$  grades, 2007 French presidential election experiment and 2012 French presidential poll.

Experts in judging figure skaters, divers, gymnasts, or wines may, however, have the finesse to discern refinements that demand finer scales; in figure skating, for example, judges use a scale of 13 grades (for each part of a performance) and in diving a scale of 21 (for each dive).

## 8 Resisting strategic manipulation

The Gibbard-Satterwaite impossibility theorem [29, 50] proves that no method of ranking satisfying Axioms 1-7—so based on comparisons—is strategy-proof when there are at least three candidates. Yet, “It is important to ask under what circumstances it would be possible to design non-trivial strategy-proof decision rules, because strategy-proofness, when attainable, is an extremely robust and attractive property” ([11] Sect. 5).

Suppose a method  $M$  satisfying Axioms 1\* and 2-8 ranks candidate  $A$  above  $B$ ,  $A \succ_M B$ . Because of monotonicity, a voter who prefers  $B$  to  $A$  can try to manipulate by lowering  $A$ 's grade and raising  $B$ 's. The method resists to strategic manipulation if  $A$  remains above  $B$  after the manipulation.

**Definition:** A method  $M$  is *strategy-proof in ranking* if for every two candidates  $A$  and  $B$  s.t.  $A \succ_M B$  and every voter who grades  $B$  above  $A$ , he can neither raise  $B$ 's global measure<sup>16</sup> nor lower  $A$ 's.

<sup>16</sup>Examples of global measures are a candidate's mean of his grades for point-summing, his number of ticks for AV, his majority-gauge for MJ.

As will be shown, there is no such method. So formulate a less demanding property.

**Definition:** A method is *partially strategy-proof in ranking* if for every two candidates  $A$  and  $B$  s.t.  $A \succ_M B$  and every voter who grades  $B$  above  $A$ , if he can raise  $B$ 's global measure then he cannot lower  $A$ 's, and if he can lower  $A$ 's then he cannot raise  $B$ 's.

Point-summing methods are clearly not partially strategy-proof in ranking.

**Theorem 5** (1) *No method of ranking satisfying Axioms 1\* and 2-8 is strategy-proof in ranking on the entire domain of possible opinions.* (2) *The majority-gauge rule  $\succ_{MG}$  is partially strategy-proof in ranking on the entire domain when the language of grades is sufficient.*

**Proof.** (1) Suppose a method  $\succeq_M$  satisfying the axioms is strategy-proof in ranking on the entire domain. Axiom 8—IIA—implies it must be strategy-proof on any two candidates  $A$  and  $B$  with profile  $\Phi$ . If a voter  $j$  gave  $A$  the higher grade change both of  $j$ 's grades, giving  $A$  the highest and  $B$  the lowest; if voter  $j$  gave  $B$  the higher grade change both, giving  $B$  the highest and  $A$  the lowest; and if voter  $j$  gave both the same grade change nothing (she is indifferent). Strategy-proofness or monotonicity implies  $\succeq_M$  must give the same outcome for this new profile on  $A$  and  $B$ . Doing the same for every voter in turn results in an opinion profile  $\Phi'$  for which the method  $\succeq_M$  necessarily gives the same outcome between  $A$  and  $B$  as does  $\Phi$ .

In  $\Phi'$  let  $m_A$  be the number of voters who give  $A$  the highest and  $B$  the lowest grade,  $m_B$  the number who give  $B$  the highest and  $A$  the lowest, and  $m_{AB}$  the number who gave both the same grade ( $m_{AB}$  is also the number who gave the same grade in  $\Phi$ ). If  $m_A > m_B$  then  $A$ 's grades dominate  $B$ 's and monotonicity implies  $A \succ_M B$ ; and if  $m_A = m_B$  their sets of grades are identical so by Theorem 3  $A \approx_M B$ . But this implies  $\succeq_M$  coincides with the majority rule  $\succeq_{MR}$ . Since the domain is unrestricted, Condorcet's paradox shows transitivity is violated, a contradiction.

(2) A voter who gives the same grade to two candidates has no preference between them (when the scale is sufficient) so has no incentive to see one ranked above the other.<sup>17</sup>

Suppose  $A \succ_{MG} B$ ,  $A$ 's majority-grade is  $\alpha^A$  and  $B$ 's  $\alpha^B$ , and voter  $j$  ranks  $B$  above  $A$ ,  $\alpha_j^B \succ \alpha_j^A$ . If  $j$  can raise  $(p_B, \alpha^B, q_B)$  then  $\alpha_j^B \preceq \alpha^B$ , implying  $\alpha_j^A \prec \alpha_j^B \preceq \alpha^B \preceq \alpha^A$  so  $j$  cannot lower  $(p_A, \alpha^A, q_A)$ ; whereas if  $j$  can lower  $(p_A, \alpha^A, q_A)$  then  $\alpha_j^A \succeq \alpha^A$ , implying  $\alpha_j^B \succ \alpha_j^A \succeq \alpha^A \succeq \alpha^B$  so  $j$  cannot raise  $(p_B, \alpha^B, q_B)$ . ■

**Lemma** *A method of ranking  $\succeq_M$  that satisfies Axioms 1\* and 2-8 and is strategy-proof on the limited domain of polarized pairs of candidates must be consistent with the majority rule  $\succeq_{MR}$  on polarized pairs, when the language of grades is sufficient.*

**Proof.** The proof is very similar to that of Theorem 5. Suppose  $\succeq_M$  satisfies Axioms 1\* and 2-8 and is strategy-proof on pairs of polarized candidates. Take any such pair with profile  $\Phi$ . If voter  $j$  gave  $A$  the higher grade change both grades, giving  $A$  the highest and  $B$  the lowest; if voter  $j$  gave  $B$  the higher grade change both, giving  $B$  the highest and  $A$  the lowest; and if voter  $j$  gave both the same grade change nothing (he is indifferent). Since the new opinion profile on the two candidates is again polarized, strategy-proofness or monotonicity implies that  $\succeq_M$  gives the same outcome. Doing the same for all voters gives an opinion profile  $\Phi^*$  for which the outcome is the same as for  $\Phi$ . The last part of (1) in the proof of Theorem 5 now shows that  $\succeq_M$  between any two polarized opinion profiles is  $\succeq_{MR}$ . ■

**Theorem 6** *A method of ranking  $\succeq_M$  that satisfies Axioms 1\* and 2-8 and is strategy-proof on the limited domain of polarized pairs of candidates must coincide with the majority-gauge rule  $\succeq_{MG}$  when the language of grades is sufficient.*

**Proof.** The Lemma together with Theorem 4 proves it. ■

Theorems 5 and 6 provide new theoretical evidence that majority judgment (with a sufficient number of grades) best resists strategic manipulation. It says that the majority-gauge rule is always partially strategy proof and when two candidates are polarized—and sizable parts of an electorate often are—the majority-gauge rule is strategy-proof. To appreciate the importance of those results in practice, consider the merit profile of Table 4a from which we can compute that:

- The majority-gauge of Hollande is (45.05%, *Good*+, 43.28%); and
- The majority-gauge of Sarkozy is (49.25%, *Fair*+, 39.62%).

<sup>17</sup>Indeed, his primary objective may be for them to receive the majority-grade he has given, but because of the majority grade mechanism, it is strategy proof to be honest in that case.

From the opinion profile in Table 5a, we can deduce that 40.31% of all voters strictly grade Sarkozy above Hollande. They are of three types:

- Type 1 voters strongly prefer Sarkozy to Hollande because they give to Sarkozy a high grade ( $\succeq$  *Good*) and to Hollande a low grade ( $\preceq$  *Fair*); They are the most motivated to manipulate.
- Type 2 voters like both candidates because give a high grade to both (to Hollande a grade  $\succeq$  *Good*, and so to Sarkozy a grade  $\succeq$  *Very Good*); They are less motivated to manipulate.
- Type 3 voters dislike both candidates because give a low grade to both (to Sarkozy a grade  $\preceq$  *Fair* and so to Hollande a grade  $\preceq$  *Poor*). They are less motivated to manipulate.

From Table 5a, we can compute that type 1 voters are the largest fraction because they form 76.09% of the 40.31% who prefer Sarkozy to Hollande. Type 2 constitute 19.20% and type 3 only 4.71%. How those three types of voters can manipulate? Just up Sarkozy’s grade to *Outstanding*, down Hollande’s to *To Reject*. The effects on majority-gauges are very limited:

- Type 1 voters (76.09%, most tempted to manipulate) have no effect whatsoever. Type 1 voters give to Sarkozy a grade above his majority grade and to Hollande a grade below his majority grade. Consequently, increasing Sarkozy’s grade does not change the total of grades above his majority grade and decreasing Hollande’s grade does not change the total of grades below his majority grade. As such: even if all of type 1 voters manipulate, the majority-gauges of Hollande and Sarkozy remain exactly the same and so the ranking remains unchanged.
- Type 2 voters (19.20%, not motivated to manipulate) can only partially manipulate by decreasing Hollande’s majority-gauge, but cannot increase Sarkozy’s majority-gauge;
- Type 3 voters (4.71%, not motivated to manipulate) can only partially manipulate by increasing Sarkozy’s majority-gauge, but cannot decrease Hollande’s majority-gauge.

This example illustrates why majority-gauge rule resists strategic manipulation in practice. Voters that are the most tempted to manipulate (typically, the largest fraction) have no effect on the majority-gauges after manipulation. The other voters (a smaller fraction) have limited impact because they can only partially manipulate. Moreover, they are not very motivated to manipulate because either they like both candidates or dislike both and so perhaps voting honestly provide them more satisfaction than manipulating the result.

Now suppose that all type 1’s up Sarkozy’s grade to *Outstanding*, and down Hollande’s to *To Reject* and that all voters of types Types 2 & 3 that are “sufficiently motivated” (e.g. grades differ by at least two levels) do the same. This implies that 86.21% manipulate among those who prefer Sarkozy to Hollande.

With majority judgment manipulation fails: because Hollande’s M-G decreases from (45.05%, *Good*+, 43.28%)  $\searrow$  to (44.64%, *Good*−, 46.95%) and Sarkozy’s M-G increases from (49.25%, *Fair*+, 39.62%)  $\nearrow$  to (49.66%, *Fair*+, 39.62%).

With a point-summing method —Where *Outs.* has value 6 points, *Exc.* 5 points, ..., *Poor* 1 point, *Rej.* 0 point— manipulation is successful. Before manipulation: Hollande’s average score is 3.00 and Sarkozy’s average score is 2.48 and after the manipulation: Hollande’s average score is 2.56 and Sarkozy’s average score is 2.94.

Thus, if it is true that all methods are manipulable, some are more manipulable than others, and majority judgment is the less<sup>18</sup> manipulable among those that satisfy axiom 1\* and axioms 2-8.

Strategy-proofness is a very important property of a method *if* it implements correct decisions. Dictatorship, for example, is strategy-proof, yet that attribute is not a sufficient reason to adopt it. Strategy-proofness is no virtue in a method that implements wrong decisions (as MR sometimes does when there are two candidates). First and foremost a method should first guarantee a correct decision when voters or judges express themselves honestly; subject to that caveat, the method should resist manipulation as best as possible.

When a candidate grades dominate all other candidates —and in practice such a *dominating candidate* often exists<sup>19</sup>— common sense suggests that he is the correct winner of the election. By theorem 3, when voters are

<sup>18</sup>For more theoretical and practical evidences, see [4, 8].

<sup>19</sup>In practice grade dominations are frequent. The LAMSADE Jury (Table 2b) had 6 competitors, 15 comparisons, 13 were dominations. The 2012 Socialist presidential primaries had 6 competitors, 15 comparisons: in two different precincts 13 were dominations (the rankings were different) [31]. The 2007 French presidential election had 12 candidates, 66 comparisons: in the Orsay experiment that included 1,733 participants, 50 were dominations [4, 5]. The French 2012 national presidential poll reported on above (Table 3a) had 10 candidates, 45 comparisons, 34 were dominations. There were 10 possible names for the newly formed computer sciences laboratory at Paris-Diderot University, 45 comparisons, 40 were dominations. LaPrimaire [36] had 5 candidates, 10 comparisons, all were dominations.

honest, all methods satisfying axiom 1\* and 2-8 guarantee that the dominating candidate –whenever he exists– is elected. Theoretical, experimental and practical evidences, show that using majority judgement is the best chance to still elect the dominating candidate when voters are strategic.

## 9 Conclusion

The intent of this article is to convince readers of a number of main points. Several concern the drawbacks of methods based on the traditional paradigm where voters compare candidates and the “ideal” winner is a candidate who is preferred by a majority to another candidate:

- The domination paradox shows that ideal has a major flaw: majority rule for electing one of two candidates is not necessarily, with the exception of polarized electorates, a good method.
- Condorcet consistency in every situation, in consequence, is not a desirable property, contrary to the widely held view, and should certainly not be considered axiomatic.
- Comparisons as inputs to methods of voting are insufficient expressions of opinions and should be replaced by ordinal measures.

Several others concern methods based on voters’ evaluations of candidates:

- The essential property or “ideal” is that a candidate whose grades dominate another’s should lead the other (no domination paradox).
- An infinity of different methods that obey the essential properties (Axioms 1\* and 2 through 8) meet that ideal.
- Among them MJ is the one that best combats strategic manipulation.
- In particular, MJ is the one method that agrees with the traditional ideal when the electorate is polarized (i.e., when the domination paradox is excluded and the traditional ideal is acceptable).

**Acknowledgment:** It is a deep pleasure to honor the memory of Kenneth J. Arrow. He was one of the most influential figures of the 20th century in operations research and the economic sciences broadly conceived. Of particular interest to us are his monumental contributions to the theory of social organization and voting. But he was so much more, a towering figure as a man of the world: stimulating, vibrant, bright, fast, spirited, open, generous. His was the judgment that counted.

## References

- [1] Arrow, K. J. 1951. *Social Choice and Individual Values*. Yale University Press, New Haven, CT. 2nd edition 1963, Wiley, New York, NY.  
<http://www.aec.gov.au/Elections/index.htm>
- [2] Balinski, M. 2019. Réponse à des Critiques du Jugement Majoritaire. *Revue Economique*, to appear.
- [3] Balinski, M., and R. Laraki. 2007. “A theory of measuring, electing and ranking”. Proceedings of the National Academy of Sciences USA, **104**(21), 8720-8725.
- [4] Balinski, M., and R. Laraki. 2011. *Majority Judgment: Measuring, Ranking, and Electing*. MIT Press, Cambridge, MA.
- [5] Balinski, M., and R. Laraki. 2011. “Election by majority judgment: experimental evidence.” In B. Dolez, B. Grofman, and A. Laurent (eds.), *In Situ and Laboratory Experiments on Electoral Law Reform: French Presidential Elections*: 13-54. Springer, Heidelberg.
- [6] Balinski, M., and R. Laraki. 2013. “How best to rank wines: majority judgment.” In E. Giraud-Héraoud and M.-C. Pichery (eds.), *Wine Economics: Quantitative Studies and Empirical Applications*: 149-172. Palgrave Macmillan, Basingstoke, UK.
- [7] Balinski, M., and R. Laraki. 2014. “What should ‘majority decision’ mean ?” In [25]: 103-131.

- [8] Balinski, M., and R. Laraki. 2014. "Judge: *Don't vote !*" *Operations Research* 62: 483-511.
- [9] Balinski, M., and R. Laraki. 2019. "Majority judgment vs. approval voting." *Operations Research*, Special Issue Arrow, to appear, .
- [10] Barber, M., and N. McCarty. 2013. "Causes and consequences of polarization." In Jane Mansbridge and Cathie Jo Martin (eds.), *Negotiating Agreements in Politics*, Report of the task force, American Political Science Association, Washington, DC.
- [11] Barberà, S. 2010. "Strategy-proof social choice." In K. J. Arrow, A. K. Sen, and K. Suzumura (eds.), *Handbook of Social Choice and Welfare. Vol. 2*: chapter 25, 731-831. North-Holland, Netherlands.
- [12] Barberà, S. and B. Moreno. 2011. "Top monotonicity: A common root for single peakedness, single crossing and the median voter result." *Games and Economic Behavior* 73: 345-359.
- [13] Bogomolny, A. 2011. Review of [4]. Mathematical Association of America. Accessed February 1, 2016. <http://www.maa.org/press/maa-reviews/majority-judgement-measuring-ranking-and-electing>
- [14] Brams, S. J. 2010. Preface of J.-F. Laslier and M. Remzi Sanver (eds), *Handbook on Approval Voting*: vii-ix. Springer-Verlag, Berlin and Heidelberg.
- [15] Brams, S. J. 2011. "Grading candidates." (Book review of [4].) *American Scientist* 99: 426-427.
- [16] Brams, S. J., and P. C. Fishburn. 1983. *Approval Voting*. Birkhauser, Boston, MA.
- [17] Brams, S. J., and P. C. Fishburn. 2001. "A nail-biting election." *Social Choice and Welfare* 18: 409-414.
- [18] Brennan Center. 2010. "Ungovernable America? The causes and consequences of polarized democracy." Jorde Symposium. Accessed February 14, 2015, <https://www.brennancenter.org/blog/ungovernable-america-jorde-symposium>
- [19] Chernoff, H. 1954. "Rational selection of decision functions." *Econometrica* 22: 422-443.
- [20] Condorcet, le Marquis de. 1785. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. L'Imprimerie Royale, Paris.
- [21] Dahl, R. A. 1956. *A Preface to Democratic Theory*. The University of Chicago Press, Chicago, IL.
- [22] Dasgupta, P., and E. Maskin. 2008. "On the robustness of majority rule." *Journal of the European Economics Association* 6: 949-973.
- [23] Dasgupta, P., and E. Maskin. 2004. "The fairest vote of all." *Scientific American* 290 (March): 92-97.
- [24] Edelman, P. H. 2012. "Michel Balinski and Rida Laraki: Majority judgment: measuring, ranking, and electing." *Public Choice* 151: 807-810.
- [25] Elster, J. and S. Novak (eds.). 2014. *Majority Decisions*. Cambridge University Press, Cambridge, UK.
- [26] Etter, E., J. Herzen, M. Grossglauser, and P. Thiran. 2014. "Mining democracy." In *Proceedings of the Second ACM Conference on Online Social Networks*. ACM, New York, NY, 1-12. Accessed Sept. 15, 2016, <http://cosn.acm.org/2014/files/cosn060f-etterA.pdf>
- [27] Felsenthal, D. S., and M. Machover. 2008. "The majority judgment voting procedure: a critical evaluation." *Homo oeconomicus* 25: 319-334.
- [28] Fleurbaey, M. 2014. Book review of [4]. *Social Choice and Welfare* 42: 751-755.
- [29] Gibbard, A. 1973. "Manipulation of voting schemes: a general result." *Econometrica* 41: 587-601.
- [30] Goodin R. E. and C. List 2006. "A Conditional Defense of Plurality Rule: Generalizing May's Theorem in a Restricted Informational Environment". *American Journal of Political Science*, 50,940-949.
- [31] Gonzalez-Suitt, J., A. Guyon, T. Hennion, R. Laraki, X. Starkloff, S. Thibault, and B. Favreau. 2014. "Vers un système de vote plus juste?" Département d'Économie, École Polytechnique, Cahier No. 2014-20.
- [32] Hammond, P. J. 1976. "Equity, Arrow's conditions, and Rawls' difference principle." *Econometrica* 44: 793-804.

- [33] Harris – A Nielson Company. October 2014. “Ratings rebound some for President Obama’s job performance.” Accessed January 16, 2015, <http://www.harrisinteractive.com/NewsRoom/HarrisPolls.aspx>
- [34] Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky. 1971. *Foundations of Measurement. Vol. 1: Additive and Polynomial Representations*. Academic Press, New York.
- [35] Kurrild-Klitgaard, P. 1999. “An empirical example of the Condorcet paradox of voting in a large electorate.” *Public Choice* 107: 1231-1244.
- [36] LaPrimaire.org. 2016. Results. Accessed February 2, 2017. <https://articles.laprimaire.org/résultats-du-1er-tour-de-laprimaire-org-c8fe612b64cb> <https://articles.laprimaire.org/résultats-du-2nd-tour-de-laprimaire-org-2d61b2ad1394>
- [37] LaPrimaire.org. 2016. Tutorial. Accessed February 2, 2017. <https://laprimaire.org/citoyen/vote/tutorial>
- [38] Madison, J. 1787. In A. Hamilton, J. Madison, and J. Jay, *The Federalist Papers*, No. 10.
- [39] May, K. O. 1952. “A set of necessary and sufficient conditions for simple majority decision.” *Econometrica* 20: 680-684.
- [40] Miller, G. A. 1956. “The magical number seven, plus or minus 2: Some limits on our capacity for processing information.” *Psychological Review* 63: 81-7.
- [41] Moulin, H. 1988. “Condorcet’s principle implies the no show paradox.” *Journal of Economic Theory* 45: 53-64.
- [42] Moulin, H. 1988. *Axioms of Cooperative Decision Making*. Cambridge University Press, Cambridge.
- [43] Nash, J. F. 1950. “The bargaining problem.” *Econometrica* 18: 43-58.
- [44] Pew Research Center. 2016. “March 2016 political survey.” Accessed May 5, 2016,
- [45] Pew Research Center. 2016. “As election nears, voters divided over democracy and ‘respect.’” Accessed January 10, 2017,
- [46] Pons, V., and C. Tricaud. 2017. “Expressive voting and its costs: evidence from runoffs with two or three candidates.” Harvard Business School, Working Paper 17-107.
- [47] Puppe, C., and A. Slinko. 2015. “Condorcet domains, median graphs and the single crossing property.” Working paper, July.
- [48] Saari, D. G. 2001. “Analyzing a nail-biting election.” *Social Choice and Welfare* 18: 415-430.
- [49] Sadurski, W. 2008. “Legitimacy, political equality, and majority rule.” *Ratio Juris* 21: 39-65.
- [50] Satterthwaite, M. A. 1973 “Strategy-proofness and Arrow’s conditions: existence and correspondence theorems for voting procedures and social welfare functions.” *Journal of Economic Theory* 10: 187-217.
- [51] Terra Nova. 2011. “Rendre les élections aux électeurs : le jugement majoritaire.” Accessed February 17, 2015, <http://www.tnova.fr/note/rendre-les-lections-aux-lecteurs-le-jugement-majoritaire>
- [52] Tocqueville, A. de. 1831. *De la démocratie en Amérique*, Vol. 1. Edition M.-Th. Gérin, Librairie Médicis, Paris, 1951.
- [53] Weber, R. J. 1977. “Comparison of public choice systems.” Cowles Foundation discussion paper 498. Yale University, New Haven, CT.
- [54] Wikipedia. 2014. “Academic grading in Denmark.” Accessed May 7, 2014.