The strange "Majority Judgment"

Jean-François Laslier

The strange "Majority Judgment"

Jean-François Laslier*

"Majority Judgment" is an evaluative voting rule that picks a candidate with the best median evaluation. This paper deals with the following question: what does one do when choosing according to the best median? This principle amounts to finding which half-population should be neglected in order to satisfy, in the sense of a Rawlsian compromise, the other half. This principle clashes with the definition of democracy as the participation of everyone, not half of the population plus one. Moreover, providing the highest possible level of satisfaction to that half-population that is the easiest to satisfy is not what is achieved by the rules of choice called "majoritarian," which respect the Condorcet principle. On the contrary, these rules favor consensual solutions, in particular in standard political environments. This explains why the "Majority Judgment" often yields surprising outcomes.

* École d'économie de Paris/Paris School of Economics, CNRS. *Correspondence:* 48 boulevard Jourdan 75014 Paris, France. *Email:* jean-francois.laslier@ens.fr

The author would like to thank Renaud Blanch, Sylvain Bouveret, François Durand, Jérôme Lang, Antoinette Baujard, Isabelle Lebon, Annick Laruelle, Jörgen Weibull, and Herrade Igersheim for their comments.

This is the translation of the paper in French "L'étrange Jugement majoritaire" in *Revue Economique* (2018). It was edited by Cadenza Academic Translations.

INTRODUCTION

Majority Judgment is a voting system that aims to elect one candidate among several (Balinski and Laraki [2011]). This method has some advantages related to the fact that it is a "voting by evaluations" process, which allows each voter to express an opinion about all the candidates independently, not about only one of them. Majority Judgment has been promoted actively, and attracted considerable media interest in France, but most of the documents explaining this method, traditional written ones, comics, online texts and videos, ignored the numerous difficulties raised by the principle of comparing medians, on which it is based, and which singles it out from the usual evaluation methods. Yet, these difficulties have been underlined by a number of authors in the case where the comparison of the medians is meant to produce a collective choice, a "vote."

In this article, I mainly propose to clarify the following question: What do we really do when we build on the median judgments to choose among candidates?

Unlike the impression conveyed by the term "Majority Judgment," introduced by Michel Balinski and Rida Laraki, the choice defined by this method is erroneously called a "majority" choice. Summarizing a statistical distribution by its median certainly presents an analogy with the "majority rule," but as a voting rule, "Majority Judgment" acts quite differently. As we will see, Majority Judgment consists in applying a recognized principle of social justice (the principle of the minimum, or MaxMin) allowing half of the population to be disregarded. Allowing half of the population to be disregarded is not a democratic principle, that is why, in the usual political environments, Majority Judgment leads to paradoxical recommendations.

The second section lays down the study frame. The third one presents a few examples to illustrate the "MaxMed," or median maximization. We use both typical examples to illustrate logical points, and cases that are less anecdotal from the political point of view. The fourth section places an increased focus on generality, presenting Majority Judgment as a rule for social choice based on quantiles. This allows for a better understanding of the operating logic in comparing medians in general, and in the examples from the third section in particular. The fifth section completes the previous one with historical elements about the similar methods used in the past and about the issue of strategic voting.

DEFINITION AND FRAME

The proposed method is based on judgments that the voters express in semantic form about the candidates, using, for example, adjectives like *Excellent*, *Very good*, *Good*, *Average*, *Mediocre*, *Inadequate*, *To be rejected*. The "vote" consists, for each voter, in attributing one of the adjectives to each of the candidates. The ballot papers are then counted as follows: for each candidate, the median of the statistical distribution of the evaluations he or she received is computed. Balinski and Laraki call this "majority grade." By definition, at least half of the voters assess the candidate at the level of the median or above, and at least half of the voters assess him or her at the level of the median or under. Then, one chooses a candidate with the best majority grade.

As Majority Judgment takes the evaluation of the candidates by voters as data, this can be seen as a particular form of evaluative voting, and it inherits the advantages linked to these methods. In particular, these evaluative systems, in which the voter freely assesses each candidate, can escape some problems of vote splitting. This point distinguishes them specifically from the so-called single-name voting methods, where the voter only votes for one candidate. Single-name methods, with one or two rounds, create endless difficulties for the voters (the "useful vote" dilemmas) as for the parties (thinning out of the party offer or, on the contrary, an excess of competitor candidacies).

The classical way to combine evaluations, however, uses cardinal coding.[1] This is not the case with Majority Judgment, which uses adjectives that are supposed to have a meaning that is quite specific and that is common to all voters, so that each judgment expressed carries an "absolute" meaning. This possibility, absent from the traditional vote methods, is quite appealing. This is an answer to the need for expression, which is one of the components of the democratic feeling. All the voters giving the same meaning to the words "Mr. X is an acceptable candidate" is a hypothesis that can seem dubious and unverifiable, but this difficulty is independent from the purpose of this article. Thus, the objections raised in what follows against the best median are still valid even under the assumptions that the judgments are tangible and objective data.

Clearly, if the evaluation scale is not very detailed and the candidates are numerous, there is a good chance that several candidates have the same

1. The aggregation is made in an additive way. The voting terms *range voting*, *utilitarian voting*, *score voting*, or simply *evaluative voting* are used. See for example Igersheim, Baujard, and Laslier (2016).

majority grade and that another mean will be needed to separate them. This induces a practical difficulty: the challenge of deciding between identical grades. This point is not crucial yet, and we will not linger over it, as the subject of the article is the comparison principle of the median in itself, rather than its implementation.

Last clarification: even if the evaluations are endowed with the objective and absolute character desired, it is not a foregone conclusion that the voters will genuinely give their true evaluations of the candidates. In this article, we will, broadly speaking, set out the most favorable hypothesis: that voters genuinely provide their true evaluations. What does one do then, when choosing to elect the candidate according to the best median evaluation? The voting strategy issue will be addressed only briefly, in the fifth section, with the indications of the possible advantages and disavantages of this voting method.

THE PARADOXES OF MAJORITY JUDGMENT

Most of the following examples were mentioned before the publication of the works by Balinski and Laraki and can be presented with ordinal or verbal evaluations, and that is what we will do.

Non respect of the majority

The most famous paradox about comparing medians is that it does not respect the "majority" principle, that is, when there are only two options, choosing the one the majority prefers. Here is an extremely simple example of this phenomenon.

Example 1. There are three voters and two candidates, A and B. The judgments (with marks out of 20, and with adjectives) are:

	Prefer A		Prefer B
(number of voters)	(1)	(1)	(1)
A	20/20 ("Excellent")	9/20 ("Mediocre")	9/20 ("Mediocre")
B	11/20 ("Quite good")	0/20 ("To be rejected")	10/20 ("Acceptable")

The three evaluations of A are 20, 9, and 9, so the median evaluation of A is 9. The evaluations of B are 11, 10, and 0, so the median of B is 10. The

Majority Judgment chooses B in this case, whereas 2 out of 3 voters prefer A.

The example is even more suprising when the number of voters is multiplied. In the following society, which includes 101 voters, only one prefers B, and B still has the best median. Majority Judgment goes against the near unanimity of the voters.

Example 1, continued. There are 101 voters and two candidates. The judgments are as follows:

	Prefer A		Prefer B
(number of voters)	(50)	(50)	(1)
A	20/20	9/20	9/20
B	11/20	0/20	10/20

The example we give here may be striking because of the term "Majority Judgment," which Balinski and Laraki have adopted for a method based on comparing medians. This term surely creates confusion. However, the example is instructive for another reason that we will present now.

The majority principle is questionable. When detailed evaluations are available, comparable from one voter to another, a good argument against the majority principle is: why choose the option that the majority prefers *when the minority members lose much more than what the majority members benefit from*? This point, called the "tyranny of the majority," is an acceptable argument against the majority principle and is very often discussed in political philosophy (see for example Roemer [1998]). If 49% of an electorate *strongly* prefer A to B and 51% *slightly* prefer B to A, the majority rule doesn't seem to be advisable.

Yet, the previous example shows that the problem with Majority Judgment not respecting the majority principle is different. In the example, by the way, the majority expresses strong preferences (20 versus 11, and 9 versus 0) and the minority (only one voter!) expresses a weak preference (10 versus 9). The majority rule not taking into account the intensity of the preferences is a problem, from which the "majority tyranny" possibly derives, but Majority Judgment doesn't deal with the problem. Majority Judgment combines evaluations that could take this point into account, and that certainly drives voters to believe that the expressed intensity of their preferences is finally taken into account, but the use of the median prevents this. We will see that Majority Judgment, on the contrary, is particularly susceptible to the "tyranny of the majority" problem in political environments (Example 4).

For this reason, justifying the logic of the comparison of medians in the previous example seems impossible, and such examples are probably the reason why, most of the time, the supporters of decision or voting systems

based on evaluations have rejected the systems comparing the medians. Indications on the

history of this discussion can be found on the website of the Center for Range Voting,[2] an American center promoting this system.

A small disturbing candidate (*spoiler effect*)

Numerous examples show that the comparisons of medians in general, and Majority Judgment in particular, are non-intuitive. Such examples have been visible on the internet since the 1990s (as far as I know). More recently, in the academic literature, several authors have proposed typologies of these paradoxes as well as systematic comparisons, see Felsenthal and Machover (2008), Gehrlein and Lepelley (2011), or Felsenthal and Nurmi (2017). Here is another example, due to Warren Smith, that is in touch with political reality.

The 2000 US presidential election often serves as an example to illustrate the effects of vote dispersion (*spoiler effect*) in single-name voting systems, and, by contrast, to illustrate the advantage of the evaluative voting systems. This election had the Republican George W. Bush and the Democrat Al Gore in the running, as well as a third candidate, the independant Ralph Nader, who was representing a left wing more radical than the Democratic candidate. In the American system, without a second round, the Democratic candidate lost the votes that went to the independent candidate, and that led to the election of G. W. Bush. In order to reflect on the possible effect of Majority Judgment in such a situation, Warren Smith gives a simple example with three groups of voters.

Example 2. This table shows the evaluations of three candidates (Bush, Gore, and Nader), on a scale from 0 to 20, for 100 voters. Each column represents a group of voters and its weight:

(49%)	(3%)	(48%)
Gore: 20/20	Nader: 20/20	Bush: 20/20
Nader: 6/20	Gore: 4/20	Gore: 2/20
Bush: 0/20	Bush: 0/20	Nader: 0/20

2. See the webpage https://www.rangevoting.org/MedianVrange.html accessed May 3, 2017. The website is run by Warren Smith, a mathematician keen on voting rules and a propagandist of Range Voting, that is, using additive range scales.

The electorate is divided into three unequal parts: Nader's supporters represent only 3% of the population, and the rest is equally shared between Gore and Bush. Each camp rates its candidate as the best (20/20) and the other evaluations follow a political logic: Democrats are slightly sympathetic towards Nader's ideas, while Republicans completely reject them.

In the absence of Nader, Gore would win, according to all criterias one can think of, Majority Judgment included. What is more, note that Gore would also win in the absence of the Nader voters, and that all Nader voters prefer Gore to Bush. We find it hard to see which voting system would not choose Gore in this example. What does Majority Judgment say?

The median evaluations are 0 for Bush, 4 for Gore, and 6 for Nader. Consequently, Nader would be elected with the Majority Judgment method. Suprisingly, the election result is actually determined by the fact that the Gore electors rate Nader slightly higher (6/20) than the Nader electors rate Gore (4/20). The fact that nearly all of the electorate rates Nader negatively doesn't count.

This example has the benefit, among others, of showing that it is not true that all systems of evaluative voting avoid the *spoiler effect*. It all depends on how the ballot papers are counted. In this society, with Majority Judgment, Gore electors would be wrong to believe that they can rate Nader 6/20 without causing damage.

Problems relating to monotonicity

In social choice theory, "monotonicity" is the idea that, if an elected candidate gets more and more support, he or she should remain elected. Several specific variants of this requirement exist. Felsenthal and Nurmi (2017)'s monography is entirely dedicated to this issue. It turns out that Majority Judgment is not monotonic, in the sense that additional support for a winner can cause him or her to lose.

In the case of a fixed population and for a candidate viewed in isolation, how does the median evaluation vary when one of the evaluations varies? Generally, the median simply doesn't vary at all. It varies in only one case: when there is almost equality (up to only one unit) between two grades, one a little below 50%, another a little above, so that only one vote can change the candidate's median evaluation. And in this case, the change goes "in the right direction": the median increases or decreases following the change in evaluation. Consequently, in a given population, the best median cannot vary in opposition to the nature of a change in an individual evaluation. Hence,

Majority Judgment is monotonic for a fixed population, in accordance with intuition.

The situation is different with a variable population, and Majority Judgment is then subject to a startling paradox. The following illustration is adapted from Zahid (2009).

Example 3. Imagine a society made up of 11 members who hesitate between two restaurants, A and B. They decide to choose with the Majority Judgment method, but at the time of the vote, two of the friends are late and cannot participate. The 9 persons already there give their evaluations and the results are listed in the table below. It shows the median evaluations are *Quite Good* for A and *Acceptable* for B. Therefore, Majority Judgment chooses option A.

	Excellent	Very good	Good	Quite good	Acceptable	Mediocre	To be rejected
A	1	2	0	3*	3	0	0
B	0	2	2	0	2*	1	2

The whole company gets ready to leave for A, when the last two guests arrive. They indicate that, for them, choice A is *Excellent* and choice B is only *Very good*. Everything seems to work out for the best in the best of all possible worlds, since the two last guests confirm the general wish, which is the one decided on with Majority Judgment.

	Excellent	Very good	Good	Quite good	Acceptable	Mediocre	To be rejected
A	3	2	0	3*	3	0	0
B	0	4	2*	0	2	1	2

However, Majority Judgment itself says the contrary: considering the last two votes, and as the new table shows, the society now chooses B instead of A. Indeed, A retains a *Quite good* median whereas B becomes *Good*. Change of plan!

We also note that, according to Majority Judgment, the company would have been in a curious situation if the two last guests had sent the following message: “don’t worry about us, one restaurant or the other, it doesn’t matter.” Actually, as the reader can check, if the latecomers find both restaurants *Excellent*, the society must choose B, but if they find them both *Acceptable*, the society must choose A. We will discuss this matter again; it is important from a theoretical point of view, and it is known as “independence with respect to non-concerned voters,” or “separability.”

The median voter and Majority Judgment

A fundamental result in the fields of political economy and vote theory is known as the "median voter theorem." Let's recall briefly what it is.

Suppose that the voters, like the candidates, are located along a left to right axis in such a way that each voter evaluates the candidates using the simple logic of distance: among two candidates, I always prefer the one nearest to me. The voter who has as many voters to his or her left as to his or her right is called the "median voter."

Now let's consider a candidate who is located precisely where the median voter is. We will call this point the "center." We can easily see that this centrist candidate defeats, by a majority, any candidate placed elsewhere: for a challenger on the left side, all the right half of the electorate plus the median elector him/herself prefer the centrist, and vice versa. In political terms, the center defeats the right with the help of the left, and defeats the left with the help of the right. This observation forms the "median voter theorem."

This reasoning is the one Balinski and Laraki (2011) use to justify an evaluative profile being summarized by its median. Indeed, by definition, the median of an evaluative profile is such that any other evaluation is rejected, following the majority rule, either by or a majority group who find it too high, or by another group, also a majority group, who find it too low.

However, it would be wrong to infer that Majority Judgment, which compares the different medians, verifies the median voter theorem. On the contrary, against the simple logic of the theorem, Majority Judgment performs surprisingly in "unidimensional" environments, where this theorem nevertheless applies.

To explain this point, let's first consider a unidimensional society, "flat" and perfectly symmetrical, in which the voters are uniformly spread on the segment [0.1] (Figure 1). In this society, the center, located at 0.5, according to the median voter theorem, is chosen by any rule that respects the Condorcet principle.[3] It is also chosen by Majority Judgment, after tied candidates have been separated. This central point also has other properties; in particular, it is the optimal choice from both the usual utilitarian criteria point of view and the egalitarian criteria one. (Regarding the classical unidimensional model, see for example Mueller et al. [2010].) The example we propose now (adapted from Laslier [2012]) constitutes a slight disruption of this society,

3. A Condorcet winner is a candidate who is positioned in such a way that no other candidate can achieve a strict majority against him or her. A rule respects the Condorcet principle if it chooses a Condorcet winner where one exists.

which departs from its, non plausible, perfect symmetry.

Example 4. Imagine that society swings slightly towards the political right, as illustrated in Figure 2. The left, still stationary between 0 and 0.5, now only comprises 49% of the population, and the right is now, by a slight margin, the majority. It comes as no surprise that the "center," which is the Condorcet winner, shifts slightly to the right but still stays close to the center. According to the median voter theorem, the center is precisely at the median of the new distribution of the ideal points, that is at point *c*, such that:

$$50 = 49 + 51.\ (c - 0.5) / 0.5,$$

that is:

$$c = 0.5 + 1 / 102 = 0.5098\ldots \cong 0.51.$$

The utilitarian or egalitarian criteria indicate optimums that likewise shift slightly to the right, but even less than the majority criterion, as, in the new configuration, majority voting has the classical drawback of neglecting the intensity of preferences: in this case, we see that left-wing voters tend to be (slightly) further from *c* than right-wing voters, whose distribution is slightly more concentrated. We see that this example is neither "pathological" nor excessive. It features a reasonable hypothesis about unidimensional politics, majority rule, the left, the right, and the center.

Figure 1.

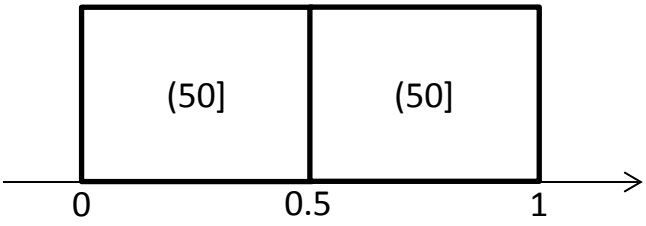
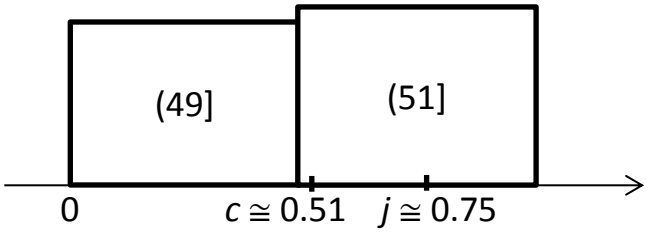


Figure 2.



For this new society, what does Majority Judgment say? Since individual evaluations follow the distances exactly, the chosen point minimizes the median distance to the ideal points of the voters, and, after a few moments of reflection, we can see that the minimal median distance is obtained around the middle of the segment [1/2, 1], that is not at the center of the total population but at the center of the right-wing electorate. Indeed, as 50% of the population is evenly spread in the interval ranging from c to 1, placed precisely at the middle of this interval, that is at

$$j = (c + 1) / 2 = 0.7549\ldots \cong 0.75,$$

all the individuals of this half-population are located at a distance from j smaller than $1 - j \cong 0.25$, a value that is consequently the best median (Figure 2). Intuitively, the logic of the best median consisted in looking for the most homogeneous half-population and in implementing the Rawls criterion (see the next section) as if the rest of the population didn't exist. In the example, by construction, the most homogeneous half-population is the right-wing electorate.

Therefore, we note that a very low and quite reasonable variation in the voters' political distribution led the choice of Majority Judgment to present a large variation. The result seems to be consistent with a kind of "majority tyranny" even though the unidimensional structure allows for certain centrist compromises to be made, as described by the median voter theorem and chosen by the "Condorcet consistent" rules. Such choices, the "Condorcet winners," are the ones that political theory, like everyday language, calls "majority" choices.

In this case, suggesting that the center of the right-wing segment of society constitutes a collective "democratic" choice and considering this choice to be a "majority" is surprising. The example is not linked to the precise shape of the distribution of the ideal points (see Laslier [2012]), which is somewhat simplistic here. We will see in the last section, as well as in general, what exact meaning the word "Majority" has in "Majority Judgment."

THE LOGIC OF MAJORITY JUDGMENT

The logic of comparing medians is not intuitive and is even difficult to understand. Giving examples is not enough. This section presents this logic from a general perspective. Majority Judgment belongs to the non-utilitarian social choice category based on comparable ordinal evaluations quantiles.

General theorems, studied since the 1960s, consequently apply to this method.

In this section, we will neglect the issues of tie-breaking, with the assumption that the best median is obtained by a single candidate. This allows for a simplifying of the presentation down to its core, the very idea behind comparing medians.

Majority Judgment as a criterion based on a quantile

In the case where the evaluations of the different states of the world, made by different people, are comparable with one another, a usual proposal of the theory of justice is called (by the economists) the Rawls criterion or Maximin criterion (Rawls [1971], [1974]; Fleurbaey [1996]). This principle relies on focusing on the individual(s) who are, for a particular policy, the least favored ones. Formally, if *u(i,x)* is the utility of the individual *i* in the policy *x*, the Rawlsian social evaluation of *x* is:

$$\text{Rawls}(x) = \text{Min}_i \{u\,(i,x)\}.$$

The Rawls criterion chooses the policy that maximizes this social evaluation, therefore:

$$\text{Max}_x \text{Min}_i \{u\,(i,x)\}.$$

Neglecting ties, the Rawls criterion simultaneously indicates a policy *x* and an individual *i* such that, in the state of the world *x*, *i* is the the least favored individual, and, in any other state of the word, the individual *i* him/herself, or another individual, would be less favored than *i* is in *x*.

We see that this criterion needs the comparison of the possible situations of different individuals in different states of the world: an "interpersonal" comparison of the utilities. However, it doesn't need a fully quantitative notion of the utility, but only to compare utility levels. For example, the Maximin criterion doesn't require being able to compute whether what an individual gains when going from *x* to *y* is more or less important than what another individual loses in the same move. The evaluations are said to be "ordinally comparable" (Bossert and Weymark [2004]; Fleurbaey and Hammond [2004]).

The Maximin criterion is obviously a very important theoretical landmark; it appears like a kind of cut-off point within the family of utilitarian

criteria,[4] which are increasingly egalitarian (d'Aspremont and Gevers [1977], Deschamps and Gevers [1978]). But its practical interest is limited in the case of large populations, insofar as the idea of making the collective choice depend on an extreme point of the evaluation (the worst evaluation) can seem excessive. Evaluation criteria that are more relevant to economic policy agree to neglect a—possibly small—fraction of the population. In this way, we can consider the maximization of income for the first decile as a social choice criterion, which amounts to neglecting the poorest 10% of the population; or for the first quartile, making the scope of the neglected population larger. In some aggregation problems, "not taking the worst income into account" is even perfectly normal and reasonable. As Elkind, Faliszewski, and Slinko (2011) judiciously note: "*it is not unusual for a professor to grade the students on the basis of their five best assignments out of six or in some sport competitions to select winners on the basis of one or more of their best attempts.*"

Majority Judgment is a Maximed; it belongs to the family of social evaluation criteria based on quantiles, with the threshold defined as half of the population. By definition, the best median is the low point of the distribution of evaluations in half a population, this half-population being such that no other half-population admits a low point higher than this best median.[5] We can then characterize Majority Judgment as the use of a classical criterion of a collective choice (the Rawls criterion) when one agrees to neglect half of the population.

This approach makes it possible to understand some of the paradoxes discussed above, example 4 in particular (a flat, slightly off-centered society). Majority Judgment looks for the most homogeneous half-population to apply the Rawls criterion. In the example, this half-population can be found in the denser part of the political spectrum, that is on the right. We must then apply the Rawls criterion, or Maximin, to this half-population. This can be done easily: the subgroup on the right chooses, at its center, the point that realizes the Rawlsian compromise of this subgroup. We understand then what the term "Majority" corresponds to in the expression "Majority Judgment": one agrees to consider only half of the population with regard to the application of the Rawls criterion, and to neglect the other half.

What is this half-population that the Rawls criterion is applied to? In order to develop a degree of intuition on this issue, we can bear in mind the case of the "spatial" vote (Enelow and Hinich [1984]). In a space with any number of dimensions, the utilities are supposed to be a function (naturally decreasing) of the distance of the ideal choice of the individual (its "position") to the planned choice. For a given group of individuals, the Rawls criterion

4. As noted by a referee of the *Revue économique*, this point can be observed on the numerical representation (through sums) of the collective welfare functions.

5. Cf. Laslier (2012), lemma 1.

applied to this group looks for the point, let's call it *x*, that minimizes the distance between *x* and the position of the group member who is the furthest from *x*. This distance measures the heterogeneity of the group spatially: if it is low, the group can find a good consensus, with the members' positions not far away from each other. To find the point for the best median evaluation in a spatial model, it is therefore necessary to look for the most homogeneous half-population in that sense.

Majority Judgment as an aggregation criterion of comparable ordinal evaluations

As seen above, Majority Judgment belongs to the general category of collective choice procedures based on comparable ordinal evaluations. Social choice theory naturally questioned the nature of the individual evaluations: is it reasonable to postulate the quantitative nature of the individual utilities? Can we avoid comparing different levels of well-being from one individual to another? All these issues have been addressed by numerous economic, philosophical, epistemological, psychological, and mathematical studies (see the collection of Elster and Roemer [1991] for multidisciplinary contributions, and d'Aspremont and Gevers [2002] or Blackorby, Bossert, and Donaldson [2002] for more technical aspects).

As the practice of applied economics is widely utilitarian, a central issue of normative economics in the 1950s and 1960s was the founding principles of public policies evaluation criteria in general, and of utilitarianism in particular. Utilitarianism numerically codes individual evaluations, which are supposed to be comparable across individuals, and then compares the sums of the individuals' evaluations. The results (Debreu [1960], Krantz et al. [1971], Pivato [2013], Macé [2015]) show the critical role of the "separability" axiom or "independence of unconcerned individuals" for utilitarianism.

D'Aspremont and Gevers (2002, 495) introduce the axiom of separability as follows: "*What do we do when some individuals are completely unconcerned by the issues at stake, so that their evaluation function remains constant over X? Can the social ranking be affected by the level of their constant evaluation count? If the answer is negative, we express it as an interprofile statement known as Separability axiom."* In other words, separability is the idea that, to estimate if A is socially preferable to B, getting the opinion of the individuals for whom A or B matters is enough, leaving aside the individuals who do not care.

The axiomatization theorems of utilitarianism say, in essence, that on the basis of three or more alternatives, the separability requirement forces the

collective choice to be of the utilitarian type (combining numerical values with the evaluations, and then adding them), or, at the extreme end, of the Rawlsian type (the Maximin). We can then deduce that the other methods, like the ones maximizing the median, break that requirement. That is what can be observed in the examples, like example 2 (the choice of a restaurant).

We note that, unlike in the case of Balinski and Laraki's report (2010, chapter 3), social choice theorists have always been interested in the comparable ordinal evaluations. The theory produced results that are relevant to study choice procedures using this type of evaluation and therefore also relevant for the best median evaluation method. The famous "Arrow**'**s impossibility theorem," in addition to being a classic in popular science, is an extremely important point of reference in these studies, but it is wrong that research has only focused on "Arrowian" social choice, that is based on ranking and not allowing direct interpersonal comparisons.

Majority Judgment as Bucklin Voting

Instead of looking "from the bottom" with the Rawls criterion, we can explain Majority Judgment "from the top" in the following way. As the evaluations can be compared, for each evaluation level (u), starting from the highest, we count the number of individuals who reach u or better, for the different candidates. Logically, when u decreases, all these numbers increase. We stop at the first value of u, let's call it u^*, for which a candidate, call him/her x, satisfies half of the population at level u. This value u^* is the median of the evaluations of x and, since u^* is the first level where we stopped, u^* is the best possible median. Candidate x, who allows this value to be reached, is the choice at the best median, or Maximed. The same misuse of language as explained above qualifies the so-designated candidate as "majority candidate": at the threshold u^*, half of the population is "u^*-*satisfied*" by x.

This idea was discovered and re-discovered several times as a voting rule, deriving the evaluations from the ranks. When one only knows the ranking of the candidates operated by each voter, one takes as the "utility" (either evaluation or satisfaction level) of the candidate x for the voter i the Borda elementary score in the preference of i: 0 for the worst candidate, 1 for the candidate coming just before, etc., until n–1 for the favourite candidate among n.

In the 1980s, Murat Sertel gave several results related to this criterion, which he calls the Majoritarian Compromise (Sertel [1986], Sertel and Yilmaz [1999], Laffond and Lainé [2010], Polat [2013]). Gilbert Basset and Joseph Persky did the same proposal under the name Robust Voting, insisting on the

statistical robustness of the median, that is its insensitivity to extreme values (Basset and Persky [1999]). Steven Brams calls it Fallback Voting (Brams [2008]), but most theoreticians talk about Bucklin voting (Tideman [2006], Brandt et al. [2016]), for a reason that will be explained further. Procedures generalizing this rule are defined in Elkind, Faliszewski, and Slinko (2011) under the name of *M-voting rules*.

MAJORITY JUDGMENT IN PRACTICE

The history of Bucklin Voting

Several voting systems are known as "Bucklin Voting," sharing some common features with Majority Judgment. Unlike Majority Judgment, which uses an exogenous scale of values, Bucklin Voting makes use of candidate rankings (possibly incomplete ones). The main feature it has in common with Majority Judgment is the principle (presented in 4.3) of finding the first candidate that can claim the support of half of the voters. The pure Bucklin system looks for the candidate who is *ranked at least at rank k** by half of the voters; Majority Judgment looks for the first candidate who is *evaluated at least at level u** by half of the population.

These systems were used in the United States for different elections before being abandoned. The history of Bucklin Voting may be instructive for understanding what the effective functioning of Majority Judgment might be, if it was applied, as it shows the real use, for political elections, of a voting method whose goal is to produce a "from the top majority" candidate, as described.

Several American states and cities adopted variations of the Bucklin Voting between 1910 and 1930. I mention here a few essential elements of this history; to read further on this topic, see the dedicated pages on the different American websites promoting Approval Voting,[6] Single Transferable Vote,[7] and Range Voting.[8]

The method seems to have been introduced in order to limit the vote-splitting effects that can lead to the election of a minority candidate. This is indeed a fundamental problem for single-name voting systems, especially with one round only, and one of the reasons that has led to the search for "evaluative voting" type solutions.

6. https://electology.org/, accessed May 3, 2017.
7. http://archive.fairvote.org/, accessed May 3, 2017.
8. www.rangevoting.org/BrownSmallwood.html, accessed May 3, 2017.

The first practical difficulty seems to be that a lot of voters, even under the Bucklin system, voted for only one candidate (bullet voting). Bucklin Voting, as we mentioned, uses the candidates' ranking, partial or not, as an input, not evaluations strictly speaking. By giving only one name, voters, under a Bucklin system, declined to express their opinions in a more detailed way, but also declined to see their votes counted a second time for another candidate. In the case (quite common in the United-States) of an election having only a small number of important candidates, such a behavior can characterize a strategic vote.

It is not clear whether bullet voting simply reflected a habit or if it was the mark of a real strategic behavior. Whatever the case may be, the system was used for several years, and the ranking of at least two or three candidates was sometimes compulsory, specifically to avoid this bullet voting, which appeared to be widespread and was interpreted as "strategic." As far as I know, this voting system was abandoned progressively wherever it was used.

Here is a possible reading of the Bucklin Voting story. The system seemed attractive enough for several cities to adopt it. But the system didn't live up to its promises because, over time, too many voters declined, for strategic reasons, to make use of the options that the system offers (here, ranking the candidates). If this reading is correct, it obviously provides lessons for electoral reforms in general and for Majority Judgment in particular. This point will be discussed in the following section.

The manipulability of Majority Judgment

The "manipulability" of a voting system refers to the fact that the best interest of a voter can be to cast a ballot that doesn't reflect his or her true opinion. The way these manipulations can happen depends on the voting system.

In single-name voting systems, the aim of manipulation is to vote for a candidate for "useful" reasons, instead of casting a "useless" vote for the preferred candidate. With a large number of voters, manipulation by only one person generally has no effect, but the concept of manipulation is better understood as a group phenomenon. In any case, manipulating requires having an idea of what the result of the vote could be (subjective or poll-based anticipation).

The practical importance of the problem of manipulation in relation to politics is a sensitive topic. Indeed, there are several conceptions of what an election is in societal terms, and even in the mind of each voter: to vote is both to influence the result and to express one's opinion, and these two objectives

can result in a trade-off (Blais et al. [2015]). Numerous historical and empirical studies describe the "strategic vote" phenomenon for the commonly used voting rules (Cox [1997]), but for innovative voting rules, the question is still abstract.

Probabilistic models have been used to explore the theoretical possibilities of manipulation; they produced numerical values of "manipulation probabilities." The real relevance of this approach has yet to be proved, but at least it provides a few theoretical landmarks. These studies show a high-degree of manipulability of the evaluating rules, such as Range Voting and Majority Judgment (see Durand [2015]), as well as the Borda Count (Lepelley and Mbih [1994]) and Bucklin Voting (Gehrlein and Lepelley [2013]). Generally, the probabilistic approach in fact provides arguments in favor of the Single Transferable Vote system[9] (Walsh [2010]).

The manipulability of the voting rules can be studied in the laboratory, where the participants are paid only on their results. In this case, strategic behaviors are noticeable when the votes are repeated, because repetition allows the voters, first, to master the mechanics of the voting rules, and second, to understand in which ways they can manipulate or not (Forsythe et al. [1993], Blais, Laslier, and Van der Straeten [2016]). *A contrario*, the concept of manipulation is difficult to comprehend outside the laboratory for voting rules that are not standard or not very intuitive.

In any case, understanding what the "manipulation" of a given voting system is seems essential.

Under Range Voting, the Borda Count, or Bucklin Voting, manipulation is generally the *exaggeration* of one's evaluation, in one way or another, in order to influence the result as much as possible (Laslier and Núñez [2014]). In the case of the "Single Transferable Vote" the situation is very different, and exaggerating one's preferences is usually useless; manipulation is rather a *downgrading* of a candidate that one likes, who can perform quite well but without being able to win. This is probably less natural (Laslier [2016]). Under Approval Voting, it is less clear what the act of manipulation consists of, and, apart from very exceptional cases, it is never in the interest of voters to approve or disapprove of candidates against their preferences (Dutta, Laslier, and De Sinopoli [2006]). What about Majority Judgment?

In the computation that defines the median of a statistical distribution, only the relative positions matter. For example, if candidate A's "majority grade" is *Fair*, my appreciation of A comes only as: higher than, equal to, or lower than *Fair*. Giving *Excellent*, *Very good*, or *Good* won't change anything

9. Definition of the Single Transferable Vote: the voter ranks all the candidates. Then, in an iterative way, the candidate who is the least frequently top-ranked is eliminated, but all ballot papers are kept, so that each ballot paper is transferred, if needed, from its first choice to its second, then its third, etc.

in this case. However, this remark doesn't say much about how to manipulate Majority Judgment voting, a voting process that chooses, among several candidates, the one who has the best median evaluation.

As for any voting rule, the "strategic" vote depends on anticipations. If I anticipate my vote having an impact, it is for deciding between two candidates. With Majority Judgment, this happens most of the time when two candidates, let's say A and B, have the same majority grade, let's say *Fair*. In this case, Majority Judgment decides between them by counting the number of voters who gave higher or lower evaluations than *Fair*. All the voters who prefer A to B will want to award a grade higher than *Fair* to A and a lower grade to B, and vice versa for the ohers.

Majority Judgment is then manipulated following the same logic as the Borda Count or Range Voting, that is the intuitive logic to "influence in the direction that is favorable to me" by exaggerating my preferences: if the election is decided between A and B, to get A elected I will have to increase the value of A and to depreciate B.

The issue of the "strategic" vote calls for a number of comments.

First, the reasoning we just followed applies the same way to the Borda rule, Range Voting, and Majority Judgment, but doesn't apply to Approval Voting, which only uses two grades. In the case of Approval Voting, the strategic prescription simply entails approving the candidate one prefers and not the other (Laslier [2009]).

Secondly, specialists have noticed for a long time that the procedure that consists in summarizing one (and only one) distribution by its median has excellent strategic properties (Dummett and Farquharson [1961], Moulin [1980]). The canonical example is the following. In a world of pure left-right politics, if voters, instead of voting or making evaluations, could only indicate their own position on the left-right axis, and if the "elected" position was the median of the announced positions, then each person would have no better choice than to indicate their real position. Balinski and Laraki [2011] use this argument to justify the use of the median, and it is actually valid when the evaluation concerns one candidate and one only. It can be described precisely as follows. If I am strategic, if there is only one "candidate" who is going to be collectively evaluated by the median of the individual evaluations, and if my aim is the collective evaluation being the closest as possible to my personal evaluation, then I have no better choice than to "honestly" provide my real evaluation of this candidate. Unfortunately, the argument does not apply to the case of several candidates.

Thirdly, Majority Judgment is different from Range Voting by virtue of the curious following property, directly related to the definition of the median. Majority Judgment doesn't consider the intensity of preferences other than by comparison with the "majority grade." If the race is on between A and

B with the grade *Fair*, then the opinion of the voters attributing *Excellent* to A and only *Good* for B—and expressing that opinion—is not taken into account, because both grades are above the majority grade *Fair*. This point, linked to the ordinal feature, contrasts with Range Voting: between A and B, if I give a higher rank to A than to B, the difference is taken into account in the same way, whether these ranks are 4 and 5, or 0 and 1. Example 1 illustrates this detail to the point of caricature. Under the rules of Majority Judgment, it is not good to be optimistic in general (or pessimistic in general) and to express it. For my vote to have some influence between A and B, one of my judgments will have to be above the majority judgment, the other below.

The testimonies show that this point can be the source of a misunderstanding of the method by the people who tried Majority Judgment out on the websites that promote it. Since I am provided with the option to say that, for me, A is *Excellent* and B is only *Good*, it seems that, in doing so, I lean more in favor of A than in favor of B. Yet in a way that is perhaps counterintuitive, this is not the case with Majority Judgment.

CONCLUSION

This article reviewed the following question: what do we do when we choose among several candidates using the best median criterion? Its goal was to clarify some points of confusion, whose origin appears to be the wrong extrapolation of some properties of the median to properties of the choice comparing the medians.

An initial phase of confusion arises due to the fact that the aspect we can call "majoritarian" of the notion of median ("half of the population higher, half lower") is to be found in the choice between several medians. As shown above, this is not the case, and the choice of the best median actually leads to something totally different, which does not seem to be very attractive from a democratic point of view.

A second factor that causes confusion is the transposition of the very real statistical "robustness" of the median to a strategic robustness of the social choice function that chooses a candidate with the best median. Majority Judgment is not a method whose goal is to summarize the evaluations of different people regarding a single candidate. Instead, it aims to be a voting system, and thus seeks to choose one candidate among several. The difference between the two is substantial. The problem is one of aggregation.

We observed that, as a collective choice method, the principle that forms the basis of Majority Judgment amounts to determining which half-population is to be neglected in order to satisfy—in a very specific way—the

other half. The best median principle thus appears to be in contradiction with the definition of democracy as the participation of everyone; not half of the population plus one. This is probably because it is based on a not very intuitive principle that this system tends to produce startling results, theoretically and practically. Moreover, providing the highest possible level of satisfaction to the half-population that is the easiest to satisfy while neglecting the other half is not what is achieved by the rules of choice, which theory refers to as "majoritarian," and which respect the Condorcet principle. These rules, on the contrary, tend to promote consensual solutions, particularly in political environments.

REFERENCES

ASPREMONT, C. d' and GEVERS, L. [1977]. "Equity and the Informational Basis of Collective Choice," *The Review of Economic Studies*, 44 (2): 199–210.

ASPREMONT, C. d' and GEVERS, L. [2002]. "Social Welfare Functionals and Interpersonal Comparability." In ARROW, K. J., SEN, A. K. and SUZUMURA, K. (eds.), *Handbook of Social Choice and Welfare*. Amsterdam : Elsevier, p. 459–541.

BALINSKI, M. and LARAKI, R. [2011]. *The Majority Judgment*. Cambridge (Mass.): The MIT Press.

BASSET, G. W. Jr. and PERSKY, J. [1999]. "Robust Voting," *Public Choice*, 99 (3–4): 299–310.

BLACKORBY, C., BOSSERT, W. and DONALDSON, D. [2002]. "Utilitarianism and the Theory of Justice." In ARROW, K. J., SEN, A. K. and SUZUMURA K. (eds.), *Handbook of Social Choice and Welfare*. Amsterdam: Elsevier, p. 543–596.

BLAIS, A., LASLIER, J.-F., POINAS, F. and VAN DER STRAETEN, K. [2015]. "Citizens' Preferences about Voting Rules: Self-Interest, Ideology, and Sincerity," *Public Choice*, 164 (3–4): 423–442.

BLAIS, A., LASLIER, J.-F. and VAN DER STRAETEN, K. (eds.) [2016]. *Voting Experiments*. Berlin: Springer.

BOSSERT, W. and WEYMARK, J. [2004]. "Utility in Social Choice." In BARBERA, S., HAMMOND, P. and SEIDL, C. (eds.), *Handbook of Utility Theory*. Boston: Kluwer, 2: 1099–1177.

BRAMS, S. J. [2008]. *Mathematics and Democracy*. Princeton: Princeton University Press.

BRANDT, F., CONITZER, V., ENDRISS, U., LANG, J. and PROCACCIA, A. (eds.) [2016]. *Handbook of Computational Social Choice*. New York: Cambridge University Press.

COX, G. W. [1997]. *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Cambridge (Mass.): Cambridge University Press.

DEBREU, G. [1960]. "Topological Methods in Cardinal Utility Theory." In ARROW, K. J., KARLIN, S. and SUPPES, P. (eds.), *Mathematical Methods in the Social Sciences*. Palo Alto: Stanford University Press, p. 16–26.

DESCHAMPS, R. and GEVERS, L. [1978]. "Leximin and Utilitarian Rules: A Joint Characterization," *Journal of Economic Theory*, 17: 143–163.

DUMMETT, M. and FARQUHARSON, R. [1961]. "Stability in Voting," *Econometrica*, 29 (1): 33–43.

DURAND, F. [2015]. Vers des modes de scrutin moins manipulables, PhD thesis in computer sciences, Université Pierre et Marie Curie, Paris.

DUTTA, B., LASLIER, J.-F. and DE SINOPOLI, F. [2006]. "Approval Voting: Three Examples," *International Journal of Game Theory*, 35 (1): 27–38.

ELKIND, E., FALISZEWSKI, P. and SLINKO, A. [2011]. "Homogeneity and Monotonicity of Distance-Rationalizable Voting Rules," *AAMAS 2011: Section C6 – Voting Protocols*, International Foundation for Autonomous Agents and Multiagent Systems, p. 821–828, http://www.ifaamas.org/Proceedings/aamas2011/resources/fullpapers.html#s-C6/.

ELSTER, J. and ROEMER, J. (eds.) [1991]. *Interpersonal Comparisons of Well-Being*. Cambridge (Mass.): Cambridge University Press.

ENELOW, J. M. and HINICH, M. J. [1984]. *The Spatial Theory of Voting: An Introduction*, Cambridge (Mass.): Cambridge University Press.

FELSENTHAL, D. and MACHOVER, M. [2008]. "The Majority Judgment Voting Procedure: A Critical Evaluation," *Homo Oeconomicus*, 25 (3–4): 319–334.

FELSENTHAL, D. and NURMI, H. [2017]. *Monotonicity Failures Afflicting Procedures for Electing a Single Candidate*. Berlin: Springer.

FLEURBAEY, M. [1996]. *Théories économiques de la justice*. Paris: Economica.

FLEURBAEY, M. and Hammond, P. [2004]. "Interpersonally Comparable Utility." In BARBERA S., HAMMOND P. and SEIDL, C. (eds.), *Handbook of Utility Theory*. Boston: Kluwer, 2: 1179–1285.

FORSYTHE, R., MYERSON, R. B., RIETZ, T. A. and WEBER, R. J. [1993]. "An Experiment on Coordination in Multicandidate Elections: The Importance of Pools and Election Histories," *Social Choice and Welfare*, 10 (3): 223–247.

GEHRLEIN, W. and LEPELLEY, D. [2003]. "On Some Limitations of the Median Voting Rule," *Public Choice*, 117 (1–2): 177–190.

GEHRLEIN, W. and LEPELLEY, D. [2011]. *Voting Paradoxes and Group Coherence*. Berlin: Springer.

IGERSHEIM, H., BAUJARD, A. and LASLIER, J.-F. [2016] "La question du vote. Expérimentations en laboratoire et In Situ," *Actualités économiques*, 92 (1–2), mars–juin 2016: 151–189.

KRANTZ, D. H., LUCE, D., SUPPES, P. and TVERSKY, A. [1971]. *Foundations of Measurement*. Vol. 1: *Additive and Polynomial Representations*. New York: Academic Press.

LAFFOND, G. and LAINÉ, J. [2010]. "Does Choosing Committees from Approval Balloting Fulfill the Electorate's Will?" In LASLIER, J.-F. and SANVER, M. R. (eds.), *Handbook on Approval Voting*. Berlin: Springer,

p. 125–150.

LASLIER, J.-F. [2009]. "The Leader Rule: A Model of Strategic Approval Voting in a Large Electorate," *Journal of Theoretical Politics*, 21 (1): 113–136.

LASLIER, J.-F. [2012]. "On Choosing the Alternative with the Best Median Evaluation," *Public Choice*, 153 (3): 269–277.

LASLIER, J.-F. [2016]. "Heuristic Voting under the Alternative Vote: The Efficiency of "Sour Grapes" Behavior," *Homo Oeconomicus*, 33 (1–2): 57–76.

LEPELLEY, D. and MBIH, B. [1994]. "The Vulnerability of Four Social Choice Functions to Coalitional Manipulation of Preferences," *Social Choice and Welfare*, 11 (3): 253–266.

MACÉ, A. [2015]. "Voting with Evaluations: When Should We Sum, What Should We Sum?" *Working Paper*, Aix-Marseille School of Economics.

MOULIN, H. [1980]. "On Strategy-Proofness and Single Peakedness," *Public Choice*, 35 (4): 437–455.

MUELLER, D., FACCHINI, F., FOUCAULT, M., FRANÇOIS, A., MAGNI-BERTON, R. and MELKI, M. [2010]. *Choix publics. Analyse économique des décisions publiques*. Bruxelles: De Boeck.

NÚÑEZ, M. and LASLIER, J.-F. [2014]. "Preference Intensity Representation: Strategic Overstating in Large Elections," *Social Choice and Welfare*, 42 (2): 313–340.

PIVATO, M. [2013]. "Formal Utilitarianism and Range Voting," *Mathematical Social Sciences*, 67: 50–56.

POLAT, A. O. [2013]. Median Rule and Majoritarian Compromise, Phd thesis in economics, Department of Economics, Bilkent University, Ankara.

RAWLS, J. [1971]. *A Theory of Justice*. Cambridge (Mass.): Harvard University Press.

RAWLS, J. [1974]. "Some Reasons for the Maximin Criterion," *The American Economic Review*, 64 (2): 141–146.

ROEMER, J. [1998]. "Why the Poor Do Not Expropriate the Rich: An Old Argument in New Garb," *Journal of Public Economics*, 70 (3): 399–424.

SERTEL, M. [1986]. *Lecture Notes in Microeconomics*, Istanbul Bogaziçi University, not published.

SERTEL, M. and YILMAZ, B. [1999]. "The Majoritarian Compromise Is Majoritarian-Optimal and Subgame-Perfect Implementable," *Social Choice and Welfare*, 16 (4): 615–627.

TIDEMAN, N. [2006]. *Collective Decisions and Voting: The Potential for Public Choice*. Aldershot: Ashgate.

WALSH, T. [2010]. "An Empirical Study of the Manipulability of Single Transferable Voting." In COELHO, H., STUDER, R. and WOOLDRIDGE,

M. (eds.), *ECAI 2010: 19th European Conference on Artificial Intelligence, Vol. 215: Frontiers in Artificial Intelligence and Applications*. Amsterdam: IOS Press, p. 257–262.

ZAHID, M. A. [2009]. "Majority Judgement Theory and Paradoxical Results," *Mimeo*.