Probability Models for the Analysis of Voting Rules in a Federal Union

Marc Feix
SUBATECH, Ecole des Mines de Nantes
Dominique Lepelley
Université de la Réunion
Vincent Merlin
CREM, Université de Caen and CNRS
Jean Louis Rouet
MAPMO and ISTO, Université d'orléans

March 20007

Two-tiers elections: Country divided into jurisdictions where “local” elections are held. Local winners are aggregated at the “federal” level. Some problems :

1. US 2000 election. Bush get 271 votes against 267 in the Electoral College thanks to his short victory in Florida.

2. European Union: What is the best decision rule for the council of minister of the enlarged EU27 ?

3. Gerrymandering: redesigning the boundaries of jurisdictions. http://www.nationalatlas.gov/prin

4. UK 2005, MP elections: Some internet sites tried to coordinate vote swapping between Liberal Democrats and Labour voters in order to beat Conservatives.

The Probabilistic Approach

How many seats per jurisdiction ?

The whole literature on apportionment, see Balinski and Young (1982)

First criteria: equalize power. Penrose (1946), Banzhaf (1968) and the whole literature on power indices.$\Longrightarrow$ Square root rule or proportional rule.

Second criteria: Maximize the total utility. Felsenthal and Machover (1999), Barberà and Jackson (2004), Beisbart, Bovens Hartmann (2005).$\Longrightarrow$ Square root rule or proportional rule.

Third criteria: Minimize the probability of conflict : Feix, Lepelley, Merlin, Rouet and Vidu.

All these works are based upon a priori probabilistic voting models.

Question: Is there an apportionment method that minimizes the probability of conflicts in federal systems ?

In *The Probability of Conflicts in a U.S. Presidential Type Election* (Economic Theory, 2004) we first compute the probability of the paradox when there are two parties, when **all the states, constituencies, regions, etc. have the same population size**. Thus, we obtain a benchmark result, when there is no bias due to the apportionment method, or a different turnout across the states.

In a second paper, *Majority Efficient Representation of the Citizens in a Federal Union*, we introduce states **with different population sizes**. According to the voting model, we find different optimal apportionment rules. Then we compare our results to the ones obtained in the literature when the objective is to give equal power to each citizen.

States with unequal populations.

A finite set $N = \{1, \ldots, i, \ldots, M\}$ of states (or regions, districts, etc.) form a political union.

We assume that m_i voters live in state i.

Two parties, A and B compete in all the states.

The winner in state i is the party who obtains a majority of voters on his side (abstention is not allowed).

Each state is represented by a_i mandates in the union, and the winner in state i gets all the mandates.

Thus, the position that is officially adopted by the union is the one which obtains a majority of mandates at the federal level.

The objectives

We adopt here the following normative criteria to evaluate the different apportionment methods : *an apportionment method is said to be majority efficient if it minimizes the probability that a decision is taken with a majority of mandates at the federal level though it is supported by a minority of voters over the whole union.*

More precisely, we try to identify the optimal value of the parameter α in the apportionment rule:

$$a_i = m_i^\alpha$$

We also try to identify the influence of the dispersion of the population on the magnitude of the paradox.

Probability models

The Impartial Culture assumption, IC.

Each voter selects a party with equal probability. When m_i is sufficiently large, the distribution of the votes follows a normal law. In each state, the excess of ballots for A or B is then given by

$$\varepsilon_i \sqrt{m_i}$$

where ϵ_i is drawn randomly according to the Gauss distribution:

$$(2\pi)^{-1/2} \exp(-\varepsilon^2/2).$$

The popular over the whole union is given by

$$\text{sign}\left(\sum_i \varepsilon_i \sqrt{m_i}\right), \qquad (1)$$

while the decision taken by the representatives is given by

$$\text{sign}\left(\sum_i a_i \,\text{sign}(\varepsilon_i)\right) \qquad (2)$$

The Impartial Culture Assumption, IAC.

This assumption considers that every distribution of the votes between the two candidates is equally likely to occur. Thus, ϵ_i is now drawn from the distribution $f(\epsilon) = 1/2$ if $-1 < \epsilon < 1$ and 0 otherwise. The excess of ballots in favor of candidate A in state i is given by $\epsilon_i m_i$ and we have to compare

$$\text{sign}\left(\sum_i \varepsilon_i m_i\right) \tag{3}$$

for the popular vote with the vote of the representatives

$$\text{sign}\left(\sum_i a_i \,\text{sign}(\varepsilon_i)\right) \tag{4}$$

The three state case

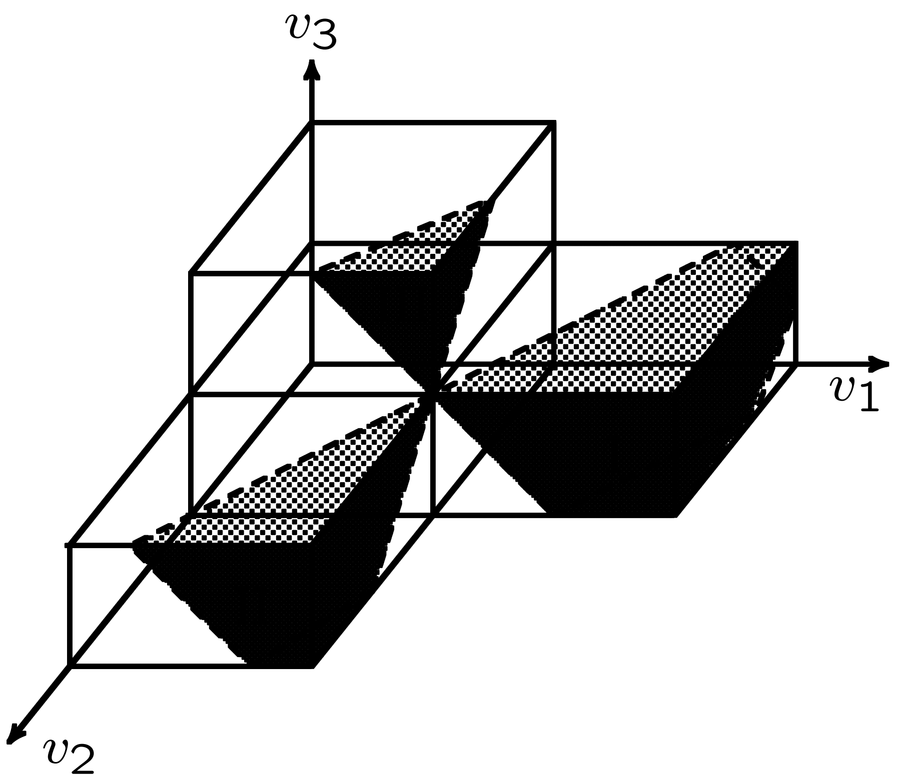
The distribution of the population is given by the vector $m = (m_1, m_2, m_3)$, with $\sum_{i=1}^{3} m_i = 1$, $m_1 \geq m_2 \geq m_3$. The only interesting case in $a_1 = a_2 = a_3$.

Proposition 1 *Let $P^3_{IAC}(m, \infty)$ be the likelihood of the majority paradox for three states of large population under IAC for the distribution m. Then, for $m_1 < \frac{1}{2}$ we get :*

$$P^3_{IAC}(m,\infty) = \frac{m_1^3 + m_2^3 + m_3^3 - (m_1 - m_2)^3}{24 m_1 m_2 m_3} \frac{-(m_1 - m_3)^3 - (m_2 - m_3)^3}{24 m_1 m_2 m_3}$$

And for $n_1 > \frac{n}{2}$, we obtain:

$$P^3_{IAC}(m,\infty) = \frac{(m_2 + m_3)^3 - (m_2 - m_3)^3}{24 m_1 m_2 m_3} + \frac{6 m_2 m_3 (m_1 - m_2 - m_3)}{24 m_1 m_2 m_3}$$



The volumes for the referendum paradox under IAC, with $n_1 = 4$, $n_2 = 3$ and $n_3 = 2$

The probability $P^3_{IAC}(m, \infty)$..

$n_3 \rightarrow$ $n_2 \downarrow$	0^+	0.05	0.10	0.15	0.20	0.25	0.30	0.333
0^+	0.25^-	$--$	$--$	$--$	$--$	$--$	$--$	$--$
0.05	0.25	0.240	$--$	$--$	$--$	$--$	$--$	$--$
0.10	0.25	0.237	0.229	$--$	$--$	$--$	$--$	$--$
0.15	0.25	0.236	0.224	0.214	$--$	$--$	$--$	$--$
0.20	0.25	0.234	0.220	0.206	0.194	$--$	$--$	$--$
0.25	0.25	0.233	0.216	0.200	0.183	0.166	$--$	$--$
0.30	0.25	0.231	0.213	0.193	0.172	0.150	0.124	$--$
0.333	0.25	0.230	0.210	0.188	0.164	0.143	0.128	0.125
0.35	0.25	0.230	0.208	0.185	0.161	0.140	0.127	$--$
0.40	0.25	0.228	0.204	0.178	0.156	$--$	$--$	$--$
0.45	0.25	0.225	0.200	$--$	$--$	$--$	$--$	$--$
0.50	0.25	$--$	$--$	$--$	$--$	$--$	$--$	$--$

Proposition 2 *Let $P^3_{IC}(m,\infty)$ be the likelihood of the majority paradox for three states of large population under IC for the distribution m. Then:*

$$P^3_{IC}(m,\infty) \;=\; \frac{\sum_{i=1}^{3} \text{arccos}\left(\sqrt{m_i}\right)}{\pi} - 0.75$$

The probability $P^3_{IC}(m,\infty)$.

$n_3 \rightarrow$ $n_2 \downarrow$	0^+	0.05	0.10	0.15	0.20	0.25	0.30	0.333
0^+	0.25^-	$--$	$--$	$--$	$--$	$--$	$--$	$--$
0.05	0.25^-	0.208	$--$	$--$	$--$	$--$	$--$	$--$
0.10	0.25^-	0.202	0.192	$--$	$--$	$--$	$--$	$--$
0.15	0.25^-	0.199	0.187	0.181	$--$	$--$	$--$	$--$
0.20	0.25^-	0.197	0.184	0.177	0.172	$--$	$--$	$--$
0.25	0.25^-	0.196	0.182	0.174	0.169	0.166	$--$	$--$
0.30	0.25^-	0.195	0.181	0.173	0.167	0.164	0.163	$--$
0.333	0.25^-	0.194	0.180	0.172	0.167	0.164	0.162	0.162
0.35	0.25^-	0.194	0.180	0.171	0.166	0.163	0.162	$--$
0.40	0.25^-	0.194	0.179	0.171	0.166	$--$	$--$	$--$
0.45	0.25^-	0.194	0.199	$--$	$--$	$--$	$--$	$--$
0.50	0.25^-	$--$	$--$	$--$	$--$	$--$	$--$	$--$

Computer simulations

We assume that $a_i = m_i^\beta$ for the IC case and $a_i = m_i^\alpha$ for the IAC case. We then run several computer simulations for different values of the vector m in order to find the optimal values of α and β.

Consider the vector of population weights (m_i^2) in equation (1). With this new variable, it becomes identical to equation (3). Similarly, if we assume that $2\beta = \alpha$, equations (2) and (4) are now the same. Thus, we can state that, to some extent, the results with the IAC assumption are "transposable" to the IC case for populations raised to the square and a distribution of the mandates such as $\beta = 2\alpha$.

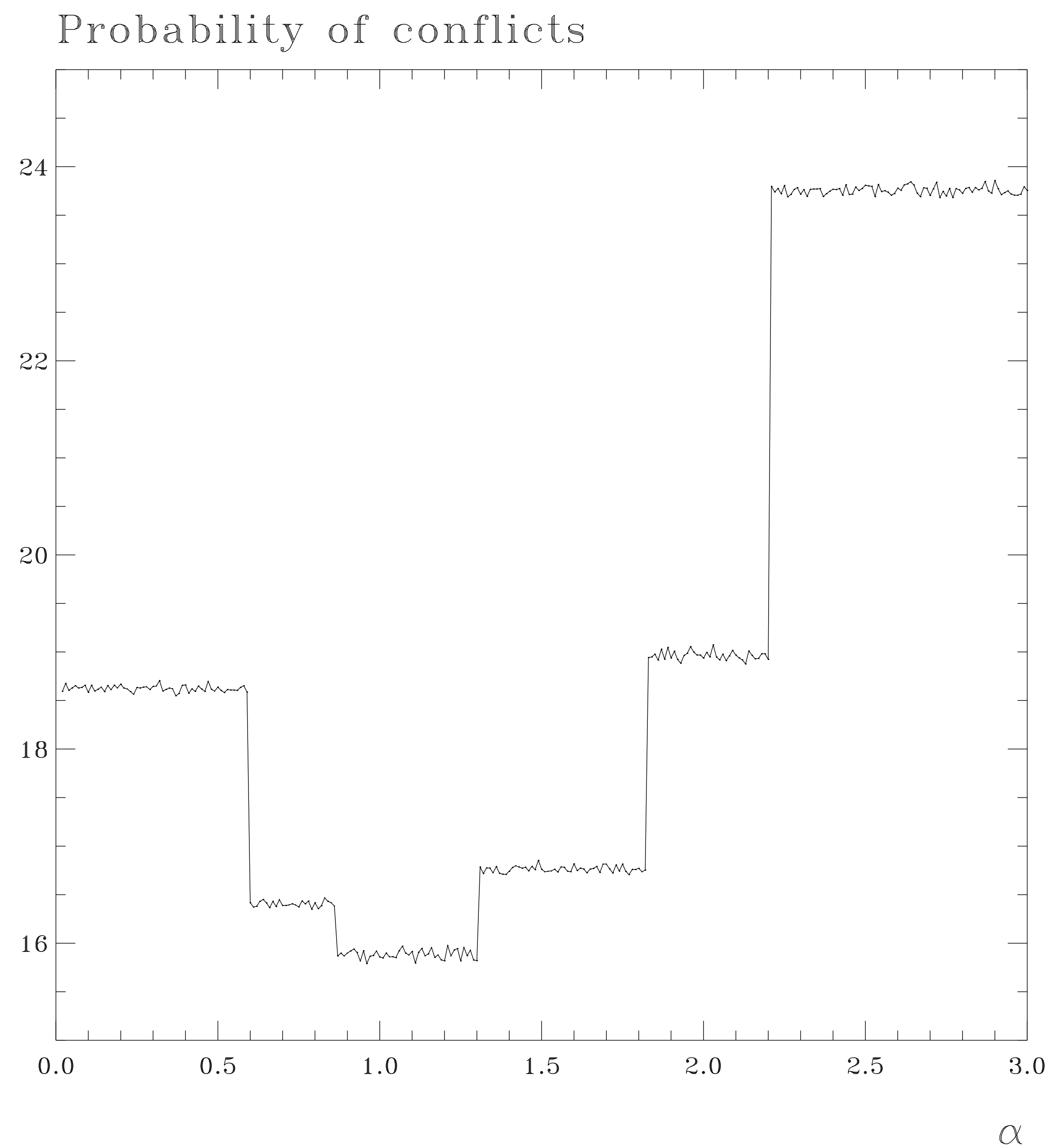


Figure 1: likelihood of conflicts between the popular vote and the vote by states, estimated after 1 000 000 draws for IAC model and $m_i = 1, 1.4, 1.8, 2.3, 3.2$.

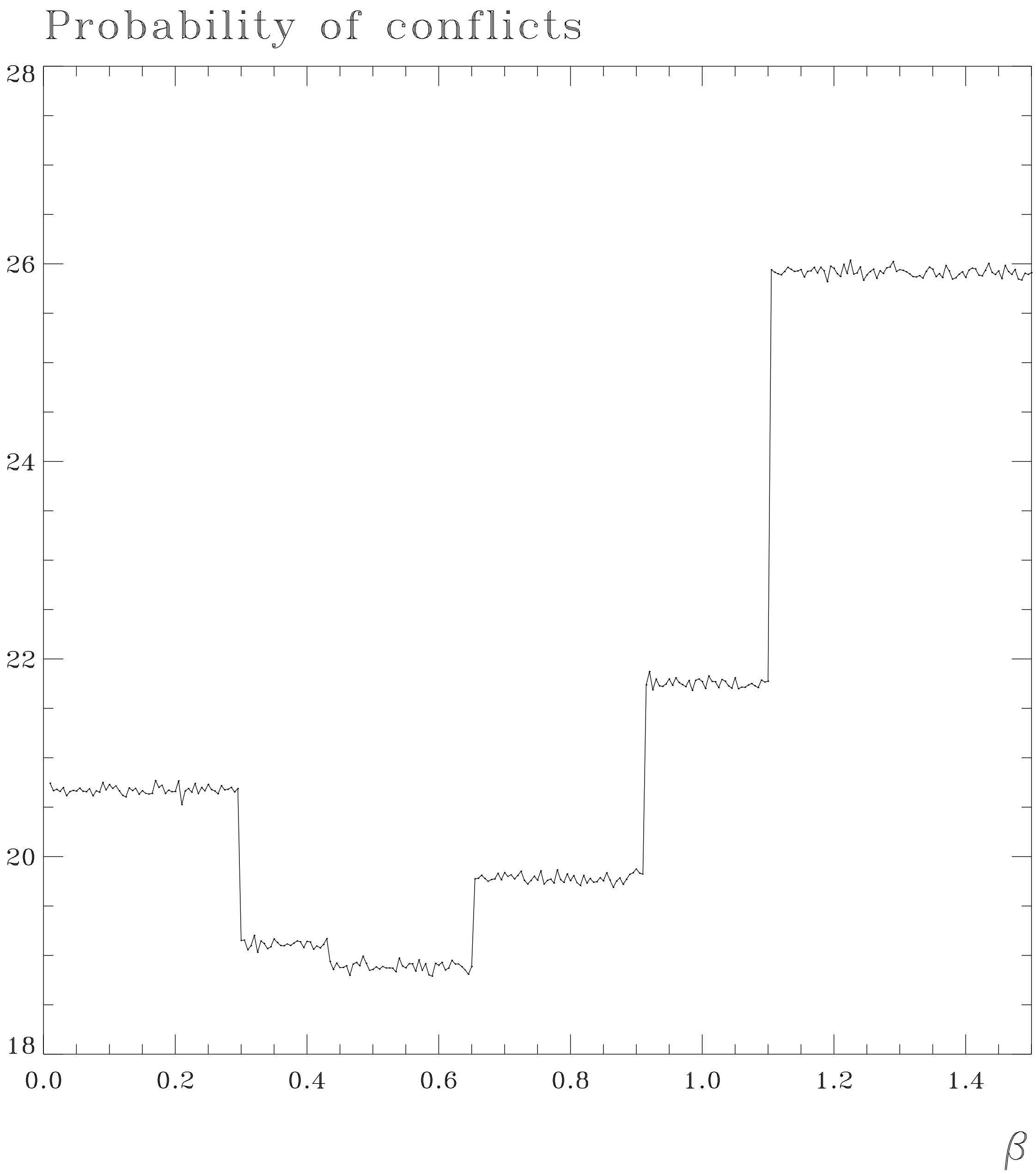


Figure 2: likelihood of conflicts between the popular vote and the vote by states, estimated after 1 000 000 draws for IC model and $m_i = 1, 1.96, 3.24, 5.29, 10.24$.

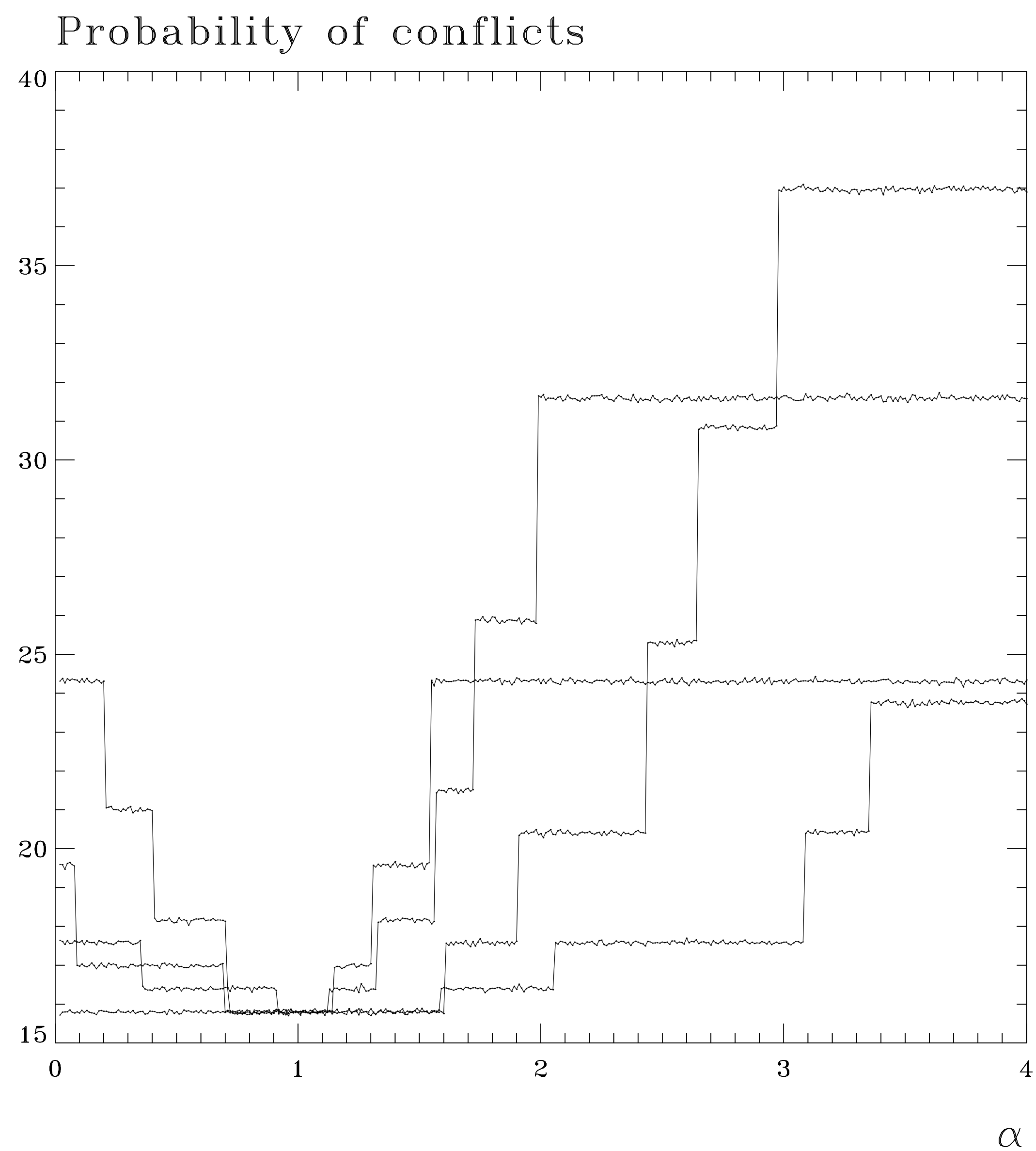


Figure 3: Testing the optimality of the $a_i = m_i$ rule for IAC model and the 5 state case, estimated after 1 000 000 draws.

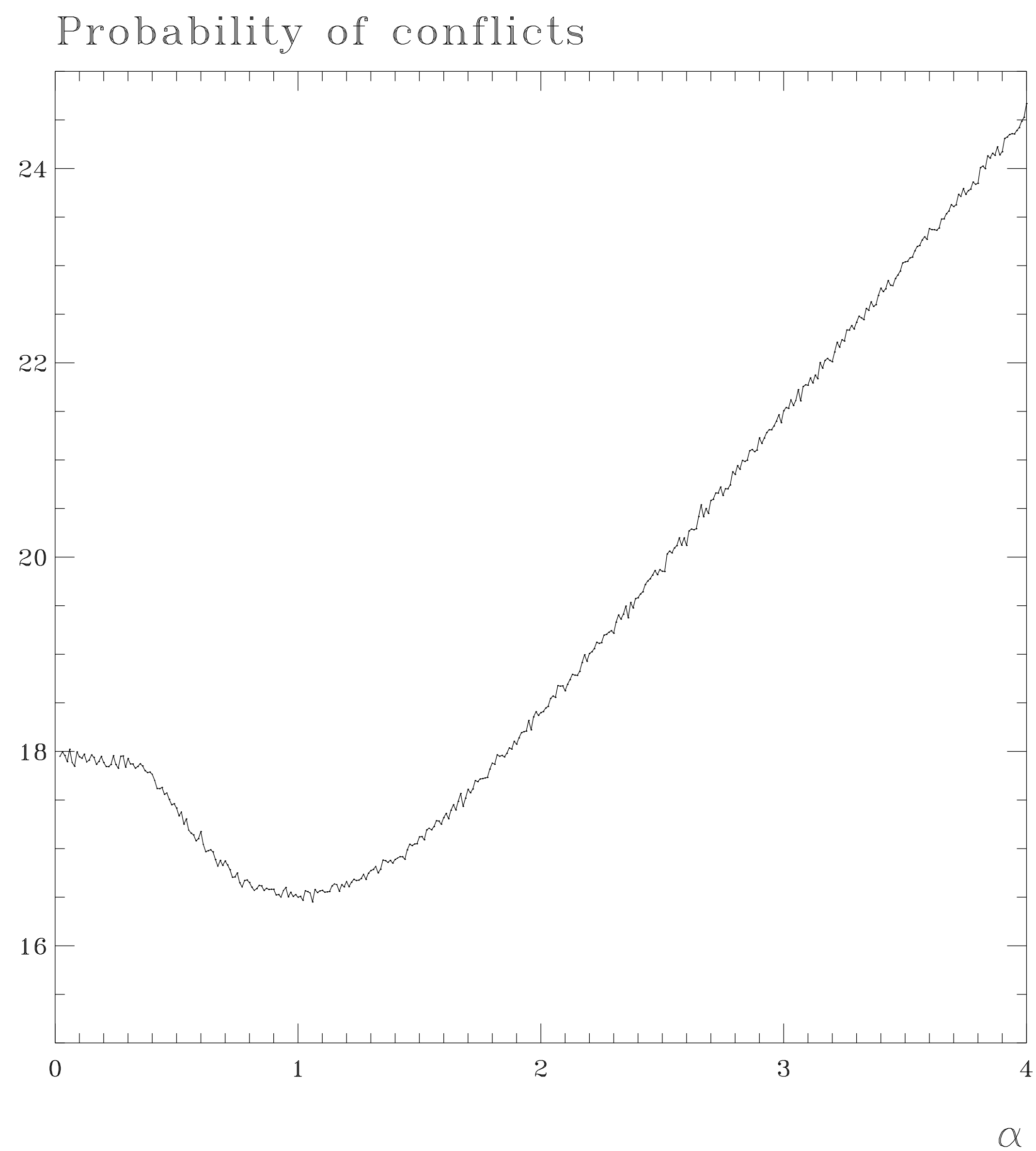


Figure 4 likelihood of conflicts between the popular vote and the vote by states, estimated after 1 000 000 draws for IAC model and m_i= 1, 1.31, 1.53, 1.77, 2.12, 2.54, 2.97, 3.15, 3.39, 3.44, 3.75, 3.93, 4.16, 4.21, 4.34, 4.85, 5.54, 5.72, 5.99.

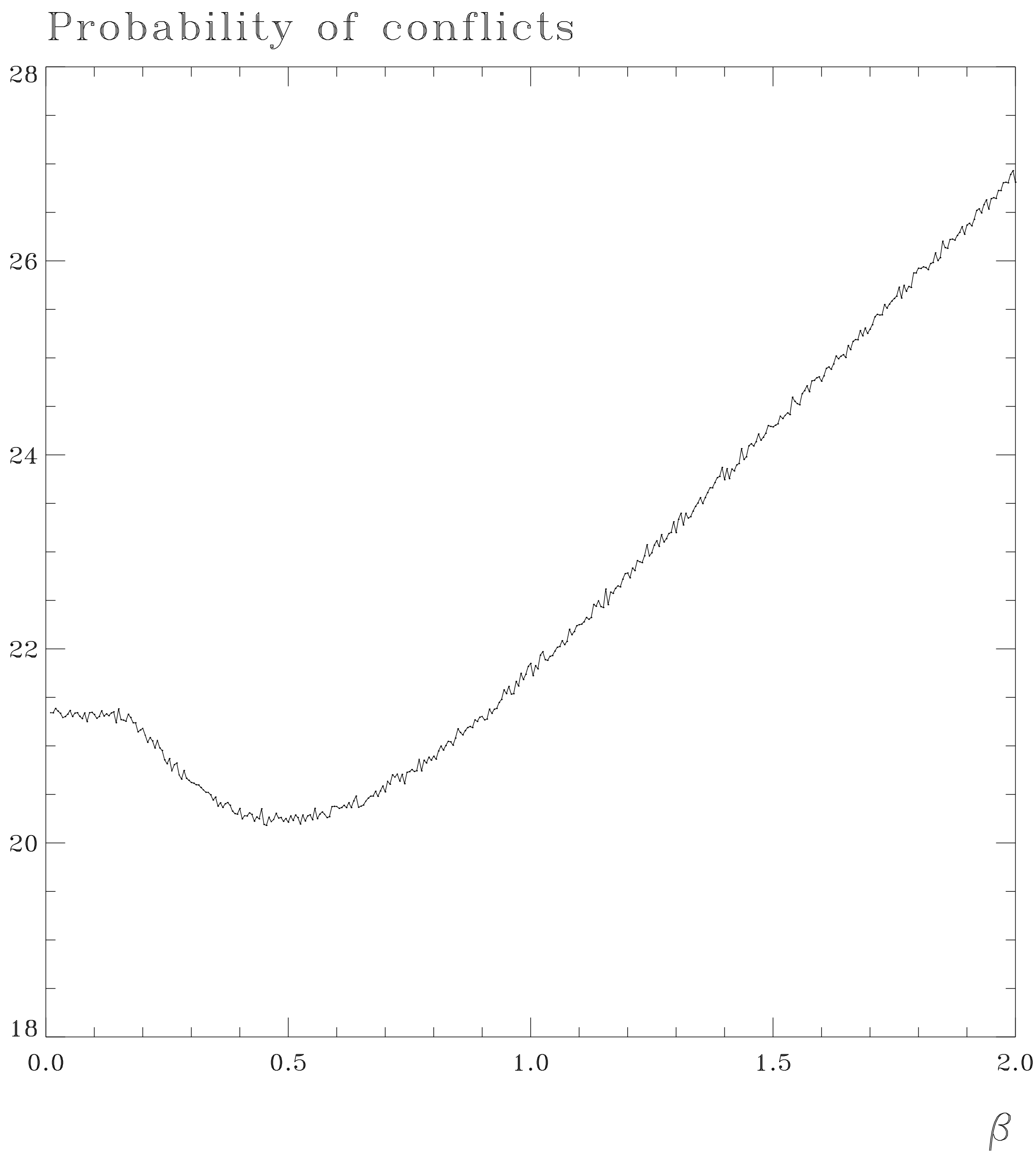


Figure 5: likelihood of conflicts between the popular vote and the vote by states, estimated after 1 000 000 draws for IC model and m_i = 1, 1.71, 2.34, 3.13, 4.49, 5.81, 6.45, 8.82, 9.92, 11.49, 11.83, 14.06, 15.44, 17.31, 17.72, 18.84, 23.52, 30.69,32.72, 35.88.

CONJECTURE For N states characterized by the populations m_i voting under the IC model (respectively IAC), the number of representatives a_i for state i should vary as $m_i^{1/2}$ (respectively m_i) in order to minimize the likelihood of conflicts. (All the votes being taken through majority votes).

Thus, we would get the same results independently of the criteria:

- It is known that equal power among the citizens is achieved (on average, for N large, with no dominant state) with the square root rule for the Banzhaf power index (IC), and with the proportional rule for the Shapley-Shubik index (IAC)

- Barberà and Jackson have discovered that the total utility of an union is maximized with a the square root rule with a model close to IC. Similarly, they found the proportional rule with a model close to IAC.

Some preliminary results

1) The conjecture may not be true for certain specifications of the population.

2) For IC, the average optimal value seems to be slightly below $\alpha = 0.5$. However, it is getting closer to 0.5 as N increases.

3) For IAC, the average optimal value seems to be $\beta = 1$ for all values of N.

4) The probability of the paradox seems to increase when a country approaches 50% of the population of the union. Another key parameter is the difference in population size between the top two states.

Question 1. The Probability of Disputable Election Outcome (Vincent Merlin, Jack Nagel).

For direct voting, an election is disputable if a party wins with less than $(50+\epsilon)\%$ of the vote nationwide, with ϵ small.

For indirect voting, an election is disputable whenever the constituencies where a party wins with less than $(50+\epsilon)\%$ are pivotal - either party that could catch them all would win a majority of seats.

We answer this question for direct and indirect voting mechanisms, using the majority rule over two candidates. We assume that all the constituencies are identical (equal population, no bias in favour of one party) and that abstention is not allowed. We use the IAC assumption: in each state, the result for party A is drawn randomly in the interval [0, 1] with a uniform distribution.

For an odd number of constituencies (N = 2k+1), we prove that the probability that the election is disputable in the Electoral College is approximated by

$$P_{EC}^{N}(\epsilon) \approx \frac{(2k+1)!(k+1)!}{k!(k+1)!2^{2k-1}}\epsilon$$

while the probability that the election is too close to call nationwide is approximated by:

$$P_{N}^{N}(\epsilon) \approx 2\epsilon\frac{\sqrt{6\pi N}}{\pi N} - 4\epsilon^{3}\frac{\sqrt{6\pi N}}{\pi N^{2}}$$

For equal size states, we get

$$P_{N}^{N}(\epsilon) < P_{EC}^{N}(\epsilon)$$

if ϵ is small.

What happens with unequal states ?

Question 2. Can we get better probability models ?

More recently, Gelman, Katz and Bafumi (British Journal of Political Sciences, 2004) show that the IC assumption can be tested, and that it is rejected when compared to electoral data from US. IAC has not been tested, but we need a model where the variance of the results from state to state does not depend on the the size of the population state.

Can we adapt the models to better fit with data ? This is a new trend in the literature, see the recent book Behavioral Social Choice, by Regenwetter, M., Grofman, B., Marley, A. A. J., Tsetlin, I. (2006).

Two PhD students from Caen working on this subject : Rahhal Lahrach, with data from French Local elections, and Thomas Senné, with US data.