Université Paris-Dauphine - PSL

Cycle Pluridisciplinaire d'Etudes Supérieures

Social choice theory, Arrow's impossibility theorem and majority judgment

Victor Elie

supervised by
Miquel Oliu Barton

June 25, 2017

Contents

1 Introduction

For centuries people have tried to aggregate individual preferences into a collective decision. Many methods arose, and many were used throughout History by different societies and civilizations on scales varying from small groups of people to entire nations. Yet, no one ever agreed on an optimal system. In recent centuries social choice theory started being formalized, and powerful results were proved, impossibility theorems, which proved that paradoxes were inevitable in voting systems.

The objective of this paper is to get into the basics of social choice theory, prove and understand *Arrow's impossibility theorem*, and to explore a recently developed voting system: *majority judgment*.

2 Social Choice Theory

2.1 Definitions

Social choice theory analyses the extent to which individual preferences can be aggregated into social preferences in the most satisfactory manner.

Let X be a finite set of alternatives and let there be I agents indexed by $i = 1, \ldots, I$. Every agent i has a rational preference relation $\succeq_i$ defined on X. Strict preference is denoted by $\succ_i$ and the indifference relation is denoted by $\sim_i$. We denote by $\mathscr{R}$ the set of all possible rational preference relations on X and $\mathscr{P}$ the set of all possible strict rational preference relations on X. We define an agent's rational preference relation as a ranking of the elements of X corresponding to his preferences.

Definition 1. A social welfare functional defined on a given subset $\mathscr{A} \subset \mathscr{R}^I$ is a function $F : \mathscr{A} \to \mathscr{R}$ that assigns a rational preference relation $F(\succeq_1, \ldots, \succeq_I) \in \mathscr{R}$ interpreted as the social preference relation, to any profile of individual preference relations $(\succeq_1, \ldots, \succeq_I) \in \mathscr{A}$.

For any profile $(\succeq_1, \ldots, \succeq_I)$, we denote by $F_p(\succeq_1, \ldots, \succeq_I)$ the strict preference relation derived from $F(\succeq_1, \ldots, \succeq_I)$.

Definition 2. A social welfare functional $F : \mathscr{A} \to \mathscr{R}$ is *Paretian* if for any pair of alternatives $\{x, y\} \subset X$ and any preference profile $(\succeq_1, \ldots, \succeq_I) \in \mathscr{A}$, we have that x is socially preferred to y, that is $xF_p(\succeq_1, \ldots, \succeq_I)y$, whenever $x \succ_i y$ for every i.

Definition 3. The social welfare functional $F : \mathscr{A} \to \mathscr{R}$ satisfies the *pairwise independence condition* (or the *independence of irrelevant alternatives condition*) if the preference between two alternatives $\{x, y\} \subset X$ depends only on the profile of individual preferences over the same alternatives.

Formally, for any pair of alternatives $\{x, y\} \subset X$ and for any pair of preference profiles $(\preceq_1, \ldots, \preceq_I) \in \mathscr{A}$ and $(\preceq_1^{'}, \ldots, \preceq_I^{'}) \in \mathscr{A}$ with the property that for every i,

$$x \preceq_i y \Leftrightarrow x \preceq_i^{'} y \text{ and } y \preceq_i x \Leftrightarrow y \preceq_i^{'} x$$

we have that,

$$xF(\preceq_1, \dots, \preceq_I)y \Leftrightarrow xF(\preceq_1^{'}, \dots, \preceq_I^{'})y$$

and

$$yF(\preceq_1, \dots, \preceq_I)x \Leftrightarrow yF(\preceq_1^{'}, \dots, \preceq_I^{'})x$$

Example 1 (The specific case of two alternatives)**.** We place ourselves in the situation where there are only two alternatives x and y. The data of our problem is the relative preference of each agent $i \in [\![1, I]\!]$ over x and y. We can describe this data by a profile

$$(\alpha_1, \dots, \alpha_I) \in \mathbb{R}^I,$$

where α_i takes the value 1, 0, -1 which correspond respectively to $x \prec_i y$, $x \sim_i y$, $y \prec_i x$.

Let $(\beta_1, \dots, \beta_I)$ a vector of non negative numbers, we can define the following social welfare functional:

$$F(\alpha_1, \dots, \alpha_I) = \text{sign} \sum \beta_i \alpha_i$$

where $\text{sign}(a) = 1$ if $a > 0$, $\text{sign}(a) = 0$ if $a = 0$, $\text{sign}(a) = -1$ if $a < 0$.

If all β_i are equal to 1, we are in the case of majority voting, indeed $F(\alpha_1, \dots, \alpha_I) = 1$ if and only if the number of agents preferring x to y is strictly superior to the number of agents preferring y to x. Similarly $F(\alpha_1, \dots, \alpha_I) = -1$ if and only if the number of agents preferring y to x is strictly superior to the number of agents preferring x to y.

If there is a j in $[\![1, I]\!]$ such that $\beta_j > 0$ and $\forall i \neq j$ $\beta_i = 0$, j is what we call a dictator. More generally, in the case where there are two alternatives, we say that a social welfare functional is *dictatorial* if there exists an agent h called a *dictator* such that for any profile $(\alpha_1, \dots, \alpha_I)$, $\alpha_h = 1$ implies $F(\alpha_1, \dots, \alpha_I) = 1$ and $\alpha_h = -1$ implies $F(\alpha_1, \dots, \alpha_I) = -1$.

Example 2 (Borda count)**.** Let $\succeq_i \in \mathscr{R}$. First of all we suppose that there are no indifference relations on in $\succeq_i$. $\forall x \in X$ we define $c_i(x)$ which is equal to the ranking of x in the preference relation $\succeq_i$. If indifference is possible then $c_i(x)$ is the average of rank of the alternatives indifferent to x.

For any profile $(\succeq_1, \dots, \succeq_I) \in \mathscr{R}$, we define $F(\succeq_1, \dots, \succeq_I)$ such that $xF(\succeq_1, \dots, \succeq_I)y$ if and only if $\Sigma_i c_i(x) \leqslant \Sigma_i c_i(y)$.

We can easily see that the Borda count is Paretian, indeed if $x \succ_i y$ for all i, $c_i(x) < c_i(y)$ for all i, therefore $\Sigma_i c_i(x) < \Sigma_i c_i(y)$. On the other hand it doesn't satisfy the pairwise independence condition, which we can see with this example. Suppose there are two agents and three alternatives $\{x, y, z\}$. For the preferences

$$x \succ_1 y \succ_1 z$$

$$y \succ_2 z \succ_2 x$$

we have that $\Sigma_i c_i(x) = 4$ and $\Sigma_i c_i(z) = 5$, x is thus socially preferred to z. For the preferences

$$y \succ_1 x \succ_1 z$$

$$z \succ_2 y \succ_2 x$$

we have that $\Sigma_i c_i(x) = 5$ and $\Sigma_i c_i(z) = 4$, z is thus socially preferred to x. But the relative ranking of z and x has not changed, the pairwise independence condition is not satisfied. Other similar examples with more alternatives and more agents can be found.

Example 3 (Condorcet paradox)**.** We suppose there are three agents and three alternatives $\{x, y, z\}$, we try to use a majority voting social welfare functional. If we have the following preferences:

$$x \succ_1 y \succ_1 z$$

$$y \succ_2 z \succ_2 x$$

$$z \succ_3 x \succ_3 y$$

By majority, we have x preferred to y, but also y preferred to z and z preferred to x. We have a cyclical preference which contradicts *transitivity*. The definition of a social welfare functional implies transitivity, indeed the function's output is a ranking of alternatives, thus if a is preferred to b and b is preferred to c, we must have that a is preferred to c.

This is closely linked to the fact that in this theory, individuals are only defined by a ranking that they make of alternatives. In fact we could imagine other ways of defining individuals, like by a grade they would give to an alternative, but we will go into further details later.

This situations shows us that we have a paradox when using majority voting in this specific case. This brings us to the central theorem in social choice theory which tells us that the Condorcet paradox is not linked to specific characteristics of majority voting but actually with the *pairwise independence condition*.

2.2 Arrow's impossibility theorem

Theorem 1 (Arrow's impossibility theorem)**.** *Suppose that the number of alternatives is at least three and that we have either $\mathscr{A} = \mathscr{R}^I$ or $\mathscr{A} = \mathscr{P}^I$. Then every social welfare functional $F : \mathscr{A} \to \mathscr{R}$ that is Paretian and satisfies the pairwise independence condition is dictatorial in the following sense: there is an agent h such that any pair of alternatives $\{x, y\} \subset X$ and for any profile $(\succeq_1, \ldots, \succeq_I) \in \mathscr{A}$, we have that x is strictly socially preferred to y whenever $x \succ_h y$.*

Many proofs of Arrow's impossibility theorem exist, there are many approaches to the problem and different concepts to reach the final solution. In this paper I have decided to follow a proof published by John Geanakoplos, which bases itself on the following Lemma. For better understanding, we will place ourselves in the case $\mathscr{A} = \mathscr{P}^I$ so that all preferences are strict. The general case proves itself similarly.

Proposition 1 (Extremal Lemma)**.** *Let alternative b be chosen arbitrarily. At any profile where b is either first or last of each agent's ranking, society must as well.*

A_1	A_2	A_3	A_4	A_5	$\dots$	A_n
b	.	b	b	b		c
.	.	.	.	.		.
.	c	c	.	.		.
c	a	.	.	.		.
.	.	.	.	.		a
a	.	.	c	.		.
.	.	a	a	.		.
.	.	.	.	c		.
.	b	.	.	a		b

Table 1: Preference scheme such that b is in extreme positions, and c is always preferred to a

Proof. Let $\succ$ be the social preference $F(\succ_1, \dots, \succ_I)$. Suppose to the contrary that for such a profile we have a and c such that $a \succ b$ and $b \succ c$. By the pairwise independence condition, this remains true if each agent i changes his preferences such that $c \succ_i a$. Indeed b is always in an extreme position therefore its relation to a and c cannot be altered by such a change. By unanimity (Pareto condition) $c \succ a$ and by transitivity $a \succ c$. This is a contradiction.

□

Proof of the theorem. We first prove that for an alternative b, there exists an agent who can be called *extremely pivotal* in the sense that by changing his vote at some profile he can move b from the bottom of the ranking to the very top.

Indeed, let each voter put b at the very bottom of their ranking, by unanimity the social welfare functional must as well. If each agent successively puts b on the top of their ranking, we call $n*$ the first agent whose change will put b from the bottom to the top of the social ranking (indeed we know b will be at the top by the previous Lemma, and that this agent exists since when all agents put b at the top, the social functional must as well). We call profile I the profile just before the agent $n*$ has put b to the top of his ranking and profile II the profile just after.

A_1	A_2	$\dots$	A_{n^*-1}	A_{n^*}	$\dots$	A_n
b	b		b	.		.
.	.		.	.		.
.	.		.	.		.
.	.		.	.		.
.	.		.	.		.
.	.		.	.		.
.	.		.	.		.
.	.		.	.		.
.	.		.	b		b

Table 2: Profile I

A_1	A_2	$\dots$	A_{n^*-1}	A_{n^*}	$\dots$	A_n
b	b		b	b		.
.	.		.	.		.
.	.		.	.		.
.	.		.	.		.
.	.		.	.		.
.	.		.	.		.
.	.		.	.		.
.	.		.	.		.
.	.		.	.		b

Table 3: Profile II

A_1	A_2	$\dots$	A_{n^*-1}	A_{n^*}	$\dots$	A_n
b	b		b	.		.
.	.		.	.		.
.	.		.	a		.
.	.		.	.		.
.	.		.	b		.
.	.		.	.		.
.	.		.	c		.
.	.		.	.		.
.	.		.	.		b

Table 4: Profile III

We now show that $n*$ is a dictator over any pair a c not involving b. We choose two alternatives a and c. We construct profile III by letting agent $n*$ put a above b and b above c so that $a \succ_{n*} b \succ_{n*} c$ and we let all other agents place arbitrarily place a and c while leaving b in its extreme position. By independence of irrelevant alternatives we have $a \succ b$ since the relative positions of a and b are the same as in profile I, where b is socially ranked at the bottom. We also have $b \succ c$ since the relative positions of b and c are the same as in profile II, where b is socially ranked at the top. By transitivity, the social welfare functional must give $a \succ c$. By independence of irrelevant alternatives we can conclude that the social preference over a and c must always agree with $n*$. Indeed we first reach the conclusion for any situation where $a \succ_{n*} b \succ_{n*} c$. We can then place b anywhere in the ranking while keeping $a \succ c$, by the independence of irrelevant alternatives.

We can conclude by showing that $n*$ is also dictator for any pair a b. We take a third alternative c that takes the role of b in the previous paragraphs. There must be an agent $n(c)$ who is a dictator over every pair $\alpha\beta$ not involving c, like for instance a and b. But agent $n*$ can change the a b relative ranking (we have an example with profiles I and II). We conclude that $n(c)$ is $n*$.

$\square$

2.3 Interpretations and limits of Arrow's theorem

At first sight Arrow's theorem seems to mean that there are no good voting systems, in the sense that they must respect certain hypotheses which can be seen as "reasonable". It is therefore useful to explore these different hypotheses and to see if they are necessarily essential for a voting system.

To begin with, Arrow's impossibility theorem treats the case of ranked voting electoral systems. An agent's opinion is represented only by his relative preference over given alternatives. We will explore later in this paper other voting systems which do not use ranks of alternatives. Indeed if the agent is asked to give a judgment between: excellent, very good, good, fair, poor, to reject. The social welfare functional input will gather more information than in a ranking. The reason why is that from a list of judgments over alternatives we can sometimes deduce a preference ranking, but the opposite is never possible.

We can also find a way out of this paradox through restricting the domain of definition of the function. The main example of such a restriction is Black's condition. We make the hypothesis that we can represent the alternatives on the horizontal axis such that the preference curves become single-peaked. We can see an example in Figure 1.

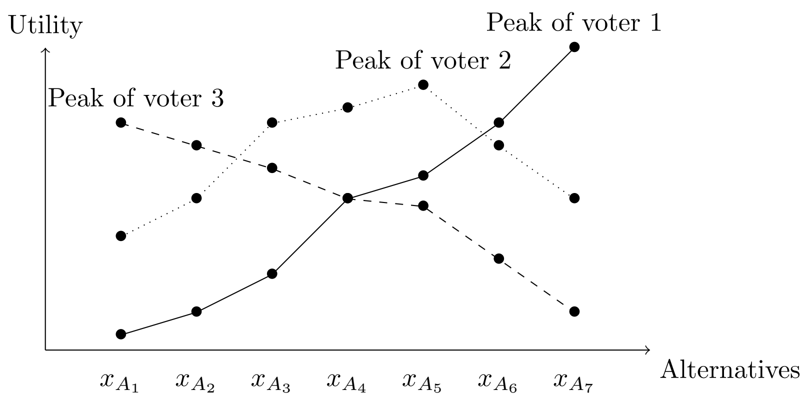

Figure 1: Single-peaked preferences

"Utility" which is represented on the vertical axis corresponds the position of an alternative in an agent's ranking. As an example we can think of the political spectrum as it is usually represented: from far-left to far-right. In this case the peak corresponds to the voter's political position and the decrease on each side of the peak correspond to the fact that the further away a candidate is from you on the political spectrum, the less you agree with him on various issues, therefore the less you wish to see him elected.

We call a Condorcet-winner an alternative which wins every head to head confrontation (i.e. it is preferred individually to any other alternative). It is a very strong position, which is why the concept of Condorcet-winner has been central throughout the development of Social Choice Theory.

The main issue of Social Choice Theory is that there is not always a Condorcet-winner. However, under Black's condition (i.e. with single-peaked preferences) there is always a Condorcet winner. Indeed, let agent A be the voter whose preferred candidate C^* is the median of the preferred candidates on the spectrum, we call A the *median voter*. Then the Condorcet-winner is C^*. If C is on the left of C^*, all voters whose peak was on the right of C^* vote for C to which we can add the *median voter*, C has therefore a majority. The result is identical if C is on the right of C^*. In the end C^* wins each head to head confrontation.

Nevertheless, empirical works have shown that the hypothesis of having single-peaked preferences is rarely satisfied in one dimension. In higher dimensions, which gives a bigger degree of freedom, the hypothesis can sometimes be satisfied.

3 Majority judgment

3.1 Social Grading Functions

Definition 4. A *Social Grading Function* is a function F that assigns to any profile Φ − any set of grades in the language Λ − one single grade in the same language for every alternative:

$$F : \Lambda^{m\times n} \to \Lambda^n$$

The following axioms sum up the minimal requirements that are needed by a social grading function (SGF).

Axiom 1 (Neutrality)**.** *The SGF judges all alternatives equally.*

Axiom 2 (Anonymity)**.** *The judges have the same influence on the grades.*

Axiom 3 (Unanimity)**.** *F is unanimous: if a competitor is given an identical grade α by every judge, then F assigns him the grade α.*

Axiom 4 (Monotonicity)**.** *F is monotonic: if two inputs Φ and Φ' are the same except that one or more judges give higher grades to a competitor I in Φ than in Φ', then $F(\Phi)(I) \succeq F(\Phi')(I)$.*

Axiom 5 (Independance of irrelevant alternatives in grading)**.** *If the lists of grades assigned by the judges to a competitor I in two profiles Φ and Φ' are the same, then $F(\Phi)(I) = F(\Phi')(I)$.*

Definition 5 (Aggregation function)**.** These axioms enable us to define a function $f : \Lambda^n \to \Lambda$ that transforms the grades given to one alternative into a final grade. For that, f needs to satisfy the following properties:

- anonymity: $f(\dots,\alpha,\dots,\beta,\dots) = f(\dots,\beta,\dots,\alpha,\dots)$.
- unanimity: $f(\alpha,\alpha,\dots,\alpha) = \alpha$
- monotonicity:

$$\alpha_j \succeq \beta_j \, \forall j \implies f(\alpha_1,\dots,\alpha_j,\dots,\alpha_n) \succeq f(\alpha_1,\dots,\beta_j,\dots,\alpha_n)$$

and

$$\alpha_j \succ \beta_j \, \forall j \implies f(\alpha_1, \ldots, \alpha_n) \succ f(\beta_1, \ldots, \beta_n)$$

Axiom 6. *F and its aggregation function are continuous.*

3.2 General principle

Majority judgment is a single-winner voting system proposed by Michel Balinski and Rida Laraki. It is based on a social grading function defined on a given language of grades Λ (e.g. in the Danish school system the grades belong to the set $\{0, 3, 5, 6, 7, 8, 9, 10, 11, 13\}$). This vote leads to a grade for each alternative, the *majority-grade*, and a ranking of alternatives: the *majority-ranking*.

If the number of agents is odd, the *majority-grade* α of an alternative is the median of the grades of the candidate. If the number of agents is even, the *majority-grade* α is the lowest of the two grades in the middle interval. For example in a sports competition if there are 10 judges who give the grades $\{3, 3, 5, 6, 6, 8, 9, 9, 9, 11\}$ to an alternative, the middle interval is $\{6, 7, 8\}$, thus the majority-grade is 6. We can understand the majority-grade as the highest grade approved at least by a majority of the agents.

This first step enables us to distinguish the alternatives by giving a majority-grade, now we will see how to rank alternatives which have the same majority-grade. To compare two alternatives, we firstly drop the majority-grade in the set of grades and find the new one out of the new set. If A's grades are $\{7, 9, 9, 11, 11\}$ and B's grades are $\{8, 9, 9, 10, 11\}$. They both have $\alpha = 9$. We drop α, the new grades are respectively $\{7, 9, 11, 11\}$ $\{8, 9, 10, 11\}$, they both have 9 as what we call their second majority-grade. We repeat the process to find 10 for B and 11 for A. An alternative's *majority-value* is the sequence of his first majority-grade, second majority-grade, up to the nth majority-grade, where n is the number of agents. The majority-values summarize the results of an election.

In the case where there is a great number of agents the system works a little bit differently. Due to the great number of agents we can safely assume that the majority-grade of each alternative will be the median of his grades. This median is both the highest grade approved by a majority and the lowest grade approved by a majority. To obtain the *majority-ranking* we only need three pieces of information:

- p, the percentage of grades better than the majority-grade.
- α, the alternative's majority-grade.
- q, the percentage of grades worse than the alternative's majority-grade.

We call *majority-gauge* the triple (p, α, q). If $p > q$ then the alternative's majority-grade can be completed by a $(+)$. If to the opposite we have $p < q$ the the majority-grade can be completed by a $(-)$. A *majority-grade* $+$ is ahead by a *majority-grade* $-$. We can see that this is equivalent to the precedent situation in which when two alternatives had the same majority-grade, we dropped them until we found two different majority-grades. Between two majority-grade$+$'s, the one having the higher percentage of grades better than the majority-grade is ahead of the other. Out of two majority-grade$-$'s, the one having the higher percentage of grades worse than the majority-grade is behind the other.

An example of this voting system is presented in *Majority-Judgment*, published by Rida Laraki and Michel Balinski. During the 2007 french presidential elections they asked a group of citizens to vote with this voting system. They asked each voter to attribute to each candidate one of the following grades: "Très bien", "Bien", "Assez Bien", "Passable", "Insuffisant", "A rejeter". Attributing no grade is interpreted as "A rejeter". The following table gives the results of the vote.

Candidate	*Excellent*	*Very Good*	*Good*	*Acceptable*	*Poor*	*To Reject*	*No Grade*
Bayrou	13.6 %	30.7 %	25.1 %	14.8 %	8.4%	4.5%	2.9%
Royal	16.7 %	22.7 %	19.1 %	16.8 %	12.2 %	10.8 %	1.8%
Sarkozy	19.1 %	19.8 %	14.3 %	11.5 %	7.1%	26.5 %	1.7%
Voynet	2.9%	9.3%	17.5 %	23.7 %	26.1 %	16.2 %	4.3%
Besancenot	4.1%	9.9%	16.3 %	16.0 %	22.6 %	27.9 %	3.2%
Buffet	2.5%	7.6%	12.5 %	20.6 %	26.4 %	26.1 %	4.3%
Bové	1.5%	6.0%	11.4 %	16.0 %	25.7 %	35.3 %	4.1%
Laguiller	2.1%	5.3%	10.2 %	16.6 %	25.9 %	34.8 %	5.3%
Nihous	0.3%	1.8%	5.3%	11.0 %	26.7 %	47.8 %	7.2%
de Villiers	2.4%	6.4%	8.7%	11.3 %	15.8 %	51.2 %	4.3%
Schivardi	0.5%	1.0%	3.9%	9.5%	24.9 %	54.6 %	5.8%
Le Pen	3.0%	4.6%	6.2%	6.5%	5.4%	71.7 %	2.7%

Table 5: Majority judgment results in three precincts of Orsay, April 22 2007

We can then determine the majority-gauge of each candidate and find the final ranking which is presented in Table 3.

	Candidate	p	α	q	Official vote of participants	National vote
1st	Bayrou	44.3 %	*Good* $+$	30.6 %	25.5 %	18.6 %
2nd	Royal	39.4 %	*Good* $-$	41.5 %	29.9 %	25.9 %
3rd	Sarkozy	38.9 %	*Good* $-$	46.9 %	29.0 %	31.2 %
4th	Voynet	29.7 %	*Acceptable* $-$	46.6 %	1.7%	1.6%
5th	Besancenot	46.3 %	*Poor* $+$	31.2 %	2.5%	4.1%
6th	Buffet	43.2 %	*Poor* $+$	30.5 %	1.4%	1.9%
7th	Bové	34.9 %	*Poor* $-$	39.4 %	0.9%	1.3%
8th	Laguiller	34.2 %	*Poor* $-$	40.0 %	0.8%	1.3%
9th	Nihous	45.0 %	*To Reject*	$-$	0.3%	1.1%
10th	de Villiers	44.5 %	*To Reject*	$-$	1.9%	2.2%
11th	Schivardi	39.7 %	*To Reject*	$-$	0.2%	0.3%
12th	Le Pen	25.7 %	*To Reject*	$-$	5.9%	10.4 %

Table 6: Majority judgment results

We can see that in this example the winner of majority judgment voting is not the winner in the two-round system. This is linked to the fact that majority judgment measures how much voters agree with a candidate, and will therefore privilege the candidate who is the most consensual. This is why François Bayrou, a centrist, wins the election in the case of Majority Judgment voting.

3.3 Advantages and shortcomings

This system is firstly interesting because the input is not a ranking but a grading system. In many situations we can think of reasons why the term "preference-ranking" is inappropriate. One can prefer an alternative A but still would rather vote for B because he is best suited for the job. A judge does not judge regarding his personal preferences but according to the Law. This is why even if a ranking is a necessary output of a social welfare functional it is not a necessary input.

Majority judgment escapes Arrow's impossibility theorem under a certain assumption. If we consider that every agent will give his grades on an absolute scale we can claim that the majority-gauge of each candidate is completely independent of the others, so that the system respects all the hypothesis of Arrow's theorem but is not a dictatorship. However some empirical studies suggest that this assumption is often false.

Another problem arises depending on the situation of the social choice. Let there be one thousand alternatives and ten voters. Having the choice between six possible grades will most probably not provide enough information to distinguish all alternatives. Asking each voter a ranking of alternatives would provide much more information. Majority Judgment thus wouldn't work well in situations in which there are too many alternatives.

Finally, a pathological example was found to illustrate the following problem: if a voter gives two alternatives grades on the same side of their majority-grade, he will have an equal impact on their grades, and thus will not have an impact on their relative ranking.

Voter category	Alternative	Excellent	Very Good	Good	Acceptable	Poor	To Reject
1000 voters	A	1000					
	B		1000				
1 voters	A				1		
	B			1			
1000 voters	A					1000	
	B						1000
Total	A	1000			1	1000	
	B		1000	1			1000

Here 2000 voters prefer A to B, one voter prefers B to A, and B is socially preferred to A.

4 Majority judgment and tactical voting

A voting system is said to be *strategy-proof* if honesty is the best policy for every voter, i.e. for each voter the best strategy is to express his true preference. In the contrary the system is *manipulable*. We next give a formal definition of manipulability. If Φ is a profile and ϕ_i is the preference order of voter i, the profile of all other voters is Φ_{-i}, so $\Phi = (\Phi_{-i}, \phi_i)$. A choice function (function which will choose one alternative) f is *manipulable* is there exist Φ and ϕ'_i such that

$$f(\phi_i', \Phi_{-i}) \succ_{\phi_i} f(\phi_i, \Phi_{-i})$$

which means that if voter i lies about his preference, he will be more satisfied by the chosen alternative. This brings us to the following central theorem.

Theorem 2 (Gibbard-Satterwaithe's Impossibility Theorem). *There is no choice function that is unanimous, non-dictatorial and strategy-proof for all preference profiles when there are at least three candidates.*

Here *unanimous* means that if every voter ranks A first, then society must choose A.

Even though this theorem tells us that no choice function is strategy-proof, we can try to decrease as much as possible the possibility of manipulation. We will now see that in the case of social grading functions there are functions with interesting strategy-related properties.

Definition 6 (Strategy-proof-in-grading). Suppose that r is a jury's final grade. A social grading function is *strategy-proof-in-grading* if when a judge's input grade is $r^+ > r$, any change in his input can only lead to a lower grade; and if, when a judge's input grade is $r^- > r$, any change in his input can only lead to a higher grade.

If a function is strategy-proof-in-grading, judges have no reason to lie about their true preferences.

Definition 7 (Order functions). An *order social grading function* assigns to each alternative the kth highest grade that it received from the jury.

If an alternative obtained the grades Good,Very Good, To reject, Acceptable, Poor, Very Good, the 3rd order function gives Good as a final grade.

Theorem 3. *The unique strategy-proof-in-grading social grading functions are the order functions.*

Proof. It is easy to see that order functions are strategy-proof-in-grading. Indeed if the final grade of alternative is r, anyone who gave a higher grade and who wishes to change the result can either:

- give a higher grade, in which case the kth highest grade remains r.
- give a lower grade, which will either not change the kth highest grade or will give a new grade $r* < r$.

The reasoning is exactly the same for a judge who gave a lower grade than the final grade. In any case a judge cannot alter the vote in his favor.

Let f be the aggregation function. Let $f(r_1, \ldots, r_n) = r$. Unanimity and monotonicity implies that r must fall between the best and the worst grades: $\max r_j \geqslant r \geqslant \min r_j$.

Suppose the judges gave the grade $r_1 \geqslant \cdots \geqslant r_n$. We will first prove that $f(r_1, \ldots, r_n) = r_k$ for some k.

Suppose $f(r_1, \ldots, r_n) = r$. If $r_j > r$ then we have:

$$\forall r_j^* \geqslant r, f(r_1, \ldots, r_{j-1}, r_j^*, r_{j+1}, \ldots, r_n) = r$$

Indeed:

- if $r_j^* > r_j$ then it is true because f is strategy-proof-in-grading.
- if $r_j^* < r_j$, then by monotonicity the value of f can either decrease or remain r. If it decreases, it would mean that by going back from r_j^* to r_j the function would increase, which is not possible since the f is strategy-proof-in-grading.

In the same way, if $r_j < r$ then we have:

$$\forall r_j^* \leqslant r, f(r_1, \dots, r_{j-1}, r_j^*, r_{j+1}, \dots, r_n) = r$$

If we suppose $r \neq r_j$ for all j. We must have $r_j > r > r_{j+1}$ for some j. By the previous statement, we can claim that for any grades r^+ and r^- such that $r^+ > r > r^-$,

$$f(\overbrace{r^+, \dots, r^+}^{j}, \overbrace{r, \dots, r}^{n-j}) = r \text{ and } f(\overbrace{r, \dots, r}^{j}, \overbrace{r^-, \dots, r^-}^{n-j}) = r$$

By monotonicity, the value of f on the left should be strictly greater than the value on the right. We have a contradiction. Thus $r = r_k$ for some k.

We know that when $s_1 \geqslant \dots \geqslant s_{k-1} \geqslant s_k = r_k \geqslant s_{k+1} \geqslant s_n$ we have $f(s_1, \dots, s_n) = r_k$. We now show that k is independant of $(r_1, \dots, r_n)$. Let $g(r_1, \dots, r_n) = k$ if $f(r_1, \dots, r_n) = r_k$ on the open set $R > r_1 > \dots > r_n > 0$ (where R is the highest grade in the language Λ). Since f is continuous, g is continuous on this set as well. Since g takes integer values, it is a constant on this set. Therefore f takes the value r_k for the same k on the entire set.

□

We observe this exact property in majority judgment since majority-grade is an order function. This shows us that this voting system is robust regarding manipulation.

Other voting systems can be very sensitive to strategy-voting, for example the two-round system which is used in France. An example of manipulability is the 2002 Presidential election. Even if many voters voted for a left-wing candidate, the two candidates who reached the second round were right-wing. This is due to the fact that the main candidate of the left Lionel Jospin saw many small candidates close to him ideologically which lead to a scattering of the left-wing votes. To obtain satisfaction, left-wing voters would better have concentrated their votes on Lionel Jospin in the first round.

5 Conclusion

We can therefore see that Majority Judgment has very interesting properties regarding both Arrow's theorem and strategy-voting. On the other hand different issues arose through he paper which explain the difficulty to implement this system in our voting practises. A question that comes up at this point is: *How much information can we ask the voter?* Indeed depending on the issue we can ask more or less effort to the voter. The French two-round system couldn't be more simple since it requires only one name. Ranking or grading ten to fifteen candidates can reveal itself to be a complicated matter for many voters and thus disturb the election process. Solving this question (which belongs to the field of psychology rather than social choice theory) could push Majority Judgment to be used in political decision-making.

6 Bibliography

References

[1] Andreu Mas-Colell, Michael D. Whinston et Jerry Green, *Microeconomic Theory*, 1995.

[2] Geanakoplos, J. ,*Three brief proofs of Arrow's impossibility theorem*, 2005.

[3] Michel Balinski and Rida Laraki, *Majority Judgment: Measuring, Ranking, and Electing*, 2010.

[4] Wikipedia, *Arrow's impossibility theorem, Majority Judgment*