Simulations in Models of Preference Aggregation

Mostapha Diss, Eric Kamwa

Simulations in Models of Preference Aggregation

Mostapha Diss* Eric Kamwa†

Abstract

Social choice theory provides a theoretical framework for analyzing how to combine individual opinions, preferences, interests or welfare so as to reach a collective decision. Social choice theory is one of the areas in economics that has seen a boom in simulations work using models based on the behavior of individuals involved in collective decision-making. The purpose of this paper is to offer to the uninitiated reader a methodological presentation of these different models, as well as the techniques for theoretical calculations and simulations, and then to report on recent developments concerning new models and advances in calculation techniques and simulations. This paper will thus give readers easy access to the models which, due to their complexity, might seem to be reserved for initiates. We take the opportunity to present and discuss the assumptions that support each of the models, and indicate how simulations may be helpful in analyzing complex problems in social choice theory.

Keywords: Social choice, voting, preference aggregation, simulations, models
JEL Codes: C15, C18, D71, D72

*CRESE EA3190, Univ. Bourgogne Franche-Comté, F-25000 Besançon, France. Email: mostapha.diss@univ-fcomte.fr.

†LC2S UMR CNRS 8053, Université des Antilles, F-97275 Schoelcher Cedex. Email: eric.kamwa@univ-antilles.fr.

Simulations dans les Modèles d'Agrégation des Préférences

Résumé

La théorie du choix social offre un cadre théorique pour analyser comment combiner des opinions, des préférences, des intérêts ou bien-être individuels afin de prendre une décision collective. La théorie du choix social est l'un des domaines de l'économie qui a connu un essor des travaux autour des simulations à partir de modèles basés sur le comportement des individus impliqués dans la prise de décision collective. L'objectif de cet article est d'offrir au lecteur non initié une présentation méthodologique de ces différents modèles, ainsi que des techniques de calculs théoriques et de simulations, puis de rendre compte des développements récents concernant les nouveaux modèles et des avancées en matière de techniques de calcul et de simulations. Cet article donnera ainsi aux lecteurs un accès facile aux modèles qui, en raison de leur complexité, peuvent sembler réservés aux initiés. Nous en profitons pour présenter et discuter des hypothèses qui sous-tendent chacun des modèles et indiquer comment les simulations peuvent être utiles pour analyser des problèmes complexes de la théorie du choix social.

Mots-clés: Choix social, vote, agrégation des préférences, simulations, modèles

JEL Codes: C15, C18, C63, D71, D72

1 Introduction

Over the past few decades, *social choice theory* has been one of the areas in economics that has seen a boom in work using models based on the behavior of individuals involved in collective decision-making. These models have helped in designing renowned and robust results in the field of preference aggregation. The frameworks developed on the basis of these models made it possible, by theoretical and/or computer simulations, to validate or invalidate several analytical results established in the literature. The aim of this paper is to offer, to the uninitiated, a methodological presentation of these different models, as well as the associated techniques of theoretical calculations and simulations, and then to report on recent developments concerning new models and advances in calculation techniques and simulations.

Computer simulations emerged in social choice theory almost at the same time as in political science. From the late sixties onwards, more and more researchers turned to computer simulations to study the behavior of agents involved in collective decision-making processes. This method of analysis differs from those that have traditionally prevailed. This new approach mainly relies on the use of statistical tools to model behaviours, as well as the use of empirical analyses on real databases (results of elections, referenda or public consultations, etc.) or data collected through opinion polls or surveys.

We could actually argue that a large literature using computer simulations in social choice theory is devoted to the evaluation of voting systems according to different normative criteria, in order to obtain a hierarchy of the voting rules under investigation. Basically, this literature can be divided into two main categories:

1. Research that aims to compare various voting rules on the basis of their ability to lead to some desirable voting outcomes. The selection of the Condorcet winner, when he/she exists, is one of these desirable voting outcomes. Indeed, there is a large literature in social choice theory devoted exclusively to the Condorcet efficiency[1] of voting rules. The question of the concordance between voting rules[2] has also been widely studied in this literature. Naturally, there are a variety of other considerations that go into this category.

2. Much of the research in social choice theory is concerned with whether a paradox can occur for a given voting rule or not. We define a voting paradox as an undesirable outcome to which a ballot can lead under a given voting rule and which may be regarded as surprising or counterintuitive. Many voting paradoxes has been the focus of numerous investigations in the literature: Condorcet's paradox, Borda's paradoxes, referendum paradox, monotonicity paradoxes, susceptibility to strategic manipulation, etc. This list of examples is of course a partial one, but they all clearly demonstrate that the notion of the probability of voting paradoxes is a central one.

[1]The Condorcet efficiency of a voting rule is defined as the conditional probability that the procedure select the Condorcet winner, given that a Condorcet winner exists.

[2]The analysis is of interest since it allows to know how the choice of the voting rule is susceptible to impact the determination of the winner.

All these questions and many others have been investigated via simulations. For a detailed survey of early research on the two items, the reader can refer to Felsenthal (2012), Felsenthal and Nurmi (2018), Gehrlein and Lepelley (2011, 2017), Nurmi (1999, 1989), and Saari (1995), among others.

Throughout this paper, the only illustration considered is the likelihood of Condorcet's paradox which is one of the themes that has most mobilized researchers from both the traditional line and the simulations approach. This choice can be understood, given the historical and theoretical importance of this paradox in the social choice literature which has largely been dominated by studies that are associated with this paradox. According to this paradox, it is possible, by aggregating the preferences of a group of individuals who are asked to rank three propositions (let us say A, B, and C) in order of preference, that a majority of voters prefers A to B, another majority prefers B to C and another prefers C to A. In such a case, we get a majority cycle which is the main drawback of the voting rule suggested by Condorcet (1785) as an alternative to that proposed by Borda (1781). Indeed, at the end of the 18th century, Borda and Condorcet, both members of the Paris Royal Academy of Sciences, proposed alternative voting rules to the one that was in use in the academy. The Borda rule picks as the winner the candidate with the highest Borda's score.[3] Condorcet criticized this rule in that it allows the existence of a candidate that is preferred by more than half of the electorate to the Borda winner; he proposed a rule based on pairwise comparisons. According to this rule, a candidate should be declared the winner if he/she beats all the other candidates in pairwise majority; such a candidate is called the Condorcet winner. In the end, the members of the academy leaned in favor of the Borda rule to the detriment of the Condorcet rule.

Although the Borda-Condorcet debate concerned only two rules for collective decision-making, it helped lay the groundwork for what can be described as "the quest for the best rule for collective decision": a quest that is built around the comparison of the merits of different voting rules against each other. Decision rules can therefore be compared either on the basis of normative properties which they meet or not (the axiomatic approach), or on the basis of the frequency with which they satisfy or fail a given criterion (the probabilistic approach). These two ways of proceeding define the two principal approaches in the theoretical study of social choice theory. We present these two approaches in Section 2 where we show that they complement each other.

As mentioned before, the second approach experienced a boom in the late fifties, which saw much work on the occurrence of Condorcet's paradox. One of the paths taken to evaluate the Condorcet paradox is empirical studies based on data collected during real decision-making processes, elections, surveys or polls. For an overview of the work that falls within this framework, the reader may refer to Chamberlin et al. (1984), Dobra (1983), Dobra and Tullock (1981), Kurrild-Klitgaard (2001, 2008), Niemi (1970), Regenwetter et al. (2002a,b), Riker (1958, 1965), Taylor (1997) and Tideman (1992). Notice that the results of these various empirical studies are summarized in Gehrlein (2006) and Gehrlein and Lepelley (2011, 2017). However, empirical analyses are not always possible, because the data for

[3]With m candidates, the Borda rule awards $m - j$ points to a candidate each time he/she is ranked j-th in a voter's ranking. The total number of points received by a candidate defines his/her Borda's score.

such studies are rarely available, accessible or even reliable; this may limit the scope of the empirical approach. The solution to such a limit is the use of probabilistic models describing the behavior of individuals and then try to find the theoretical probabilities of voting events according to these models.

The first use of probabilistic models in social choice theory dates back to the paper by May (1948) who calculated the overall likelihood of the referendum paradox which is the one that occurred, for instance, when Donald Trump was elected in 2016.[4] It deserves to be mentioned that, some years later, Guilbaud (1952) also published an important paper on the probability of the existence of a Condorcet winner when three options are in the contest. This practice spread in the late 1960s with the work of Campbell and Tullock (1965), Garman and Kamien (1968) and Niemi and Weisberg (1968) on the probability of a cyclical majority. For the state of the art on the Condorcet paradox using probabilistic models, the reader can refer to the books of Gehrlein (2006) and Gehrlein and Lepelley (2011, 2017). The use of probabilistic models requires us to make a priori assumptions on the distribution of preferences in order to build a model describing the behavior of the individuals. We discuss the main assumptions and models in Section 3. The probabilistic model approach faces some criticisms, chiefly: *i)* even with a fairly small number of candidates in the running, probabilistic models can very quickly become intractable; *ii)* the realism of the assumptions underlying most models is questionable; and *iii)* the results obtained depend strongly on the hypothesis underlying the model and can thus vary from one hypothesis to another. We give a detailed discussion of each of these limitations later in the paper.

As mentioned before, the use of simulations has emerged as a means of transcending the main limitations of the two traditional approaches of preference aggregation analysis. In a general sense, simulation is the systematization and formulation of a model for determining the main characteristics of a system, a transaction or a process. In the framework of the aggregation of preferences, simulations mean the construction of a model which tries to simulate to the best degree possible the behaviours (i.e., preferences) of the individuals (i.e., voters). The use of simulations in social choice theory is not as recent as one might think; it was a result of the pioneering work of Arrow (1963, 1951) that the first results based on simulations appeared (see for instance, Klahr, 1966, Weisberg and Niemi, 1978, 1973), mainly around the probability of the Condorcet paradox, in which analytical calculations were no longer possible due to mathematical limits when the number of candidates or voters increases. We come back to these different studies later in the paper, where we take the opportunity to provide a brief history of simulations in social choice theory and to present the different approaches adopted as well as review their scope.

The rest of the paper is organized as follows: First of all, we must familiarize the reader with the object of social choice theory, namely the aggregation of preferences. Section 2 is therefore dedicated to this end. In Section 3, we present the main models or hypotheses on which the theoretical works are based. Section 4 is devoted to the methods of simulations that have been developped in contrast to that of the theoretical approach. Section 5 concludes.

[4]More precisely, Donald Trump won enough states to secure the majority in the Electoral College, while Hillary Clinton received 2.87 million more votes than Trump.

2 Aggregation of individual preferences

2.1 Preference aggregation: a brief history

The aggregation of preferences is at the heart of social choice theory, the essential purpose of which is to study ways of coherently aggregating individual preferences into a collective choice or a collective ranking of candidates. Given a group of individuals, who have to choose between at least two options (alternatives or candidates), a collective decision procedure, also called an aggregation rule, associates with each state of nature a collective ranking of options or a subset of winners, ideally a singleton. This theory leans on fundamental microeconomic principles and aims to understand the decision-making of rational individuals as regards to economic phenomena or beyond. According to List (2013), social choice theory is not a single theory but a cluster of models and results concerning the aggregation of individual inputs. Indeed, it covers, by its vast field of applications, a multitude of contexts where the problem is formally similar: a group of individuals (e.g., experts, judges, jury, voters, etc.) who face a set of options (e.g., resource allocations, economic projects, candidates in a competition or an election, etc.), must reach a collective decision based on the opinions and interests of the different members. In other words, the scope and stakes of this theory are, in fact, potentially far-reaching, and may interest, in addition to economics and politics, areas as diverse as management, psychology, computer science and philosophy. This theory is now a recognized branch of modern microeconomics.

Historically, since the seminal works of Arrow (1963, 1951), the interest of economists in the question of collective choice has had its source in the *new economy of well-being* developed in the 1940s, thanks in particular to the works of Pigou (1920), Bergson (1938), and Samuelson (1947). The traditional economy of well-being was for a long time dominated by an almost total adherence to the utilitarian approach of Bentham (1789), and settled down, with the work of Edgeworth (1881), Marshall (1960), and Pigou (1920), in a very different framework from that of social choice theory. It was only from the late 1940s, with the work of Arrow (1963, 1951), Black (1976, 1958, 1948), and May (1952, 1971), that the context of the collective decision has been conjoined with that of the welfare economy, leading to the birth of the social choice theory in its modern form.

More precisely, in utilitarian calculations, the preferences of individuals are represented by numerical utility functions defined on all social states, and judgments on social interest are obtained by maximizing the sum of individual utilities. This implies that one can measure satisfaction, or happiness, in the form of utilities and that these utilities are comparable between individuals. This cardinal approach was called into question in the 1930s and finally abandoned in favor of the new welfare economy, which banned any possibility of interpersonal comparison of individual utilities and which gave an important place to the criterion of *Pareto efficiency*. This criterion tells us that one allocation is preferable to another if it allows an increase in the level of satisfaction of one or more individuals without reducing the utility of others. This principle proved to be insufficient, however, since many possible allocations can be Pareto optima. Hence, an economic theory of collective choice has become indispensable. The notion of a collective aggregation function introduced by Arrow

(1963, 1951) fits into this framework insofar as it allows aggregating individual preferences into a collective preference, so that society can choose between the different optima in the sense of Pareto. Arrow (1963, 1951) showed that inconsistencies related to collective choice are not a surprise as they affect a very wide class of aggregation procedures. The method adopted in Arrow's theorem consists of choosing a certain number of properties which it seems reasonable to impose on the aggregation procedures and then demonstrating that the satisfaction of these properties taken together leads to an inconsistency, in the sense that such an aggregation procedure does not exist. In other words, there is no aggregation procedure that verifies all the desirable conditions taken into account.[5]

2.2 The axiomatic approach and the probabilistic approach

By accepting the idea of discarding or weakening some properties and adding others, Arrow's theorem has given rise to countless contributions using what we call the *axiomatic approach*, which is the more common approach in social choice theory. Moreover, the majority of the convincing results in this field have been obtained in the form of impossibility theorems, which highlights the difficulty of designing a method allowing a reasonable aggregation of the opinions expressed by individuals in a decision procedure. However, the most important limitation of the axiomatic approach is that it does not offer information on the frequency of situations where aggregation procedures violate the desirable properties taken into account.

The *probabilistic approach* has been developed in the framework of social choice theory in order to deal with this limit of the axiomatic approach. This approach starts by using models describing the behavior of individuals involved in the aggregation process and then to quantify the probability of occurrence of certain types of collective outcomes for a given aggregation rule under the assumption fixed on the distribution of individual preferences. Most of the models that are widely used in the literature are based on two pioneer assumptions: the *Impartial Culture* (IC) and the *Impartial and Anonymous Culture* (IAC) assumptions. These models will formally be defined and discussed later. Most of the models used are special cases of the multinomial law, and one of their limitations lies in the fact that even for a limited number of alternatives and individuals, the multinomial law becomes difficult to manage— except, in this case, by resort to numerical or computer simulations. Indeed, simulations have made it possible to validate or invalidate several results established in the literature.

[5]The desirable conditions are among others: i) *completeness and transitivity* according to which the collective ranking is a complete and transitive binary relation; ii)the condition of *independence of irrelevant alternatives* requires that the relative rankings of two alternatives depends only on the relative rankings of these candidates in the voters' rankings; the *Pareto principle* according to which when all the voters have the same strict preference over a pair of candidates, the social ranking is the same as the voters' unanimous preference; and iv) the *non-dictatorship* imposes that a given individual cannot impose his choice to the whole society.

2.3 Individual preferences

Let N be the set of $n \geq 2$ individuals who have to decide on the set A of $m \geq 3$ alternatives.[6] In order to decide, each voter must make a judgment of his/her own on the candidates in the running. This judgment is part of the process of preference formation which in social choice theory is not the subject of a study. The preferences are then assumed to be exogenous. In addition, it is generally assumed that each voter votes sincerely and acts according to his/her true preferences.

The binary relation R over A is a subset of the cartesian product $A \times A$. For $a, b \in A$, if $(a, b) \in R$, we write aRb to say that "a is at least good as b". $\neg aRb$ is the negation of aRb. If we have aRb and $\neg bRa$, we will say that "a is better or strictly preferred to b". In this case, we write aPb with P the asymmetric component of R. If we assume that voters only have strict preferences also called linear orders,[7] the *preference profile* $\pi = (P_1, P_2, ..., P_i, ..., P_n)$ gives all the linear orders of the n voters on A. The set of all the preference profiles of size n on A is denoted by $\mathcal{P}(A)^n$. A voting situation $\tilde{n} = (n_1, n_2, ..., n_t, ..., n_{m!})$ indicates the number of voters for each linear order such that $\sum_{t=1}^{m!} n_t = n$.

For $m = 3$ and linear orders assumed, Table 1 describes the possible strict rankings on $A = \{a, b, c\}$. In this table, it is indicated that n_1 voters have the ranking abc; this means that they rank candidate a at the top followed by candidate b and candidate c is the least preferred.

Table 1: Possible strict rankings on $A = \{a, b, c\}$

$n_1 : abc$	$n_2 : acb$	$n_3 : bac$
$n_4 : bca$	$n_5 : cab$	$n_6 : cba$

It should be noted that, in most social choice literature, it is explicitly admitted that the preferences of individuals are linear orders. This implies that agents cannot be indifferent between two or more alternatives. Using Monte Carlo simulations, Bjurulf (1972) and Jones et al. (1995) pointed out that this is not without impact on the results obtained; they also

[6] As mentioned before, in order to show the increasing interest for simulations in social choice theory, our illustrative example is the Condorcet's paradox which requires that there are at least three alternatives. However, notice that an entire component of the literature on probability calculations and simulations in social choice theory is ignored in this paper: the one that considers the two-alternative case. First, some voting paradoxes can occur with only two alternatives (e.g., the referendum paradox, see Miller, 2012); second (and most importantly) a large number of studies deal with the question of voting power, which can be measured as the probability of being pivotal for a voter, in a two-candidate (voting "yes" or "no") framework. The two most famous power indices, the Banzhaf index and the Shapley-Shubik index, are respectively based on IC and IAC. On this topic, see e.g., Straffin (1988) and Le Breton et al. (2016).

[7] A linear order is a binary relation that is transitive, complete and antisymmetric. The binary relation R on A is *transitive* if for $a, b, c \in A$, if aRb and bRc then aRc. R is *antisymmetric* if we have aRb and bRa, then $a = b$. R is *complete* if and only if for all $a, b \in A$, we have aRb or bRa.

emphasize that admitting the possibility of having weak orders rather than strict orders would make preferences more realistic and would greatly reduce the probability of certain events such as the Condorcet paradox. Analytically, Fishburn and Gehrlein (1980) reached the same conclusion. In a recent book, as part of their work on “Behavioral Social Choice”, Regenwetter et al. (2002b) have developed the tools to overcome the tradition of a priori preferences. In addition to highlighting the different limits of the traditional approach of social choice theory, they have developed methodologies to (re)construct preference distributions from incomplete data and a statistical sampling and Bayesian inference framework for the theoretical and empirical analysis of preference aggregation in samples drawn from practically any distribution over any family of binary relations. We say more on this in Section 4.

We are now equipped to introduce the main theoretical models used in social choice theory when dealing with voting events.

3 Simulations based on theoretical models on agents' behavior

The main purpose of using theoretical assumptions to model the behavior of a group of individuals is to derive a representation of the probability of a given event. The starting point in the different models is to assume an a priori distribution or assumption under which the samples of the individual preferences are drawn. In this paper, we will only focus our attention on the most popular and widespread models.

3.1 The main theoretical models

The impartial culture model

The impartial culture (IC) model was introduced for the first time in the social choice literature by Guilbaud (1952), who was interested in calculating the probability of the Condorcet paradox. As mentioned before, this model is among the most used in the literature, as shown by the large body of work on calculating the probabilities of electoral events produced since the seventies.

Under this model, it is assumed that all voting profiles have the same probability of appearing. This means that each individual randomly and independently chooses his/her preference on the basis of a uniform probability distribution across all linear (or weak) orders. It follows that, with linear orders, where each of $m!$ linear orders has the same probability $\frac{1}{m!}$ of being chosen by an individual, and the probability of attaining a voting situation $\tilde{n}$ is given by:

$$\text{Prob}(\tilde{n}) = \frac{n!}{\prod_{j=1}^{m!} n_j!}(m!)^{-n} \tag{1}$$

It is worth noting that the IC model is only a special case of the multinomial law. David and Mallows (1961), Gehrlein and Fishburn (1978a,b) and Plackett (1954) showed that for an infinite number of individuals, an application of the central limit theorem allows an approximation of the multinomial law by a multivariate normal law. Recall that a multivariate normal distribution is a vector of multiple normally distributed variables, such that any linear combination of the variables is also normally distributed.[8] The central limit theorem states that averages calculated from independent, identically distributed random variables have approximately normal distributions, regardless of the type of distribution from which the variables are sampled, provided it has finite variance. However, the use of the multinomial law, as well as its approximation, can quickly become intractable even with a relatively small number of individuals. The various probability calculation techniques suggested under the IC model (e.g., Gehrlein and Fishburn, 1978a,b, Saari and Tataru, 1999) have the disadvantage of leading to different formulas which are often not compact and are difficult to handle.

Fishburn and Gehrlein (1980) was the first to introduce an extension of IC in order to take into account the possible indifference in the agents' preferences: the *impartial weak ordering culture* (IWOC). More recently, Diss et al. (2010) have provided another extension of IC that allows the possibility for voters to have dichotomous preferences with complete indifference between two or more candidates: the *extended impartial culture* (EIC). Note that for selected values of the parameters in the EIC and IWOC models, the IC model is easily found. Also, we can easily find the IWOC model from EIC. As it can be seen, these extensions are only refinements of IC that tend to take into account the remarks of Bjurulf (1972) and Jones et al. (1995) according to which admitting the possibility of having weak orders rather than strict orders would make preferences more realistic and would greatly reduce the probability of certain events such as the Condorcet paradox.

The dual culture model

The dual culture (DC) model first introduced by Gehrlein (1978) operates as the IC model; it was defined only for strict rankings. Under DC, the probability that a given individual chooses his/her preference is the same as that of a voter with the inverted (dual) ranking. Let us illustrate this with the preferences of Table 1. In this table, the rankings *abc* and *cba*, *acb* and *bca*, *bac* and *cab* are dual; so, the distribution is as follows:

$$\begin{aligned} Prob(abc) &= Prob(cba) = p_1 \\ Prob(acb) &= Prob(bca) = p_2 \\ Prob(bac) &= Prob(cab) = \frac{1}{2} - p_1 - p_2 \end{aligned}$$

One may notice in this case that we recover the IC model when $p_1 = p_2 = \frac{1}{6}$.

[8] Recall that the normal distribution is the most common type of distribution assumed in probabilistic analyses. The standard normal distribution has two parameters (the mean and the standard deviation) and has the main following properties: i) The mean, mode and median are all equal. ii) The curve is symmetric at the center (i.e., around the mean). iii) Exactly half of the values are to the left of center and exactly half the values are to the right. iv) The total area under the curve is 1.

The impartial and anonymous culture model

The impartial and anonymous culture (IAC) model was introduced for the first time in the literature of social choice theory by Gehrlein and Fishburn (1976) and Kuga and Nagatani (1974). Under this model it is assumed that all voting situations are equally likely to be observed, then the probability of a given event is calculated according to the ratio between the number of voting situations in which the event occurs and the total number of possible voting situations. The possibility of computing the probability of an event as a ratio is not specific to IAC: with IC, the probability is the ratio between the number of preference profiles in which the event occurs and the total number of possible preference profiles. Both models are based on a notion of equiprobability, but the elementary events are preference profiles under IC and voting situations under IAC. Notice that the IAC model allows to obtain closed form representations and this is one of its main advantages compared to the IC model. The probability of getting a given voting situation $\tilde{n}$ with n voters and m candidates is given as follows:

$$\text{Prob}(\tilde{n}) = \frac{n!(m!-1)!}{(n+m!-1)!} \tag{2}$$

Under the IAC model, the number of voting situations associated with a given event can also be reduced to the solutions of a finite system of linear constraints with rational coefficients. For instances, using the labels of Table 1, the number of voting situations associated with the Condorcet paradox (of the type a is majority preferred to b, b is majority preferred to c, and c is majority preferred to a) is reduced to the solutions of the following system:[9]

$$\begin{cases} n_1+n_2+n_5 > n_3+n_4+n_6 & (a \text{ is majority preferred to } b) \\ n_1+n_3+n_4 > n_2+n_5+n_6 & (b \text{ is majority preferred to } c) \\ n_4+n_5+n_6 > n_1+n_2+n_3 & (c \text{ is majority preferred to } a) \\ n_i \geq 0 \quad \text{for} \quad i=1,2,\ldots,6 \\ \sum_{i=1}^{6} n_i = n \end{cases} \tag{3}$$

Different techniques and algorithms for finding solutions for such systems have been proposed in the literature; the reader may refer to the works of Cervone et al. (2005), El Ouafdi et al. (2020), Gehrlein and Lepelley (2011), Lepelley et al. (2008) and Wilson and Pritchard (2007).

The maximal culture model

The maximal culture (MC) model is due to Fishburn and Gehrlein (1977). MC is quite similar to IAC with the exception that there is no need to fix the number of voters in the random voting situation. It fixes an integer L ($L > 0$) and each ranking is drawn from a uniformly random distribution over $[0, L]$. According to this, for three-candidate elections, the number of possible equally likely voting situations is equal to $(L+1)^6$ and the expected number of voters in a voting situation is $3L$.

[9] The second possible cycle is defined in the same way using the symmetry of the IAC assumption with respect to candidates.

The urn model: The Pólya-Eggenberger model

According to Berg (1985a,b) and Berg and Bjurulf (1983), the IC and IAC models are in fact only two special cases of the more general Pólya-Eggenberger model; this is not the case for the MC model. This model was introduced into the social choice literature by Berg (1985a).[10] In this model, everything happens as if from an urn containing B balls including B_j balls of color j ($j = 1, 2, \ldots, m!$), where each individual involved in the collective decision process chooses his/her preference by means of a random draw of a ball in the urn, and at each draw, α balls of the same color as the one drawn by the individual are added into the urn. The quantities B and α are assumed to be positive real numbers. So, the probability of getting a given voting situation $\tilde{n}$ under the Pólya-Eggenberger model is:

$$\text{Prob}(\tilde{n}) = \frac{n!}{B^{[n,\alpha]}} \prod_{j=1}^{m!} \frac{B_j^{[n,\alpha]}}{n_j!} \tag{4}$$

where $B^{[n,\alpha]} = \prod_{i=0}^{n-1}(B + i\alpha)$ is a generalized ascending factorial and the B_j are positive numbers associated with each order such that $B = \sum_{j=1}^{m!} B_j$.

According to Berg and Bjurulf (1983), α is a parameter measuring the level of social cohesion: the larger it is, the more the preferences of the individuals tend to be homogeneous. Berg and Bjurulf (1983) showed that if we fix $B_j = 1$ for all $j = 1, 2, \ldots, m!$, the Pólya-Eggenberger model leads to the IC model for $\alpha = 0$ and to the IAC model for $\alpha = 1$. Compared to the IC and IAC models, the Pólya-Eggenberger model therefore has the advantage of taking into account all possible degrees of interdependence in the preferences that individuals adopt.

Table 2 reports the limit probability, i.e., when the number of voters tends to infinity, of the Condorcet paradox in three-candidate elections obtained under each of the above theoretical models. All these results are drawn from Gehrlein and Lepelley (2011, 2017).

Table 2: Limiting probability of the Condorcet paradox in three-candidate elections

Model	Limiting probability of the Condorcet paradox
IC	$\frac{1}{4} - \frac{3}{2\pi}\arcsin(\frac{1}{3}) \approx 0.0877$
DC	$\frac{1}{4} - \frac{1}{2\pi}\big(\arcsin(1-4p_1) + \arcsin(1-4p_2) + \arcsin(4p_1+4p_2-1)\big)$
IAC	$\frac{1}{16}$
MC	$\frac{11}{120} + \frac{99L^4+341L^3+474L^2+305L+109}{120(L+1)^5}$

The different models that we have just presented have governed most of the theoretical analyses and have been used to develop the probability representations of electoral events in

[10] The reader may refer to Johnson and Kotz (1977) for an overview of this model.

particular for three-candidate elections. Despite the fact that for the same event, the models can lead to different probabilities, Gehrlein and Lepelley (2004) put forward a certain number of arguments to justify the use of such models. Let us summarize these arguments:

- It can be useful to find out if the relative probabilities of paradoxical outcomes on various voting mechanisms behave in a consistent fashion over a number of different assumptions about the likelihood that voting situations or voter preference profiles are observed.
- With real elections, large amounts of empirical data are not available; the use of theoretical models is thus found to be very useful.
- Despite the fact that they are generally believed to represent situations that exaggerate the probability that paradoxical events will occur, the theoretical models can show that some paradoxical events are very unlikely to be observed in reality.
- Theoretical models can show the relative impact that paradoxical events can have on different types of voting situations.
- By using probability models to obtain closed form representations, it is easy to observe the impact of varying different parameters (e.g., parameters of specific measures of social homogeneity or group coherence) of voting situations or voter preference profiles; this is somewhat more difficult to do with simulation studies.
- The obtained probability representations are directly reproducible and verifiable with mathematical analysis, which is not as simple to do with simulation analysis.

The last two arguments express the main advantages of theoretical models on approaches based on simulations. However, we must also admit that very early on (and continue to do so even today, despite the increase in computer processing power) the analytical approach has shown its limits when trying to explore situations with more than three candidates. Thus, alongside the analytical approach, many works have been developed on the basis of simulations based on the a priori theoretical hypotheses that we have just presented.

3.2 Theoretical-based simulations

Initially, the studies of voting situations involved only two or three candidates and were limited to a finite or a very small number of agents. This is due to the fact that the analytical calculations, which were done by hand, rapidly became complicated and indeed unmanageable or untractable. One way to overcome the constraints and limits of the analytical approach would be to operate on real data; however, these are difficult to access, or rarely available. Even if they are available, the reliability of expressed preferences may be questioned, and there is no guarantee that the voters interviewed will all be able to really express their preferences when the number of candidates is large. To circumvent this obstacle, several authors quickly opted for the assistance of computer science through simulations. Over time, simulations have come to be no longer confined to the subfields of economics, and are spreading to almost all social sciences (see for instance, Axelrod, 1997, Fontana, 2006). The principle of simulation, in the common sense of the term, is to use a model, that is to say an abstract representation of a system or a problem, and to study the evolution of this model without operating the actual system.

In their first usage in social choice theory, the applications of the simulations focused for the most part on the evaluation of the probability of the Condorcet paradox. Among these applications, without being exhaustive, are the works of Campbell and Tullock (1965), DeMeyer and Plott (1970), Gehrlein and Fishburn (1976), Klahr (1966) and Weisberg and Niemi (1978, 1973); most of these works involve only strict orders for agent preferences. According to Jones et al. (1995) this could be justified by the performance of computers at that time. Taking advantage of the computer advances of the 90s, Jones et al. (1995) conducted an analysis of the simulated probability of the Condorcet paradox when weak preferences are allowed.

The simulations are made assuming a certain distribution a priori on the preferences of individuals, i.e., recourse to one of the theoretical models presented earlier. Once the distribution is chosen, the preferences are generated using the Monte Carlo simulation method, which is a method of estimating a numerical quantity that uses random numbers. Note that this method was introduced by Von Neumann and Ulam (1945), referring to games of chance in casinos, during the Manhattan project.[11] This method has the advantage of being easy to use. It is now applied to a very wide range of problems. Let us briefly present the methodology of Monte Carlo simulations under IC and IAC models as they are carried out as part of the aggregation of preferences for generating samples of preferences. For our presentation, we will focus on cases where only strict preferences are allowed and we will use the notation of Section 2.3.

Simulations under IC:

The goal is to generate, equiprobably, a profile of total orders with n voters and m candidates. So each of the $m!$ possible total orders is chosen equiprobably; to choose a total order is therefore equivalent to choose an integer between 1 and $m!$. An integer is chosen over this interval for each of the voters, one after the other and independently. The chance of occurrence of each profile in this process is actually $\left(\frac{1}{m!}\right)^n$. At the end of the process, which is anonymous, we count the number of voters assigned to each strict order; we then obtain a $m!$-uple of integers whose sum is equal to n. Concretely, the routine used is the following:

- We start from the profile $(0, 0, \ldots, 0)$ which is a null vector with $m!$ components.
- From step 1 (voter 1) to step n (voter n), an integer of 1 to $m!$ is equiprobably selected. If the result is j, add 1 to the component j of the profile.
- At the n^{th} stage, the profile is indeed a $m!$-uple of integers whose sum is equal to n.

[11] Project Manhattan is the code name for the research project that produced the first atomic bomb during the Second World War.

To generate a sample of size T (number of repetitions), we run the previous routine T times while keeping the result; this gives a T-tuple of profiles with total orders.

Simulations under IAC:

Given the $m!$ possible strict orders, the objective is to generate the voting situations (anonymous profiles) equiprobably. As a reminder, a voting situation is an $m!$-uple $(n_1, n_2, \ldots, n_{m!})$ of natural numbers whose sum is equal to n for which it will be necessary to randomly generate each of the components. To do this, $m! - 1$ numbers are generated equiprobably in $[0; 1]$ which we rank in increasing order, say $x_1, x_2, \ldots, x_{m!-1}$; this series will be completed by 0 the smallest possible value and 1 the largest possible value such that $0 = x_0, x_1, x_2, \ldots, x_{m!-1}, x_{m!} = 1$. Then, the value $n(x_j - x_{j-1})$ is assigned to n_j $(j = 1, 2, \ldots, m!)$ that is to say $n(x_1 - x_0)$ is assigned to n_1, $n(x_2 - x_1)$ to n_2 and so on until $n(x_{m!} - x_{m!-1})$ which is assigned to $n_{m!}$. Note that the values obtained may not be integers; they are then rounded to the lower unit. After rounding, if the sum of the numbers thus assigned is less than n, the difference observed will be added randomly and equiprobably to one of the components. By this process, voting situations have the same chances of being observed. To generate a sample of size T, we run the previous routine T times.

It is worth nothing that Feix and Rouet (2005) showed that there is a connection between the probability models (IC and IAC) and probabilistic models or distributions that are widely used in physics, particularly in quantum mechanics and statistical physics: the IC model is linked to Maxwell-Boltzmann distribution[12] and the IAC model is related to the Bose-Einstein statistic;[13] quantum statistics would thus be another gateway for calculating the probabilities of voting events. Feix and Rouet (2005) complete their analysis by simulating the probabilities of existence of the Condorcet winner under IC and IAC with a number of candidates ranging from 3 to 8 and electorates of infinite size. Their calculations show a certain convergence between IC and IAC when the number of candidates increases. Table 3 reports likelihood of the Condorcet paradox when preferences are simulated according to the IC and IAC models for voting situations with three to eight candidates.[14] The values in this table are derived from those of Table 4 and 5 by Feix and Rouet (2005).

[12] The Maxwell-Boltzmann statistic is a probability law or distribution used in statistical physics (thermal equilibrium) to determine the distribution of particles between different energy states.

[13] It describes one of two possible ways in which a collection of non-interacting, indistinguishable particles may occupy a set of available discrete energy states at thermodynamic equilibrium.

[14] These figures are consistent with those obtained under the IC model by Gehrlein (1985), Klahr (1966), Niemi and Weisberg (1968) and Weisberg and Niemi (1978).

Table 3: Limiting probabilities of the Condorcet paradox obtained by simulations under IC and IAC

	Number of candidates					
Models	3	4	5	6	7	8
IC	0.0877	0.1756	0.2513	0.3152	0.3694	0.4140
IAC	0.0624	0.1616	0.2477	0.3143	0.3691	0.4170

Other simulation results on the likelihood of the Condorcet paradox with a given number of candidates and a given number of voters, are available in the literature not only the IC and IAC models but also for many other assumptions; for an overview, the reader may refer to the papers of Fishburn and Gehrlein (1982), Gehrlein (1997), Jones et al. (1995), Pomeranz and Weil (1970) and Weisberg and Niemi (1978). It comes from all these results that the probability of the Condorcet paradox tends to increase with the number of candidates and the number of voters. The literature is now full of numerous probabilities of various electoral events, obtained by simulations of the IC and IAC models. See for instance, the works of Aleskerov et al. (2012), Brandt et al. (2016), Diss and Doghmi (2016), Kelly (1993), Lepelley et al. (2000) and others. Notice that based on a number of probabilistic models,[15] Laslier (2010) simulated the frequency of the existence of a Condorcet winner, for several profile sizes and also the likelihood of the election of the Condorcet winner, when he/she exists, for several voting rules. It comes from the simulation results that the way we should judge voting rules depends also on the context (political election, aggregation of judgments, jury, etc.) and the right model could depend on the type of collective decision problem under consideration.

It should be noted that in the days of the first simulation work in the theory of social choice, computer workstations were almost non-existent or at least expensive; access to mainframes was even more so. Thus, the simulations, which for the most part were confined to the probability of the Condorcet paradox, were limited to voting situations with three candidates and a very small number of voters. After a certain number of voters, the calculations were time consuming and the results were based on samples generated from a low number of repeats; this therefore casts doubt on the accuracy of the results. With the development of mathematical, statistical and computer techniques, over time, many (more or less complex) programming languages have been developed, as well as software and toolkits that meet the particular needs of the simulation, particularly for generating samples of preferences. Today, easily accessible Microsoft Excel spreadsheets offer many possibilities for simulations using simple macros and VBA language. We can also turn to more advanced tools such

[15] Among others, Rousseauist cultures, impartial culture, distributives cultures and spatial Euclidian cultures. Rousseauist cultures are adapted from Rousseau's ideal of a general will. Distributive cultures describe societies of complete antagonism with a context comparable to that which governs the problems where a unit of a divisible good has to be shared between individuals. Spatial Euclidian cultures are consistent with what we present in Section 4.1.

as *Maple*, *MATLAB* or *Mathematica* based on sharp programming languages more or less comprehensible only for insiders. This development of methods and techniques today makes available to researchers "turnkey" kits to effectively conduct their simulations that today can be done on personal computers or on dedicated servers, or even on supercomputers (Macal and North, 2010). The saving of time is remarkable and the accuracy of the obtained results is indisputable. The advances made in current computer simulation techniques have made it possible to correct or refine several theoretical results obtained in the past.

Like the analytical approach based on theoretical models, the theoretical simulations approach is strongly criticized. According to Tideman and Plassmann (2013), under the theoretical-based models, the analysis and consequently the results assume frequency distributions chosen just because of the convenient mathematical properties, while these distributions are far from reflecting what is happening in real elections. In fact, there is no evidence that voters' choices obey any probabilistic distribution, let alone a uniform distribution. No work has ever supported or even established that the above theoretical models reflect the reality in a particular situation. On the basis of their criticisms of theoretical models, several authors have argued for simulations based on more realistic distributions and assumptions.

4 Other approaches of agent-based modeling and simulation models

Besides the models just discussed, two main other approaches emerge: spatial voting models and models inspired by psychology. The modeling under these approaches seeks not to assume a certain behavior of voters but to determine a distribution of preferences that is closest to reality. These approaches have the common feature of analyzing and generating preferences so as to reflect or to come close to real elections' data samples.

4.1 The spatial voting models

Spatial voting models were first applied specifically to elections by Downs (1957) to study the relative positioning of political parties and voters using a spatial approach built on the pioneering work of Black (1948), Hotelling (1929), Lerner and Singer (1937), Smithies (1941) and Greenhut (1956), who addressed the problem of location between two competing firms in order to optimally choose their setting in a market of undifferentiated goods. Under a spatial model, it is assumed that both candidates and voters are placed in a unidimensional or multidimensional space according to the position they take or prefer on certain issues, each of which corresponds to a dimension. In such a setting, a voter tends to choose the candidate who is closest to his/her position while a candidate will tend to choose a position that maximizes the number of electoral votes.

Let us notice that the most basic spatial model inspired by Downs (1957), involves an election based on a single dimension under which candidates can be ordered on a left-right axis, such that for each voter, his/her utility is decreasing with the distance to his/her preferred alternatives along this axis. Given his/her location (i.e., ideal or bliss point) and knowing the locations of the candidates on the spectrum, each voter casts a vote for the candidate who is closest to his/her location. The locations of the voters along the line follow a specific distribution and the Euclidean distance serves as a tool for measuring the elector-candidate proximity. For a given voter i and party or candidate j, if we denote by v_i the voter's position and by p_{ij} the party's position as perceived by voter i, the Downsian utility (U_{ij}) of voter i is given by:

$$U_{ij} = -\alpha \times (v_i - p_{ij})^2 \tag{5}$$

In Eq. 5, the overall policy importance is captured through the parameter α.

Besides the Downsian-inspired model, we note the existence of the so-called directional models. Under the directional model developed by Rabinowitz and Macdonald (1989),[16] it is assumed that utilities are determined by both the intensity and communality of direction of voters' and candidates' positions. So, voters have a diffuse preference for certain direction on an issue but vary in the intensity with which they hold that preference. Under this model, the voter's utility is a product of the policy positions of the voter i and the party j:

$$U_{ij} = \alpha \times v_i \times p_{ij} \tag{6}$$

We owe to Rabinowitz and Macdonald (1989) the introduction of a mixed model that combines the directional and the Downsian logic: a voter's choice is determined both by a proxy of proximity and by a directional component. A voter's utility under this model is defined as follows:

$$U_{ij} = \alpha \times \big(-\beta(v_i - p_{ij})^2 + (1-\beta)v_i p_{ij} \big) \tag{7}$$

where $\beta \in [0,1]$ is a relative weight of the two components of voter utility. As one can see, when parameter β is equal to 1, we get the Downsian model; when $\beta = 0$, we get the directional model. More recently, Kedar (2005) introduced a model combining the Downsian approach with a compensatory component which captures the outcome orientation of the voters. According to Kedar (2005), when a voter is outcome-oriented, it is assumed that he/she compares the expected policy outcome P if all parties are elected and a counterfactual policy outcome P_{-p_j} where one party p_j is excluded from the policy process P; then, he/she will choose the party where the distance between the two scenarios is greatest, providing the party shifts the expected policy outcome in the desired direction. Under the compensational model, a voter's utility is defined as follows:

$$U_{ij} = \alpha\big(-\beta(v_i - p_{ij})^2 - (1-\beta)[(v_i - P)^2 - (v_i - P_{-p_j})^2] \big) + \delta_j z_i \tag{8}$$

where, given p_j the position of party and $s_j \in [0,1]$ the relative impact of party j such that $\sum_j s_j = 1$; $P = \sum_j s_j p_j$; δ_j is a vector indicating the effect of background variables z_i on voter utility for party j.

Each of the above models has been described in the one-dimensional framework; they are easily extensible and adaptable to the multidimensional framework. However, in the

[16] Please refer to Merrill and Grofman (1999) for an overview of all the so-called directional models.

multidimensional framework, resorting to Euclidean distances requires assumptions about agent preferences. These hypotheses are, for the most part, improbable and empirically unrealistic; worse, they can complicate the theoretical analyses (Enelow and Hinich, 1990). Hence the use of models based on simulations. So, utilizing data from real elections, many variations of spatial models have been used to test theoretical results and also to show how institutional context affects voter behavior. See for instance, the work of Merrill (1984), Plassmann and Tideman (2011), Tideman (1992) and Tideman and Plassmann (2013).

4.2 The behavioral social choice approach

Popularized by Regenwetter et al. (2006), *behavioral social choice* is the counterpart of the traditional social choice theory which integrates into the analysis some realistic psychological factors (limited rationality, cognitive biases, etc.) that can influence individuals's choices in a real world. A behavioral approach usually tries to confront "what should be" (the normative aspect) with "what is" (the empirical aspect). So, behavioral social choice compares how supposedly rational individuals should make their decisions with how real decision makers behave empirically. It provides a framework for crafting more realistic models of social choice by embedding social choice analysis into a psychological representation of preferences and choice behavior, alongside a statistical evaluation of these models against empirical data (see for instance, Regenwetter and Grofman, 1998, Regenwetter et al., 2002a,b, Tsetlin and Regenwetter, 2003); it also develops methodologies to (re)construct preference distribution from incomplete data (Regenwetter et al., 2006).

Behavioral social choice challenges the analyses carried out in social choice theory based on a priori assumptions on the distribution of agents' preferences. The group of authors behind behavioral social choice support the idea that the results obtained from the theoretical models are highly dependent on the a priori assumptions considered in generating elections scenarios; these hypotheses, by restricting the behavior of individuals to probabilistic distributions (normal law), are themselves very far from reflecting the behavior of individuals in the real world. Thus, the results of the theoretical models based on a priori assumptions tend to promote views that are too pessimistic regarding the probability of many voting events such as the Condorcet paradox. According to Popova et al. (2013), these results may magnify gloomy predictions found in the axiomatic literature on the inability of an electorate to make a group decision.

Behavioral social choice aims to empirically analyze the rules or methods of preference aggregation by abstracting useless and/or unsubstantiated assumptions about human behavior. It turns out, therefore, that for any analysis, one has to state, very explicitly, tested and validated hypotheses about human behavior. Behavioral social choice considers empirical data on social choice from an inferential statistical point of view. If the empirical data are considered as imperfect and incomplete reflections of the voters' preference, one must evaluate the replicability of social choice outcomes and assess to what extent one can be confident about the search for correct collective outcomes. Thus, under each behavioral model, the maximum likelihood estimate is used to calculate the probabilities of the voting events, and statistical confidence levels are generated through a nonparametric bootstrap (Efron, 1979).

Generally speaking, the main idea behind a bootstrap is to make, on sample data, inferences about an estimate of sample statistics (sample mean, standard deviation, etc.) for a population statistical parameters (its mean, its standard deviation, etc.). Concretely, going from a data sample of size N with complete or incomplete voters' preference from real elections, it proceeds by carrying out sampling with replacement: a sample of size N is independently drawn from the original sample with replacement and replicated T times. For each of the T bootstrap samples, the estimates of the population parameters are evaluated; then a sampling distribution is built with all these estimates and used for the statistical inference. Ideally, it would be nice if T were large enough to ensure meaningful statistics; this is generally possible when using Monte Carlo simulations on fairly powerful computers by generating random samples.

In a behavioral social choice context, a bootstrap appears as a computer-based method for statistical inference that without relying on too many assumptions, is a way of simulating possible sources of uncertainty[17] in the results; it assesses how the results would be affected by small disturbances in the distribution of votes (preferences) and helps infer confidence levels about at which point estimates of model parameters would not be affected by such disturbances. Using this inference approach of preference aggregation, Regenwetter et al. (2006) and related papers have established the robustness of the empirical absence of majority cycles for a wide range of realistic modeling assumptions; they also came to the conclusion that the theoretical assumption quite often used in the literature give a pessimistic view, assigning high probabilities to the existence of electoral paradoxes, and indeed considering them as virtually certain when in fact in the real world this is not the case.

5 Concluding remarks

Over time, simulation models have emerged as an indispensable tool in many disciplines and fields of study. They offer a way to overcome the limits or constraints of theoretical modeling. In social choice theory as well as in political science or psychology, simulations quickly found their place as a way of dealing with the complexity of the topic and the challenges of the modeling of human behavior in a decision-making framework. They appear as a springboard allowing us to complete the analyses carried out in theoretical approach, or at least to question them. The models developed in theoretical work have shown some limits when it comes to modeling the behavior of individuals involved in a process of collective decision: models can become intractable; and indeed, given certain parameter values (number of agents, number of alternatives, etc.) some analyses are almost impossible. Moreover, the results obtained depend strongly on the assumptions on the behavior of the agents that support the models. It is also true that these assumptions are deemed relevant to a universe that is actually very far from reality.

[17] According to Regenwetter et al. (2002b, 2006) uncertainty can come from various factors, such as voters' uncertainty about their preference, unreliability of voter turnout, counting of ballots, etc.

Since social choice theory has been one of the areas in economics that has seen a boom in work using models based on the behavior of individuals involved in collective decision-making, the purpose of this paper has been to offer to the uninitiated in the social choice theory, a methodological presentation of some well-known models and the techniques of theoretical calculations and simulations, and then to report on recent developments of new models and advances in calculation techniques and simulations.

After briefly presenting the general framework of the aggregation of preferences, we presented the most widespread theoretical models and their extensions, and then discussed their strengths and weaknesses. We have particularly emphasized the two models that are most prevalent in the literature: the model of impartial culture (IC) and that of impartial and anonymous culture (IAC). The model IC, introduced by Guilbaud (1952), is based on the idea that all preference profiles are equiprobable and that each individual chooses his/her preference in a uniform probability distribution. For instance, when the individual preferences are expressed as linear orders on a set of alternatives, the IC assumption indicates that the preference relation of each voter is drawn uniformly at random from the set of all possible linear orders. On the other hand, the IAC model, introduced by Gehrlein and Fishburn (1976) and Kuga and Nagatani (1974), assumes the equiprobability of voting situations. Most theoretical results are based on these two models. Since these models are special cases of the multinomial law, one of their limitations lies in the fact that even for a limited number of alternatives and individuals, the multinomial law becomes difficult to manage. Indeed, we have shown how simulation models (notably with the Monte Carlo method) developed under these models may be helpful in analyzing complex problems in social choice theory; they have made it possible to validate or invalidate several results established in the literature. In short, the simulations implemented under these assumptions have helped to produce well-known and robust results in the field of preference aggregation.

The theoretical modeling has been strongly criticized for being based on distributions that do not reflect what happens in real elections; in fact, there is no evidence that voters' choices obey any probabilistic distribution, and no work has ever supported or even established that the theoretical models reflect the reality in a particular situation. These criticisms gave rise to the emergence of modelling that is not built on a priori assumptions on the preferences of agents. In this paper, we have presented two approaches that fall within this framework: spatial voting models and behavioral social choice. Under spatial voting models, inspired by Downs (1957), it is assumed that both candidates and voters are placed in a unidimensional or multidimensional space according to the position they take or prefer on certain issues, each of which corresponds to a dimension. In such a setting, a voter tends to choose the candidate who is closest to his/her position, while a candidate will tend to choose a position that maximizes the number of electoral votes. According to Merrill (1984), Plassmann and Tideman (2011) and Tideman and Plassmann (2013), when generating candidates and voters by means of simulations based on a spatial model, outcomes come astonishingly close to describing the distribution of actual outcomes, and ranking data simulated with the spatial model are very similar to observed ranking data. The spatial-model results thus tend to be more realistic. Behavioral social choice, popularized by the book of Regenwetter et al. (2006), provides a framework for crafting more realistic models of social choice by embedding social choice analysis into a psychological representation of preferences and choice behavior, and a statistical evaluation of these models against empirical data; it also develops methodologies to (re)construct preference distributions from incomplete data. Contrary to the theoretical models, these two approaches describe a modeling in which one confronts "what must be" with "what is", the goal being to get as close as possible to what happens in real situations of collective decision. Practice has shown that the developed models perform well in this task.

Remarkable advances in computer science and mathematical and statistical calculation techniques are giving more and more prominence to simulations. This suggests that new opportunities are opening to theorists to refine the results found in the literature, but also to revisit certain problems whose resolution was previously impossible.

References

Aleskerov, Fuad, Daniel Karabekyan, Remzi M. Sanver, and Vyacheslav Yakuba. 2012. On the manipulability of voting rules: The case of 4 and 5 alternatives. *Mathematical Social Sciences*, 64: 67-73.

Arrow, Kenneth J. 1963. *Social Choice and Individual Values*. 2nd edition, Wiley, New York.

Arrow, Kenneth J. 1951. *Social Choice and Individual Values*. 1st edition, Wiley, New York.

Axelrod, Robert. 1997. Advancing the art of simulation in the social sciences. In: Conte, R., Hegselmann, R., and Terna, P. (eds) *Simulating Social Phenomena*. Lecture Notes in Economics and Mathematical Systems. Springer, Berlin, Heidelberg, vol 456, 21-40.

Bentham, Jeremy. 1789. *An Introduction to the Principles of Morals and Legislation*. Clarendon Press, Oxford.

Berg, Sven. 1985a. A note on plurality distortion in large committees. *European Journal of Political Economy*, 1(2): 271-284.

Berg, Sven. 1985b. Paradox of voting under an urn model: the effect of homogeneity. *Public Choice*, 47: 377-387.

Berg, Sven, and Bo H. Bjurulf. 1983. A note on the paradox of voting : anonymous preference profiles and May's formula. *Public Choice*, 40: 308-316.

Bergson, Abram. 1938. A reformulation of certain aspects of welfare economics. *Quarterly Journal of Economics*, 52: 310-334.

Bjurulf, Bo H. 1972. A probabilistic analysis of voting blocs and the occurrence of the Paradox of Voting. In: Niemi, R., and Weisberg, H.F. (eds), *Probability models of collective decision making*. Charles E. Merrill, Columbus, OH, 232-251.

Black, Duncan. 1976. Justification of the Borda count. *Public Choice*, 28: 1-15.

Black, Duncan. 1958. *The Theory of Committees and Elections*. Cambridge University Press, Cambridge.

Black, Duncan. 1948. On the rationale of group decision making. *Journal of Political Economy*, 56: 23-34.

de Borda, Jean-Charles. 1781. Mémoire sur les élections au scrutin. Histoire de l'Académie Royale des Sciences, Paris.

Brandt, Felix , Christian Geist, and Martin Strobel. 2016. Analyzing the Practical Relevance of Voting Paradoxes via Ehrhart Theory, Computer Simulations, and Empirical Data. *Proceedings of the 2016 International Conference on Autonomous Agents, AAMAS'16*: 385-393.

Campbell, Colin D., and Gordon Tullock. 1965. A measure of the importance of cyclical majorities. *Economic Journal*, 75: 853-857.

Cervone, Davide, William V. Gehrlein, and William Zwicker. 2005. Which scoring rule maximizes Condorcet efficiency under IAC? *Theory and Decision*, 58: 145-185.

Chamberlin, John R, Jerry L. Cohen, and Clyde H. Coombs. 1984. Social choice observed: Five presidential elections of the American Psychological Association. *The Journal of Politics*, 46: 479-502.

Condorcet, Nicolas de C. 1785. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Imprimerie Royale, Paris.

David, Florence N., and Colin L. Mallows. 1961. The variance of spearman's rho in normal samples. *Biometrika*, 48: 19-28.

DeMeyer, Frank, and Charles R. Plott. 1970. The probability of a cyclical majority. *Econometrica*, 38: 345-354.

Diss, Mostapha, Vincent Merlin, and Fabrice Valognes. 2010. On the Condorcet efficiency of approval voting and extended scoring rules for three alternatives. In: Laslier, J-F., and Sanver, R.M. (eds) *Handbook on approval voting*. Springer, Berlin, pp 255-283.

Diss, Mostapha, and Ahmed Doghmi. 2016. Multi-winner scoring election methods: Condorcet consistency and paradoxes. *Public Choice*, 169(1): 97-116.

Dobra, John L. 1983. An approach to empirical measures of voting paradoxes : an update and extension. *Public Choice*, 41(2): 241-250.

Dobra, John L., and Gordon Tullock. 1981. An approach to empirical measures of voting paradoxes. *Public Choice*, 36: 193-194.

Downs, Anthony. 1957. An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, 65(2): 135-150.

Edgeworth, Francis Y. 1881. *Mathematical psychics: An Essay on the Application of Mathematics to the Moral Sciences*. Kegan Paul, Londres.

Efron, Bradley. 1979. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1): 1-26.

El Ouafdi, Abdelhalim, Issofa Moyouwou, and Hatem Smaoui. 2020. IAC-probability calculations in voting theory: Progress report. In: Diss M., and Merlin, V. (eds), *Evaluating Voting Systems with Probability Models, Essays by and in honor of William Gehrlein and Dominique Lepelley*, Springer, Berlin.

Enelow, James M., and Melvin J. Hinich. 1990. *Advances in the Spatial Theory of Voting*. Cambridge University Press.

Feix, Marc, and Jean-Louis Rouet (2005), L'espace des phases électoral et les statistiques quantiques. Applications à la simulation numérique. Working paper. https://halshs.archives-ouvertes.fr/halshs-00003973.

Felsenthal, Dan S. 2012. Review of Paradoxes Afflicting Procedures for Electing a Single Candidate. In: Felsenthal, D., and Machover, M. (eds), *Electoral Systems : Paradoxes, Assumptions, and Procedures*, Studies in Choice and Welfare, Springer, Berlin, Heidelberg, 19-91.

Felsenthal, Dan S., and Hannu Nurmi. 2018. *Voting Procedures for Electing a Single Candidate*. Springer.

Fishburn, Peter C., and William V. Gehrlein. 1977. An Analysis of Voting Procedures with Nonranked Voting. *Behavioral Science*, 22: 178-185.

Fishburn, Peter C., and William V. Gehrlein. 1980. The paradox of voting: Effects of individual indifference and intransitivity. *Journal of Public Economics*, 14: 83-94.

Fishburn, Peter C., and William V. Gehrlein. 1982. Majority efficiencies for simple voting procedures: Summary and interpretation. *Thery and Decision*, 14 (2): 141-153.

Fontana, Magda. 2006. Simulation in economics: evidence on diffusion and communication. *Journal of Artificial Societies and Social Simulation*, 9(2).

Garman, Mark B., and Morton I. Kamien. 1968. The paradox of voting: Probability calculations. *Behavioral Science*, 13: 306-316.

Gehrlein, William V. 1978. Condorcet winners in dual cultures. Presented at National Meeting of Public Choice Society, New Orleans, LA.

Gehrlein, William V. 1985. The Condorcet criterion and committee selection. *Mathematical Social Sciences*, 10: 199-209.

Gehrlein, William V. 1997. Condorcet's paradox and the Condorcet efficiency of voting rules. *Mathematica Japonica*, 40: 173-199.

Gehrlein, William V. 2006. *Condorcet's paradox*. Springer Berlin Heidelberg.

Gehrlein, William V., and Peter C. Fishburn. 1976. Condorcet's paradox and anonymous preference profiles. *Public Choice*, 26: 1-18.

Gehrlein, William V., and Peter C. Fishburn. 1978a. Coincidence probabilities for simple majority and positional voting rules. *Social Science Research*, 7: 272-283.

Gehrlein, William V., and Peter C. Fishburn. 1978b. Probabilities of election outcomes for large electorates. *Journal of Economic Theory*, 19: 38-49.

Gehrlein, William V., and Dominique Lepelley. 2004. Probability calculations in voting theory: An overview of recent results. In: Wiberg, M. (ed) *Reasoned choices: Essays in honor of Academy Professor Hannu Nurmi*. The Finnish Political Science Association, Turku, Finland, 140-160.

Gehrlein, William V., and Dominique Lepelley. 2011. *Voting paradoxes and group coherence*. Springer, Berlin/Heidelberg.

Gehrlein, William V., and Dominique Lepelley. 2017. *Elections, Voting Rules and Paradoxical Outcomes*. Publisher: Springer-Verlag.

Guilbaud, Georges-Théodule. 1952. Les théories de l'intérêt général et le problème logique de l'aggrégation. *Economie Appliquée*, 5: 501-584.

Greenhut, Melvin L. 1956. *Plant Location in Theory and Practice, The Economics of Space*. Chapel Hill The University of North Carolina Press.

Hotelling, Harold. 1929. Stability in Competition. The *Economic Journal*, 39(153): 41-57.

Johnson, Norman L., and Samuel Kotz. 1977. *Urns Models and their Application*. Wiley, New York.

Jones, Bradford,Benjamin Radcliff, Charles Taber, and Richard Timpone. 1995. Condorcet winners and the paradox of voting: Probability calculations for weak preference orders. *American Political Science Review*, 89: 137-144.

Kedar, Orit. 2005. When Moderate Voters Prefer Extreme Parties: Policy Balancing in Parliamentary Elections. *American Political Science Review*, 99(2): 185-199.

Kelly, Jerry S. 1993. Almost all social choice rules are highly manipulable, but a few aren't. *Social Choice and Welfare*, 2: 161-175.

Klahr, David. 1966. A computer simulation of the paradox of voting. *American Political Science Review*, 60: 384-390.

Kuga, Kiyoshi, and Nagatani Hiroaki. 1974. Voter antagonism and the paradox of voting. *Econometrica*, 42: 1045-1067.

Kurrild-Klitgaard , Peter. 2001. An empirical example of the Condorcet paradox of voting in a large electorate. *Public Choice*, 107: 135-145.

Kurrild-Klitgaard, Peter. 2008. Voting paradoxes under proportional representation: Evidence from eight Danish elections. *Scandinavian Political Studies*, 31(3): 242-267.

Laslier, Jean-François. 2010. In Silico Voting Experiments. In: Laslier, J-F., and Sanver, R.M. (eds), *Handbook on Approval Voting*, Studies in Choice and Welfare. Springer, Berlin, Heidelberg, pp 311-335.

Le Breton, Michel, Dominique Lepelley, and Hatem Smaoui. 2016. Correlation, partitioning and the probability of casting a decisive vote under the majority rule. *Journal of Mathematical Economics*, 64: 11-22.

Lepelley, Dominique, Ahmed Louichi, and Hatem Smaoui. 2008. On Ehrhart polynomials and probability calculations in voting theory. *Social Choice and Welfare*, 30(3): 363-383.

Lepelley, Dominique, Ahmed Louichi, and Fabrice Valognes. 2000. Computer simulations of voting systems. *Advances in Complex Systems*, 03(01n04): 181-194.

Lerner, Abba P., and Hans W. Singer. 1937. Some Notes on Duopoly and Spatial Competition. *Journal of Political Economy* 45(2): 145-186.

List, Christian. 2013. *Social Choice Theory*. The Stanford Encyclopedia of Philosophy.

Macal, Charles M., and Michael J. North. 2010. Tutorial on agent-based modelling and simulation. *Journal of Simulation*, 4(3): 151-162.

Marshall, Alfred. 1960. *Principles of Economics*. Macmillan, Londres, 9th edition.

May, Kenneth O. 1952. A set of independent necessary and sufficient conditions for simple majority decisions. *Econometrica*, 20: 680-684.

May, Kenneth O. 1948. Probability of Certain Election Results. *American Mathematical Monthly*, 55(4): 203-209.

May, Robert M. 1971. Some mathematical remarks on the paradox of voting. *Behavioral Science*, 16: 143-153.

Merrill III, Samuel. 1984. A comparison of efficiency of multicandidate electoral systems. *American Journal of Political Science*, 28(1): 23-48.

Merrill III, Samuel, and Bernard Grofman. 1999. *A unified theory of Voting: Directional and Proximity Spatial Models*. Cambridge: Cambridge University Press.

Miller, Nicholas R. 2012. Election Inversions by the U.S. Electoral College. In: Felsenthal, D.S., and Machover, M. (eds), *Electoral Systems : Paradoxes, Assumptions, and Procedures*, Studies in Choice and Welfare, Springer, Berlin, Heidelberg, 93-127.

Niemi, Richard G. 1970. The occurrence of the paradox of voting in university elections. *Public Choice*, 8: 91-100.

Niemi, Richard G., and Weisberg, Herbert F. 1968. A mathematical solution for the probability of the paradox of voting. *Behavioral Science*, 13: 317-323.

Nurmi, Hannu. 1999. *Voting Paradoxes and How to Deal with Them*. Springer-Verlag.

Nurmi, Hannu. 1989. *Comparing Voting Systems*. Dordrecht: B.Reidel.

Pigou, Arthur C. 1920. *The Economics of Welfare*. Macmillan, Londres.

Plackett, Robin L. 1954. A reduction formula for normal multivariate integrals. *Biometrika*, 41(3/4): 351-360.

Plassmann, Florenz, and T. Nicolaus Tideman. 2011. How to predict the frequency of voting events in actual elections. Mimeo.

Pomeranz, John E., and Roman L. Weil. 1970. The cyclical majority problem. *Communications of the ACM*, 13(4): 251-254.

Popova, Anna, Michel Regenwetter, and Nicholas Mattei. 2013. A behavioral perspective on social choice. *Annals of Mathematics and Artificial Intelligence*, 68: 5-30.

Rabinowitz, George, and Stuart E. Macdonald. 1989. A Directional Theory of Issue Voting. *American Political Science Review*, 83(1): 93-121.

Regenwetter, Michel, and Bernard Grofman. 1998. Choosing subsets: a size-independent probabilistic model and the quest for a social welfare ordering. *Social Choice and Welfare*, 15: 423-443.

Regenwetter, Michel, James Adams, and Bernard Grofman. 2002a. On the (sample) Condorcet efficiency of majority rule: An alternative view of majority cycles and social homogeneity. *Theory and Decision*, 53: 153-186.

Regenwetter, Michel, Bernard Grofman, and Anthony A. J. Marley. 2002b. On the model dependence of majority preference relations reconstructed from ballot or survey data. *Mathematical Social Sciences*, 43: 451-466.

Regenwetter, Michel, Bernard Grofman, Anthony A. J. Marley, and Ilia A. Tsetlin. 2006. *Behavioral social choice*. Cambridge, UK: Cambridge University Press.

Riker, William H. 1958. The paradox of voting and congressional rules for voting on amendments. *American Political Science Review*, 52: 349-366.

Riker, William H. 1965. Arrow's theorem and some examples of the paradox of voting. In: Claunch, J.M. (ed), *Mathematical applications in political science*, Vol. I. Dallas: Southern Methodist University Press, 41-60.

Saari, Donald G. 1995. *Basic Geometry of Voting*. Springer.

Saari, Donald G, and Maria Tataru. 1999. The likelihood of dubious election outcomes. *Economic Theory*, 13: 345-363.

Samuelson, Paul A. 1947. *Foundations of Economic Analysis*. Harvard University Press, Cambridge.

Smithies, Arthur. 1941. Optimum location in spatial competition. *Journal of Political Economy*, 49: 423-439.

Straffin, Philip D. 1988. The Shapley-Shubik and Banzhaf power indices as probabilities. In: Roth, A. (ed), *The Shapley Value: Essays in Honor of Lloyd S. Shapley*, Cambridge: Cambridge University Press, 71-82.

Taylor, Alan D. 1997. A glimpse of impossibility: Kenneth Arrow's impossibility theory and voting. *Perspect Polit Sci*, 26: 23-26.

Tideman, Nicolaus T. 1992. Collective decision and voting: Cycles. Presented at Public Choice Society Meeting, New Orleans, LA.

Tideman, Nicolaus T., and Florenz Plassmann. 2013. Developing the aggregate empirical side of computational social choice. *Annals of Mathematics and Artificial Intelligence*, 68: 31-64.

Tsetlin, Ilia A., and Michel Regenwetter. 2003. On the probability of correct or incorrect majority preference relations. *Social Choice and Welfare*, 20: 283-306.

Von Neumann, John, and Stan Ulam. 1945. The Monte Carlo Method. *Bulletin of the American Mathematical Society*, 51-9-165.

Weber, Rodney J. 1978. Comparison of voting systems. Yale University, unpublished manuscript.

Weisberg, Herbert F., and Richard G. Niemi. 1978. Probability calculations for cyclical majorities in Congressional voting. In: Niemi, R.G., and Weisberg, H.F. (eds), *Probability models of collective decision making*. Charles E. Merrill, Columbus, OH, 181-203.

Weisberg, Herbert F., and Richard G. Niemi. 1973. A pairwise probability approach to the likelihood of the paradox of voting. *Behavioral Science*, 18: 109-117.

Wilson, Mark C., and Geoffrey Pritchard. 2007. Probability calculations under the IAC hypothesis. *Mathematical Social Sciences*, 54: 244-256.