

WORKING PAPERS

N° TSE-622

February 2016

## “Correlation, Partitioning and the Probability of Casting a Decisive Vote under the Majority Rule”

Michel Le Breton, Dominique Lepelley, Hatem Smaoui

# Correlation, Partitioning and the Probability of Casting a Decisive Vote under the Majority Rule\*

Michel Le Breton<sup>†</sup>    Dominique Lepelley<sup>‡</sup>    Hatem Smaoui<sup>§</sup>

February 2016

## Abstract

The main purpose of this paper is to estimate the probability of casting a decisive vote under the majority rule for a class of random electorate models encompassing the celebrated Impartial Culture (IC) and Impartial Anonymous Culture (IAC) models. The emphasis is on the impact of correlation across votes on the order of magnitude of this event. Our proof techniques use arguments from probability theory on one hand and combinatorial and algorithmic tools for counting integer points inside convex polytopes on the other hand.

**JEL Classification Numbers:** D71, D72.

**Keywords:** Elections, Power Measurement, Voting, Random Electorate.

---

\*This is the last version of a working paper which has circulated for the first time in January 2011. Early versions have been presented at the Social Choice and Welfare Meeting in Delhi and at the Computational Social Choice (COMSOC) meeting in Krakow. We thank all participants for their comments and suggestions.

<sup>†</sup>Toulouse School of Economics, France. E-mail address: michel.lebreton@tse-fr.eu.

<sup>‡</sup>CEMOI, Université de La Réunion, France. Corresponding author. E-mail address: dominique.lepelley@univ-reunion.fr. Fax: +33 (0)2 62 93 84 79.

<sup>§</sup>CEMOI, Université de La Réunion, France. E-mail address: hatem.smaoui@univ-reunion.fr.

# 1 Introduction

The main purpose of this paper is to introduce a general model of a random electorate of  $N$  voters described by their preferences over two alternatives. Our model will admit, as special cases, the two most popular models in the literature on power measurement. The first one, called *Impartial Culture* (IC) is the basis of the celebrated Penrose-Banzhaf power index (Penrose (1946), Banzhaf (1965)). It assumes that the preferences of the voters over the two alternatives are independent and equally likely: correlation among the preferences of the voters is totally precluded. The second one, called *Impartial Anonymous Culture* (IAC) which has been pioneered independently in voting theory by Chamberlain and Rothschild (1981), Good and Mayer (1975), Fishburn and Gehrlein (1976) and Kuga and Nagatani (1974) is the basis (as forcefully demonstrated by Straffin (1977, 1988)) of another celebrated power index due to Shapley and Shubik (Shapley and Shubik (1954), Straffin (1977, 1988)). The IAC model introduces correlation among voters and the specific distributional assumption which is considered implies that the real random variable defined as the number of voters supporting the first alternative is uniform over all feasible integers. From a computational perspective, this distributional property of the IAC model makes it very handy as compared to some other models and probably explains its success. Further, as noted convincingly by Chamberlain and Rothschild, the IAC model is more attractive than the IC model in the sense that the electoral predictions of the IAC models don't display a discontinuity in the neighborhood of the outcome of a tied election.

Given a random electorate  $\lambda$ , the power of a voter is defined as the probability of being pivotal<sup>1</sup> i.e. as the probability of being able to change the electoral outcome by his or her vote. Given that we will focus on a symmetric simple game (the ordinary majority game), if the model of random electorate  $\lambda$  is fully symmetric (i.e. if the preferences are interchangeable), then all voters will have the same power denoted  $Piv(\lambda, N)$ . Both the IC and the IAC models are symmetric. For the IC model, this defines the Penrose-Banzhaf power index  $Piv(IC, N)$  while for the IAC model this defines the Shapley-Shubik power index  $Piv(IAC, N)$ . It is well known that  $Piv(IC, N)$  is of order  $\frac{1}{\sqrt{N}}$  and  $Piv(IAC, N)$  is equal to  $\frac{1}{N}$ .

The main purpose of this paper is to continue the exploration of the implications of *correlation* on the asymptotic behavior of the power index. Precisely, we will consider a general family of models of random electorate  $\lambda$  and study the asymptotic behavior of  $Piv(\lambda, N)$  with respect to  $N$ . Our motivation to do so is to depart from the IAC model

---

<sup>1</sup>Good and Mayer (1975) refers to this as the *efficacy* of a vote.

which assumes that the correlation is the same for all pairs of voters in the population. It is likely that the intensity of the correlation between the votes of  $i$  and  $j$  will depend upon some characteristics of  $i$  and  $j$  suggesting that the correlation may vary from one pair to another. Most of the paper will however be based on a particular pattern of *heterogeneity*. Precisely, we will assume that the voters are partitioned into groups and that: correlation is positive and identical for any pair of voters belonging to the same group and null for any pair of voters belonging to two different groups. We will assume that within each group the correlation is defined as in the IAC model. This gives the IC and the IAC models as special cases: the IC model emerges when all the groups are singletons and the IAC model arises when there is a unique group which is then the entire population.

While particular, this model is general enough to cover many situations. We will offer a separate treatment of two polar cases. The first case is the case where there is a bound on the size of the groups; this bound does not depend upon the size of the population. This assumption is well suited to capture *local interactions* (within the family or the workplace for instance). The second case is the case where there is a fixed number of groups; this means that the size of the groups grows with the size of the population. This assumption is well suited to describe *large scale interactions* (special interest groups, geographical territories, electoral districts, countries if the population under scrutiny is multinational,...). After offering some general results, we proceed to the study of these two cases. The analysis of the two cases uses different techniques. When  $\lambda$  describes the local case, the use of some local versions of the Central Limit Theorem allows to estimate  $Piv(\lambda, N)$ . We show that it is of order  $\frac{1}{\sqrt{N}}$  and we calculate explicitly  $\lim_{N \rightarrow \infty} \sqrt{N} Piv(\lambda, N)$ . In contrast, when  $\lambda$  describes the global case, our estimation of  $Piv(\lambda, N)$  is based on different mathematical techniques. We address the problem quite differently using a combinatorial approach based on Ehrhart theory and algorithmic tools for computing the number of integer points in parametric polytopes. We show that  $Piv(\lambda, N)$  is of order  $\frac{1}{N}$  and we calculate explicitly  $\lim_{N \rightarrow \infty} N Piv(\lambda, N)$  in some specific cases.

#### *Related Literature*

The partition random model explored in this paper has been suggested by Straffin (1977) under the name *partial homogeneity*. He suggests this model as an alternative to the existing IC and IAC models but does not derive any general result. Instead, he proceeds to some numerical calculations of the probability of being pivotal in the Canadian constitutional amendment process. Straffin writes: “*In the Canadian constitution example, it might be that neither the independence assumption nor the homogeneity assumption describe the situation very well. British Columbia and Québec, for example, might reasonably be expected to*

*behave independently, while the four Atlantic provinces may have common interests and might reasonably be considered to judge proposed constitutional amendment by a common set of values. The most reasonable thing to do might be to partition the provinces into subsets whose members are homogenous among themselves, but behave independently of the members of other subsets".*

Chamberlain and Rothschild (1981) also consider the case of a partition into two groups and study the asymptotics of the probability of being pivotal under some general conditions: the random draws of the parameter  $p$  (denoting the probability that any individual votes for the first alternative) in each of the two groups do not necessarily result from a uniform distribution (a feature shared with Good and Mayer (1975)) and the draws are not necessarily independent among the two groups.

Our model of correlation among voters aims to contribute to the existing studies of the implications of correlation on power measurement. Knowing the exact magnitude of the probability of being pivotal is interesting for itself but this information is also essential for the design of the optimal weights of representatives, as argued convincingly by Barbera and Jackson (2006). They introduce a block model which is quite similar to the model of partitions which is considered here except for the fact that instead of IAC, they assume perfect correlation within each block/group. Precisely, they describe it as follows: *"Each country is made up of some number of blocks of agents, where agents within each block have perfectly correlated preferences and preferences across blocks are independent. The blocks within a country are of equal size. These assumptions reflect the fact that countries are often made up of some variety of constituencies, within which agents tend to have correlated preferences. For instance, the farmers in a country might have similar opinions on a wide variety of issues, as will union members, intellectuals, and so forth. The block model is a stylized but useful way to introduce correlation among voters' preferences"*. They proceed to a separate analysis of the "fixed-size-block model" and the "fixed-number-of-blocks model" which parallels exactly our distinction between small and large groups. The "block model" was in fact introduced by Penrose (1952) in chapter 7 of his pioneering monograph. His work is motivated by empirical considerations. He observes that if voters were voting independently of each other, then the mean value of the statistics  $\frac{D^2}{N}$  over an indefinite period of years (where  $D$  denotes the difference between the votes of the two sides) would equal unity. This prediction is violated in the case of the twenty-six American Presidential elections that he examined. The mean value is much larger than 1. He concludes from that that *"this marked excess over the theoretical value of unity may be interpreted as indicating that the voters did not vote as random units but were grouped into blocs which voted independently.*

The approximate size of each of a set of blocs taking the place of individuals is given by the actual mean value of  $\frac{D^2}{N}$  measured over a period of years".

## 2 The Model of a Random Electorate

A random electorate is a triple  $(\mathcal{N}, X, \lambda)$  where  $\mathcal{N}$  is a finite set of voters,  $X$  is a finite set of alternatives and  $\lambda$  is a probability distribution on  $\mathcal{P}^{\mathcal{N}}$  (the set of functions from  $\mathcal{N}$  to  $\mathcal{P}$ ) where  $\mathcal{P}$  is the set of linear orders over  $X$ . In the case where  $X$  consists of two alternatives say 0 and 1, the set  $\mathcal{P}$  contains two preferences (0 is preferred to 1, 1 is preferred to 0) which will be coded 0 and 1 and  $\mathcal{P}^{\mathcal{N}}$  will be identified with the Cartesian product  $\{0, 1\}^N$  where  $N$  denotes the cardinality of  $\mathcal{N}$ . The first popular random electorate model, called *Impartial Culture* (IC), is defined by  $\lambda(P) = \frac{1}{2^N}$  for all profiles of preferences  $P = (P_1, P_2, \dots, P_N)$  in  $\{0, 1\}^N$ . The IC model assumes that the preferences of the voters are independent Bernoulli random variables with a parameter  $p$  equal to  $\frac{1}{2}$  (i.e. the electorate is not biased towards a particular candidate). In contrast, the second popular random electorate model, called *Impartial Anonymous Culture* (IAC) is defined as follows. The parameter  $p$  is drawn in  $[0, 1]$  from the uniform distribution and, conditional on the draw of  $p$ , the preferences of the voters are independent Bernoulli preferences with parameter  $p$ . The probability of profile  $(P_1, P_2, \dots, P_N)$  is therefore  $\lambda(P) = \int_0^1 p^k (1-p)^{N-k} dp$  where  $k$  is the number of coordinates equal to 0 in  $P$ . Using the formula:

$$\int_0^1 p^t (1-p)^{N-t} dp = \frac{(t)!(N-t)!}{(N+1)!} \quad (1)$$

we obtain that  $\lambda(P) = \frac{1}{(N+1)\binom{N}{k}}$ . The terminology IAC results from the fact that in the IAC model, the events  $E_k \equiv \left\{ P \in \{0, 1\}^N : \#\{i \in \mathcal{N} : P_i = 0\} = k \right\}$  for  $k = 0, 1, \dots, N$  are equally likely. Since there are  $\binom{N}{k}$  such events, the probability that  $k$  voters vote 0 is equal to  $\frac{1}{N+1}$  for all  $k = 0, 1, \dots, N$ .

A social choice mechanism is a monotonic mapping  $\Psi$  from  $\{0, 1\}^N$  into  $[0, 1]$  where  $\Psi(P)$  denotes the probability of choosing candidate 0 when the profile of preferences is  $P$ <sup>2 3</sup>. In this paper, we will focus on the standard majority mechanism *Maj* defined as follows (with  $\#N^0(P)$  denoting the number of 0 in  $P$ ):

<sup>2</sup>In this binary setting, if  $\Psi(P) \in \{0, 1\}$ , a social choice mechanism is defined alternatively by a simple game (Taylor and Zwicker (1999)).

<sup>3</sup>We will not make any distinction between preferences and behavior. There is no room for strategic behavior here: if we interpret  $\Psi$  as a direct revelation game, then voting sincerely according to his/her preference is the unique (weakly) dominant strategy.

$$Maj(P) = \begin{cases} 0 & \text{if } \#N^0(P) < \frac{N}{2} \\ 1 & \text{if } \#N^0(P) > \frac{N}{2} \\ \frac{1}{2} & \text{if } \#N^0(P) = \frac{N}{2} \end{cases}$$

If  $N$  is odd, the third eventuality never arises and the mechanism is deterministic: the probability of choosing 0 is either 0 or 1. If  $N$  is even, the third alternative arises when the electorate is split into two groups of equal size and the tie is broken by using a fair lottery. The whole paper is about evaluating the probability of being pivotal. Recall that voter  $i \in \mathcal{N}$  is *pivotal* if her or his vote is susceptible to change the result of the election. We denote by  $E_k(i)$  the event defined by  $E_k(i) = \{P \in \{0, 1\}^N : \#N^0(P_{-i}) = k\}$  for  $k = 0, 1, \dots, N-1$ , and by  $Piv(\lambda, i, N)$  the probability that  $i$  is pivotal ( $P_{-i}$  denotes the reduced profile obtained by removing the vote of  $i$ ). It is easy to see that  $Piv(\lambda, i, N) = \lambda(E_{\frac{N-1}{2}}(i))$  when  $N$  is odd, and  $Piv(\lambda, i, N) = \frac{1}{2}\lambda(E_{\frac{N-2}{2}}(i)) + \frac{1}{2}\lambda(E_{\frac{N}{2}}(i))$  when  $N$  is even. The slight difference between the even and odd cases lies in the fact that in the odd case, the vote of  $i$  is always able to change the result, whereas in the even case, her or his vote creates an equality and the electoral outcome will change with probability  $\frac{1}{2}$ , depending on the lottery result.

As we consider the standard majority mechanism, if the probability measure is symmetric, then  $Piv(\lambda, i, N)$  does not depend on  $i$  and will be denoted shortly by  $Piv(\lambda, N)$ . The probability  $Piv(\lambda, N)$  has been calculated for the two popular models of random electorate which have just been defined. For the IC model,  $Piv(IC, N) = \left(\frac{N-1}{2}\right)\frac{1}{2^{N-1}}$  when  $N$  is odd and  $Piv(\lambda, N) = \left(\frac{N-1}{2}\right)\frac{1}{2^{N-1}}$  when  $N$  is even. For the IAC model,  $Piv(IAC, N) = \frac{1}{N}$  for both cases. Using Stirling's formula,  $N! \simeq \sqrt{2\pi N} \left(\frac{N}{e}\right)^N$ , we deduce that when  $N$  gets large  $Piv(IC, N)$  behaves like  $\sqrt{\frac{2}{\pi N}} \simeq \frac{0.79788}{\sqrt{N}}$ .

In this paper, we assume that the electorate  $\mathcal{N}$  is partitioned into  $K$  groups  $\mathcal{N}_1, \mathcal{N}_2, \dots, \mathcal{N}_K$  i.e.  $\cup_{1 \leq k \leq K} \mathcal{N}_k = \mathcal{N}$  and  $\mathcal{N}_k \cap \mathcal{N}_{k'} = \emptyset$  for all  $k, k'$  such that  $k \neq k'$ . We will denote by  $N_k$  the size of group  $k$ :  $\sum_{k=1}^K N_k = N$  and without loss of generality we assume that  $N_1 \geq N_2 \geq \dots \geq N_K$ . We consider the following random electorate model.

We assume that the preferences of any voter  $i$  from group  $\mathcal{N}_k$  is the realization of a Bernoulli random variable with parameter  $p_k$  and that conditional on  $p_k$ , the preferences of any two voters in that group are independent. We assume that the coordinates of the vector  $(p_1, p_2, \dots, p_K)$  are the realizations of  $K$  independent random variables with a uniform distribution on  $[0, 1]$ . As  $Piv(\lambda, i, N)$  is the same for all voters  $i$  belonging to the same group, this probability will be denoted by  $Piv_k(\lambda, N)$  for all members of group  $\mathcal{N}_k$ .

We start our calculations by reducing the problem of computing  $Piv_k(\lambda, N)$  to a well defined *combinatorial problem* which amounts to count the number of possible decompositions

of a given integer into  $K$  integers under some specific constraints.<sup>4</sup> We will denote by  $\Pi(M, R_1, \dots, R_k, \dots, R_K)$  the set of decompositions  $(x_1, \dots, x_K)$  of the integer  $M$  into  $K$  ordered integers ( $M = x_1 + \dots + x_K$ ) under the constraint that the  $k^{\text{th}}$  integer  $x_k$  does not exceed  $R_k$ , and by  $\pi(M, R_1, \dots, R_k, \dots, R_K)$  the cardinality of this set. To illustrate the use of such decompositions in our calculations, consider a society where the number  $N$  of voters is odd and multiple of 3:  $N = 3K$  with  $K$  odd. The society is therefore divided into  $K$  groups of size 3 each (three members). Take  $K = 21$  and suppose that  $i$  is a member of group 1. Voter  $i$  will be pivotal if in the rest of the society 31 voters vote 0 and 31 voters vote 1. How to enumerate the number of possible decompositions of 31 into 21 integers such that the first one cannot exceed 2 and the other twenty ones cannot exceed 3? One possibility is  $(2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 0, \dots, 0)$  and all subsequent permutations but we can also vary the choice of integers by taking for instance  $(1, 2, 2, 2, 2, \dots, 2, 2, 0, 0, 0, 0, 0)$  where “2” appears 15 times.<sup>5</sup>

**Lemma 1:** For all  $k \in \{1, \dots, K\}$ ,

$$1. \text{ If } N \text{ is odd, } Piv_k(\lambda, N) = \pi\left(\frac{N-1}{2}, N_1, \dots, N_{k-1}, N_k-1, N_{k+1}, \dots, N_K\right) \frac{1}{N_k} \left(\prod_{l \neq k} \frac{1}{N_l+1}\right) \quad (1.a)$$

$$2. \text{ If } N \text{ is even, } Piv_k(\lambda, N) = \pi\left(\frac{N-2}{2}, N_1, \dots, N_{k-1}, N_k-1, N_{k+1}, \dots, N_K\right) \frac{1}{N_k} \left(\prod_{l \neq k} \frac{1}{N_l+1}\right) \quad (1.b)$$

*Proof.* Consider the case where  $N$  is odd. We obtain:

$$Piv_k(\lambda, N) = \sum_{\Pi\left(\frac{N-1}{2}, N_1, \dots, N_{k-1}, N_k-1, N_{k+1}, \dots, N_K\right)} \binom{N_k-1}{x_k} \left( \int_0^1 p_l^{x_k} (1-p_k)^{N_k-x_k-1} dp_k \right) \\ \times \left[ \prod_{l \neq k} \binom{N_l}{x_l} \left( \int_0^1 p_l^{x_l} (1-p_l)^{N_l-x_l} dp_l \right) \right]$$

By using formula (1), we deduce:

$$Piv_k(\lambda, N) = \sum_{\Pi\left(\frac{N-1}{2}, N_1, \dots, N_{k-1}, N_k-1, N_{k+1}, \dots, N_K\right)} \frac{1}{N_k} \left( \prod_{l \neq k} \frac{1}{N_l+1} \right) = \\ \pi\left(\frac{N-1}{2}, N_1, \dots, N_{k-1}, N_k-1, N_{k+1}, \dots, N_K\right) \frac{1}{N_k} \left( \prod_{l \neq k} \frac{1}{N_l+1} \right)$$

---

<sup>4</sup>Note that there are at most  $K$  cells i.e.  $K$  non zero integers in the decomposition and there is an upper bound on the entries of each cell. Our problem is close but not totally equivalent to the problem of counting compositions of integers with restrictions on the summands and their number as presented in Flajolet and Sedgewick (2009).

<sup>5</sup>In that case, our first result (Proposition 1) will imply that the total number of decompositions behaves as  $c4^{\frac{N}{3}}\sqrt{N}$  where  $c$  is a universal constant.



The proof in the case where  $N$  is even proceeds along the same lines.  $\square$

Let us check quickly that the IC and IAC models correspond to two extreme special cases of this general framework. The IC value is attached to the case where  $K = N$  i.e. where the partition structure consists of  $N$  singletons:

$$Piv_k(IC, N) = Piv(IC, N) = \pi \left( \frac{N-1}{2}, 1, \dots, 1, N_k - 1, 1, \dots, 1 \right) \frac{1}{2^N} = \binom{N-1}{\frac{N-1}{2}} \frac{1}{2^{N-1}}$$

since  $\pi \left( \frac{N-1}{2}, 1, \dots, 1, 0, 1, \dots, 1 \right) = \binom{N-1}{\frac{N-1}{2}}$ . The IAC value is attached to the case where  $K = 1$  i.e. where the partition structure consists of a single set: the set  $\mathcal{N}$ :

$$Piv_k(IAC, N) = Piv(IAC, N) = \pi \left( \frac{N-1}{2}, N-1 \right) \frac{1}{N} = \frac{1}{N}$$

since  $\pi \left( \frac{N-1}{2}, N-1 \right) = 1$ .

An alternative approach to the counting problem is based on probability. Let  $X_{ik}$  denote the Bernoulli random variable describing the preference of voter  $i$  in group  $k$  and let  $S_k$  and  $\widehat{S}$  denote respectively the sums  $\sum_{j \in \mathcal{N}_k} X_{jk}$  and  $\sum_{k=1}^K \sum_{j \in \mathcal{N}_k} X_{jk} = \sum_{k=1}^K S_k$ . With these notations, we can express the pivot probabilities as follows:

$$\begin{aligned} Piv_k(\lambda, N) &= \lambda \left( \widehat{S}_{-i} = \frac{N-1}{2} \right) \text{ when } N \text{ is odd and} \\ Piv_k(\lambda, N) &= \frac{1}{2} \left( \lambda \left( \widehat{S}_{-i} = \frac{N-2}{2} \right) + \lambda \left( \widehat{S}_{-i} = \frac{N}{2} \right) \right) \text{ when } N \text{ is even} \end{aligned}$$

This probabilistic approach will be very useful when we will focus on the asymptotic behavior of  $Piv_k(\lambda, N)$  when  $N$  tends to infinity. Note that all the random variables  $X_{ik}$  are symmetric in the sense that  $\Pr(X_{ik} = 0) = \Pr(X_{ik} = 1) = \frac{1}{2}$  since  $\Pr(X_{ik} = 0) = \int_0^1 p dp = \frac{1}{2}$ . We have  $E[X_{ik}] = \frac{1}{2} \times 0 + \frac{1}{2} \times 1 = \frac{1}{2}$  and  $Var[X_{ik}] = E[X_{ik}^2] - E[X_{ik}]^2 = \frac{1}{4}$ . But two random variables  $X_{ik}$  and  $X_{jl}$  are independent iff  $k \neq l$ . If not, we have:

$$\Pr(X_{ik} = 0, X_{jk} = 0) = \int_0^1 p^2 dp = \frac{1}{3} > \frac{1}{4}$$

The two variables are positively correlated:  $Cov(X_{ik}, X_{jk}) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$ ; the coefficient of correlation  $\rho$  is then equal to  $\frac{1}{3}$ .

### 3 The case of Many Small Groups

In this section, we will focus on the case where there is an exogenous upper bound  $S$  on the size of the groups in the partition  $(\mathcal{N}_1, \mathcal{N}_2, \dots, \mathcal{N}_K)$ . This implies that as  $N$  gets large, then the number of groups increases.

To motivate the general result which will be presented hereafter, it is instructive to consider the case where  $S = 2$ . In any such partition structure, the groups are either singletons or pairs. We can think of this partition as describing a society where there are singles and couples but no other family types. Consider the case where  $N$  is even and all the groups are exactly of size 2. From (1.b), we deduce that:

$$Piv_k(\lambda, N) = Piv(\lambda, N) = \pi \left( \frac{N-2}{2}, 1, 2, \dots, 2, 2 \right) \frac{1}{2} \left( \frac{1}{3} \right)^{\frac{N-2}{2}}.$$

We can check that:<sup>6</sup>

$$\pi \left( \frac{N-2}{2}, 1, 2, \dots, 2, 2 \right) = \sum_{k=0}^{\lfloor \frac{N-2}{4} \rfloor} \frac{\left( \frac{N-2}{2} \right)!}{(k!)^2 \left( \frac{N-2}{2} - 2k \right)!} \binom{\frac{N}{2} - k}{k+1}.$$

Indeed, counting how many decompositions of  $\frac{N-2}{2}$  into  $\frac{N}{2}$  integers chosen in  $\{0, 1, 2\}$  amounts first to choose how many pairs  $k$  we choose among  $\frac{N-2}{2}$ . The number of possibilities is  $\frac{\left( \frac{N-2}{2} \right)!}{(k!) \left( \left( \frac{N-2}{2} - k \right) \right)!}$ . This value of  $k$  cannot exceed  $\lfloor \frac{N-2}{4} \rfloor$ .

To reach the integer  $\frac{N-2}{2}$ , we need  $\frac{N-2}{2} - 2k$  singletons which can be chosen among  $\frac{N}{2} - k$ . The number of possibilities is  $\frac{\left( \frac{N}{2} - k \right)!}{\left( \frac{N-2}{2} - 2k \right)! (k+1)!} = \frac{\left( \frac{N}{2} - 1 - k \right)! \frac{N-k}{k+1}}{\left( \frac{N-2}{2} - 2k \right)! (k+1)!}$ .

After collecting the terms, we obtain the expression reported above.

Calculating the above sum is not an immediate combinatorial exercise<sup>7</sup> and we will mostly focus on the asymptotic behavior of  $Piv(\lambda, N)$ .

We conjecture that:

$$\lim_{N \rightarrow \infty} \Phi(N) \equiv \sqrt{N} \left( \sum_{k=0}^{\lfloor \frac{N-2}{4} \rfloor} \frac{\left( \frac{N-2}{2} \right)!}{(k!)^2 \left( \left( \frac{N-2}{2} - 2k \right) \right)!} \times \frac{N-2k}{2k+2} \right) \times \frac{1}{2} \times \left( \frac{1}{3} \right)^{\frac{N-2}{2}} \text{ exists.}$$

The following Table contains some numerical values of  $\Phi(N)$  which supports this conjecture:

<sup>6</sup> $\lfloor x \rfloor$  denotes the integer part of  $x$ .

<sup>7</sup>We were not able to derive a closed form value of this sum through the use of combinatorial identities.

$N$	102	202	1002	5002	100002
$\Phi(N)$	0.69015	0.69056	0.6909	0.69097	0.69098

Table 1: Values of  $\Phi(N)$

Interestingly, the function  $\Phi$  seems to behave asymptotically as the function  $\Upsilon$  defined as follows:

$$\Upsilon(N) \equiv \sqrt{N} \left( \sum_{k=0}^{\lfloor \frac{N-2}{4} \rfloor} \frac{\binom{N-2}{2k}!}{(k!)^2 ((\frac{N-2}{2} - 2k)!)} \right) \times \left( \frac{1}{3} \right)^{\frac{N-2}{2}}$$

The following Table contains some numerical values of  $\Upsilon(N)$  which supports this guess:

$N$	102	202	1002	10002	100002
$\Upsilon(N)$	0.69525	0.69314	0.69143	0.69103	0.69099

Table 1 bis: Values of  $\Upsilon(N)$

We now prove a generalized version of the conjecture. To proceed, we use a probabilistic approach. We assume that all the groups have a size smaller than  $S$  and we will be interested in societies where the set of voters is partitioned into groups of size  $s$  where  $s$  runs from 1 to  $S$ . We will consider societies where  $N$  gets indefinitely large but such that the proportion of the population in each type of group (described by its size) remains invariant in the population growth process. We will denote by  $\gamma^s$  the proportion of voters in a group of size  $s$ . We assume that  $\gamma^s = \frac{sK^s}{N}$  where  $K^s$  is an integer for all  $s = 1, \dots, S$  and  $N = \sum_{s=1}^S sK^s$ . The initial society contains  $K^s$  groups of size  $s$ . For any integer  $R$ , its  $R^{\text{th}}$  replica has  $N$  voters where  $N$  is defined as follows:

$$N = N(R) = R \sum_{s=1}^S K^s s$$

In this context,  $Piv_k(\lambda, N)$  is the same for all groups  $\mathcal{N}_k$  of size  $s$  and will be denoted by  $Piv_{(s)}(\lambda_R, N)$ , the probability for a voter belonging to a group of size  $s$  to be pivotal. The following proposition gives the asymptotic behavior of  $Piv_{(s)}(\lambda_R, N)$  when  $N$  gets large (equivalently where  $R$  gets large).

**Proposition 1:** Let  $\lambda_R$  be the random electorate defined above. For all  $s = 1, 2, \dots, S$ ,

$$\lim_{R \rightarrow \infty} \sqrt{N} Piv_{(s)}(\lambda_R, N) = \frac{1}{\left( \sqrt{\frac{1}{6} + \frac{\gamma^1}{12} + \frac{1}{12} \sum_{l=2}^S \gamma^l l} \right) \sqrt{2\pi}}.$$

*Proof.* For all  $R$  and all  $i = 1, 2, \dots, N(R)$ , we arrange the random variables  $X_i^R$  describing the individual votes in the  $R^{\text{th}}$  replica in a *triangular array*<sup>8</sup> defined as follows: the first  $RK^1$  variables describe the vote of voters in groups of size 1, the next  $2RK^2$  variables describe the votes of voters in groups of size 2 and so on.

We obtain

$$\sigma^2(R) \equiv \text{Var}\left(\sum_{i=1}^N X_i^R\right) = \sum_{i=1}^N \text{Var}(X_i^R) + \sum_{s=1}^S RK^s s(s-1) \text{Cov}(X_i^R, X_j^R)$$

where  $\text{Cov}(X_i^R, X_j^R)$  denotes the covariance between  $X_i^R$  and  $X_j^R$  when  $i$  and  $j$  belong to the same group. We have shown before that:

$$\text{Var}(X_i^R) = \frac{1}{4} \text{ for all } i = 1, 2, \dots, N$$

$$\text{Cov}(X_i^R, X_j^R) = \frac{1}{12} \text{ for all } i, j = 1, \dots, N \text{ if } i \text{ and } j \text{ belong to the same group}$$

We obtain:

$$\sigma(R) = \sqrt{N} \left( \sqrt{\frac{1}{6} + \frac{\gamma^1}{12} + \frac{1}{12} \sum_{s=2}^S \gamma^s s} \right)$$

A random variable  $X_i^R$  is of type  $s$  if  $R \sum_{l=1}^{s-1} lK^l < i \leq R \sum_{l=1}^s lK^l$ . We pack the  $sRK^s$  random variables of type  $s$  into  $RK^s$  random variables  $(Z_{ks}^R)_{1 \leq k \leq RK^s}$  where  $Z_{ks}^R$  is defined as follows:

$$Z_{ks}^R = r \text{ iff } \sum_{i=(k-1)s+1}^{ks} X_{is}^R = r$$

This defines a new triangular array  $(Z_{ks}^R)_{1 \leq s \leq S, 1 \leq k \leq RK^s}$  (indexed by  $R$ ) where the random variables  $Z_{ks}^R$  are independent. Hereafter, we will refer to  $Z_{ks}^R$  as a random variable of type  $s$ . We note that all random variables are integer valued: the support of a random variable of type  $s$  is  $\{0, 1, \dots, s\}$ . Let  $1 \leq i \leq N(R)$  be a member of a group of type  $s$  and for each value of the row index  $R$ , consider the random variable  $S_i^R$  defined as follows:

$$S_i^R = \sum_{j=1, j \neq i}^{N(R)} X_j^R = \sum_{l=1, l \neq s}^S \sum_{k=1}^{K^l} Z_{kl}^R + \sum_{k=1}^{K^s-1} Z_{ks}^R + W_i^R$$

---

<sup>8</sup> A triangular array is a collection of  $(y_1^k, y_2^k, \dots, y_{n(k)}^k)_{k \geq 1}$  of random variables on a probability space.

where  $W_i^R \equiv \sum_{j=2}^s 1_{X_{j_s}^R}$ . The probability that  $i$  of type  $s$  is pivotal,  $Piv_{(s)}(\lambda_R, N)$  is equal to the probability of the event  $\{S_i^R = \frac{N-1}{2}\}$  if  $N$  is odd and to half the probability of the event  $\{S_i^R = \frac{N-2}{2}\} \cup \{S_i^R = \frac{N}{2}\}$  if  $N$  is even.

We note that the span of the random variables  $Z_{kl}^R$  for  $1 \leq l \leq S$  and  $1 \leq k \leq K^l$  and  $W_i^R$  is equal to 1. Further, the distribution functions of these random variables belong to a finite set of cardinality at most  $S$ , are not degenerate and occur infinitely often (except possibly  $W_i^R$ ) in the sequence  $\left( (Z_{kl}^R)_{1 \leq l \leq S, 1 \leq k \leq K^l} \cup \{W_i^R\} \right)_{R \geq 1}$ . Let  $\epsilon > 0$ .

If  $N$  is odd, since  $E[S_i^R] = \frac{N-1}{2}$ , we deduce from the local central limit theorem on lattice distributions listed as theorem 2 in Petrov (1975)<sup>9</sup> that if  $R$  is large enough:

$$\left| \sigma(S_i^R) Piv_{(s)}(\lambda_R, N) - \frac{1}{\sqrt{2\pi}} \right| \leq \epsilon$$

Similarly, if  $N$  is even, since  $E[S_i^R] = \frac{N-1}{2}$ , we deduce from Petrov's theorem that if  $R$  is large enough:

$$\left| \sigma(S_i^R) \Pr\left(S_i^R = \frac{N-2}{2}\right) - \frac{e^{-\frac{1}{8\sigma^2(S_i^R)}}}{\sqrt{2\pi}} \right| \leq \epsilon$$

$$\left| \sigma(S_i^R) \Pr\left(S_i^R = \frac{N}{2}\right) - \frac{e^{-\frac{1}{8\sigma^2(S_i^R)}}}{\sqrt{2\pi}} \right| \leq \epsilon$$

Since  $\frac{e^{-\frac{1}{8\sigma^2(S_i^R)}}}{\sqrt{2\pi}}$  tends to  $\frac{1}{\sqrt{2\pi}}$  and  $\frac{\sigma(S_i^R)}{\sigma(R)}$  tends to 1 when  $R$  tends to  $+\infty$ , we deduce that if  $R$  is large enough:

$$\left| \sigma(R) Piv_{(s)}(\lambda_R, N) - \frac{1}{\sqrt{2\pi}} \right| \leq \epsilon.$$

□

The random variable  $S_i^R = S^R - X_i^R$  introduced in the proof of Proposition 1 counts the number of votes in favor of 1 in the population without individual  $i$ . Proposition 1 provides information on the asymptotic behavior of the probability of the event  $\{S_i^R = \frac{N-1}{2}\}$ . To illustrate Proposition 1, consider the case of an electorate, denoted  $\lambda_R^s$ , where all the groups have the same size  $s$ . In such case, we deduce from our result that:

<sup>9</sup>Chapter 7, Section 1, p 189. Petrov's theorem in what follows.

$$Piv_{(s)}(\lambda_R, N) = Piv(\lambda_R^s, N) \simeq \frac{1}{\sqrt{N}} \times 2\sqrt{\frac{3}{2\pi(2+s)}}$$

The following Table lists a sample of values of the probability of being pivotal for a sample of values of  $s$ .

$s$	1	2	3	4	5	...	10
$\sqrt{N}Piv(\lambda_R^s, N)$	0.798	0.691	0.618	0.564	0.522	...	0.399

Table 2:  $\sqrt{N} \times$  Probability of being pivotal as a function of  $s$

We can also handle mixed situations i.e. random electorates  $\lambda$  where the sizes of the groups differ across voters. For instance, when the random electorate  $\lambda$  is such that  $\gamma^1 = 0.2$ ,  $\gamma^2 = 0.3$ ,  $\gamma^3 = 0.4$  and  $\gamma^4 = 0.1$ , we obtain :  $Piv(\lambda_R) \simeq 0.65885$ . We could interpret these groups as family groups: singles, couples without children voting, couples with one children voting, and so on.

**Remark 1.** The proof strategy of Proposition 1 based on a specific version of the local central limit theorem has exploited the fact that the individuals could be partitioned in a regular way and that in each group the probability draw of the votes in the group was not changing with the size of the population. From our construction, in each group of size  $s$ , the random number of votes for 1 is described by a multinomial probability law independent of  $N$  with values in the set  $\{0, 1, \dots, s\}$ . So partitioning per se is not enough to permit a direct use of that version of the local central limit theorem; we need invariance of the law with respect to  $N$ .<sup>10</sup>

**Remark 2**<sup>11</sup>. We could alternatively look at the probability of being pivotal of a group of size  $\lfloor \epsilon\sqrt{N} \rfloor$  where  $\epsilon > 0$  is fixed instead of a group of size 1 as done until now. Such a group, acting as a block, is pivotal iff :

$$\frac{N}{2} - \frac{\epsilon\sqrt{N}}{2} \leq S_N^R \leq \frac{N}{2} + \frac{\epsilon\sqrt{N}}{2}$$

where:

$$S_N^R = \sum_{i=1}^{N(R)} X_i^R$$

<sup>10</sup>We don't mean that it is a necessary condition per se. It is a necessary condition to appeal at the theorem that we have used. We conjecture that additional results could be deduced from other local central limit theorems (like for instance Mc Donald (1979)).

<sup>11</sup>We thank S. Brams for having raised the question answered in that remark.

and  $(X_i^R)_{1 \leq i \leq N(R)}$  is an arbitrary triangular array of Bernoulli random variables of parameter  $\frac{1}{2}$ . Let us assume that this triangular array is  $m(R)$ -dependent<sup>12</sup> and such that for some  $\eta > 0$  and some constant  $K$ :

$$\text{Var} (X_{i+1}^R + \dots + X_j^R) \leq (j - i) K \text{ for all } i, j \text{ and } R,$$

$$\text{Lim}_{R \rightarrow \infty} \frac{\text{Var} (X_1^R + \dots + X_{N(R)}^R)}{N(R)} \text{ exists and is nonzero,}$$

$$\text{Lim}_{R \rightarrow \infty} \frac{m(R)^{2+\eta}}{N(R)} = 0.$$

Since the Bernoulli variables have moments of any order, we deduce from Berk's theorem that  $\frac{X_1^R + \dots + X_{N(R)}^R - \frac{N(R)}{2}}{\sqrt{N(R)}}$  is asymptotically normal with mean 0 and variance  $v$  where  $v \equiv \text{Lim}_{R \rightarrow \infty} \frac{\text{Var}(X_1^R + \dots + X_{N(R)}^R)}{N(R)}$  (see Appendix 1). We deduce that the probability of a group of relative size  $\epsilon\sqrt{N}$  to be pivotal, denoted  $\text{Piv}(\epsilon, N)$ , is approximately<sup>13</sup> equal to:

$$\begin{aligned} \text{Prob} \left\{ \frac{N}{2} - \frac{\epsilon\sqrt{N}}{2} \leq S_N^R \leq \frac{N}{2} + \frac{\epsilon\sqrt{N}}{2} \right\} &\simeq \text{Prob} \left\{ -\frac{\epsilon}{2} \leq N(0, \sqrt{v}) \leq \frac{\epsilon}{2} \right\} \\ &\simeq \epsilon \sqrt{\frac{1}{2\pi v}} \end{aligned}$$

This weak version of the pivotality result holds in a much larger class of electorates. The notion of  $m$ -dependency matches different possibilities. First, we could continue to consider partitions into groups whose size can even increase slowly with  $N$ . What is essential, as reflected by the other two conditions of Berk's theorem, is to bound in an appropriate way the variance of any pack of random variables and to have the variance of the electorate to behave asymptotically as the size of the electorate. For the sake of illustration of such construction, consider the case where the  $N(R)$  voters are partitioned into consecutive blocks where each block has a size  $m = m(R) = \lfloor N^{\frac{1}{4}} \rfloor$ . We assume that  $m(R)$  is even and that within each block the Bernoulli random variables describing the votes are correlated as follows. Let  $Z = (Z^1, Z^2, \dots, Z^{m-1}, Z^m)$  be a  $m$ -dimensional random vector such that (i) the coordinates are independent random variables, (ii) the first  $m - 1$  coordinates  $Z^j$   $j = 1, \dots, m - 1$  are

<sup>12</sup>For a definition of the notion of  $m$ -dependency, see Appendix 1, footnote 20.

<sup>13</sup>Note that here, in contrast to the case where we consider a single voter or even a finite set of voters whose size does not depend upon  $N$ , the limit probability of the group of being pivotal does not change if we change the quota  $\lfloor \epsilon\sqrt{N} \rfloor$  into the quota  $\lfloor \epsilon\sqrt{N} \rfloor + C$  where  $C$  is a constant.

Bernoulli random variables with parameter  $\frac{1}{2}$  and (iii) the random variable  $Z^m$  takes the values 0 and 1 with the probability  $\frac{p}{2}$  and the value  $\frac{1}{2}$  with the probability  $1 - p$  where  $p \in [0, 1]$ . Given  $Z$ , the vector  $X$  is constructed as follows:

$$X_i^R = Z^i \text{ for all } i = 1, \dots, m-1$$

$$X_m^R = \begin{cases} 1 & \text{if either } Z^m = 1 \text{ or } Z^m = \frac{1}{2} \text{ and } \text{Maj}(Z^1, Z^2, \dots, Z^{m-1}) = 1 \\ 0 & \text{if either } Z^m = 0 \text{ or } Z^m = \frac{1}{2} \text{ and } \text{Maj}(Z^1, Z^2, \dots, Z^{m-1}) = 0 \end{cases}$$

In words, the  $m^{\text{th}}$  player of each block votes independently of the other voters with some probability or follows the majority opinion of the other voters with the complement probability. It is immediate to check that  $X_m^R$  is a Bernoulli random variable with parameter  $\frac{1}{2}$ . Let us evaluate how his vote is correlated with the vote of any other voter  $i$ . From above we deduce that:

$$\begin{aligned} \text{Cov}(X_m^R, X_i^R) &= \text{Prob}(X_m^R = 1 \text{ and } X_i^R = 1) - \frac{1}{4} \\ &= \text{Prob}(Z^m = 1 \text{ and } X_i^R = 1) + \\ &\quad \text{Prob}(Z^m = \frac{1}{2}, \text{Maj}(Z^1, Z^2, \dots, Z^{m-1}) = 1 \text{ and } X_i^R = 1) - \frac{1}{4} \\ &= \frac{p}{4} + \frac{1-p}{2} \text{Prob}(\text{Maj}(Z^1, Z^2, \dots, Z^{m-1}) = 1 \mid X_i^R = 1) - \frac{1}{4} \end{aligned}$$

So we are left with the evaluation of the conditional probability  $\text{Prob}(\text{Maj}(Z^1, Z^2, \dots, Z^{m-1}) = 1 \mid Z_i^R = 1)$  which is equal to  $\sum_{k=\frac{m-2}{2}}^{m-2} \binom{m-2}{k} \frac{1}{2^{m-2}} = \frac{1}{2} + \frac{1}{2^{m-1}} \binom{m-2}{\frac{m-2}{2}}$ . If  $m$  is large, we deduce from Stirling's formula that:

$$\frac{1}{2^{m-1}} \binom{m-2}{\frac{m-2}{2}} \simeq \frac{1}{2} \sqrt{\frac{2}{\pi m}}$$

Collecting the terms together, we derive  $\text{Prob}(\text{Maj}(Z^1, Z^2, \dots, Z^{m-1}) = 1 \mid Z_i^R = 1) \simeq \frac{1}{2} \left(1 + \sqrt{\frac{2}{\pi m}}\right)$  and therefore:

$$\text{Cov}(X_m^R, X_i^R) \simeq \frac{p}{4} + \frac{1-p}{4} \left(1 + \sqrt{\frac{2}{\pi m}}\right) - \frac{1}{4} \simeq \frac{1-p}{\sqrt{\pi m}} \text{ for all } i = 1, \dots, m-1$$

The covariance matrix of each block is almost diagonal. We obtain that in each block  $\text{Var}(X_1^R + \dots + X_m^R) \simeq \frac{m}{4} + \frac{2(1-p)\sqrt{m}}{\sqrt{\pi}}$ . Conditions (ii) and (iii) of Berk's theorem follow easily. Here  $v = \frac{1}{4}$ . Property (iv) also holds true since for any  $\eta < 2$ ,  $\frac{m(R)^{2+\eta}}{N(R)} = N^{-\frac{1}{2} + \frac{\eta}{4}}$  tends to 0 when  $N$  tends to infinity. In that example, correlation (which concerns only a unique



voter) collapses as the number of voter grows. We can provide examples where it is not the case

Second, we could even abandon completely the idea of partitioning of voters into blocs and move instead towards local interaction. Indeed,  $m$ -dependence (even with  $m$  independent of  $N$ ) does not force partition. For instance, let  $Y_1, Y_2, \dots$  be a sequence of independent Bernoulli variables with parameter  $\frac{1}{2}$  and let  $X_i$  be defined as follows:

$$X_1 = Y_1 \text{ and } X_i = \begin{cases} Y_i & \text{with probability } p \\ Y_{i-1} & \text{with probability } 1 - p \end{cases}$$

where  $p \in [0, 1]$ . It is easy to check that the  $X_i$  are Bernoulli random variables with parameter  $\frac{1}{2}$ . Note also that for all  $i \geq 3$ <sup>14</sup>,  $Cov(X_i, X_{i-1}) = \frac{p(1-p)}{4}$  and  $Cov(X_i, X_{i+m}) = 0$  whenever  $m > 1$ . Indeed,  $X_i X_{i-1} = 1$  is the union of the following four disjoint events: either:

$$\{X_i = Y_i = 1 \text{ and } X_{i-1} = Y_{i-1} = 1\}, \text{ or } \{X_i = Y_i = 1 \text{ and } X_{i-1} = Y_{i-2} = 1\} \text{ or}$$

$$\{X_i = Y_{i-1} = 1 \text{ and } X_{i-1} = Y_{i-2} = 1\} \text{ or } \{X_i = Y_{i-1} = 1 \text{ and } X_{i-1} = Y_{i-1}\}$$

Using independence, we deduce that the probabilities of the first three events are respectively  $\frac{p^2}{4}$ ,  $\frac{p(1-p)}{4}$  and  $\frac{(1-p)^2}{4}$  while the probability of the fourth one is equal to  $\frac{p(1-p)}{2}$ . Therefore the probability of the event  $\{X_i X_{i-1} = 1\}$  is equal to  $\frac{p^2 + 2p(1-p) + (1-p)^2}{4} + \frac{p(1-p)}{2} = \frac{1}{4} + \frac{p(1-p)}{4}$ .

In this construction, the sequence  $Y$  is a sequence of independent signals : voter  $i$ ' s preference follows his own signal with some probability  $p$  and follows the previous signal with probability  $1 - p$ . This makes every voter correlated with his two adjacent neighbors. There is no way in which we can partition the voters into blocks of size 2.

## 4 The Case of Few Large Groups

In this section, we consider the polar case of a society divided into a finite (possibly large) number of groups. This means that as  $N$  gets larger and larger, the number of voters in each group gets larger and larger. This peculiarity prevents us from applying the same probabilistic approach as in the preceding section To circumvent this difficulty, we will tackle the problem from a different combinatorial angle. For fixed values of  $K$ , the general problem of computing the number  $\pi(M, R_1, \dots, R_k, \dots, R_K)$  can be phrased as counting the exact

---

<sup>14</sup>When  $i = 2$ ,  $Cov(X_1, X_2) = \frac{1-p}{4}$ .

number of integer solutions of a system of linear inequalities with integer coefficients, where the variables are  $x_k$  ( $k = 1, \dots, K$ ) and the parameters are  $M$  and  $R_k$  ( $k = 1, \dots, K$ ), which is equivalent to count the number of integer points inside a parametric convex polytop. There is a well established mathematical approach for performing such a calculation, based on Ehrhart's theory (Ehrhart, 1962, 1967, 1977) and efficient counting algorithms<sup>15</sup>. We refer to Lepelley et al. (2008) and Wilson and Pritchard (2007) for more details on the use of these tools in probability calculations under IAC hypothesis in voting theory. Most of the results presented in the following subsections have been obtained by applying (parameterized) Barvinok's algorithm (Barvinok, 1994; Barvinok and Pommersheim, 1999)<sup>16</sup>.

## 4.1 A Preliminary Result

It can be noticed that, when the number  $N_1$  of voters in the largest group represents more than 50% of the total number of voters, then the probability of casting a decisive vote only depends, in each group, on the value of  $N_1$ . More precisely, we have the following general result (Recall that  $\lfloor x \rfloor$  denotes the integer part of  $x$ ).

**Proposition 2:** If  $N_1 \geq \lfloor \frac{N}{2} \rfloor + 1$ , then  $Piv_1(\lambda, N) = \frac{1}{N_1}$  and  $Piv_k(\lambda, N) = \frac{1}{N_1+1}$  for  $k = 2, 3, \dots, K$ .

*Proof.* Let  $x_k$  be the value of the  $k^{th}$  term in the decomposition of  $\lfloor \frac{N-1}{2} \rfloor$ :  $x_1 + \dots + x_k + \dots + x_K = \lfloor \frac{N-1}{2} \rfloor$ . If  $N_1 \geq \lfloor \frac{N}{2} \rfloor + 1$ , then  $N_2 + N_3 + \dots + N_k + \dots + N_K \leq \lfloor \frac{N-1}{2} \rfloor$ . Consequently, for  $k = 2, 3, \dots, K$ ,  $x_k$  can take any integer value between 0 and  $N_k$  (including 0 and  $N_k$ ) and when  $x_2, x_3, \dots, x_K$  are set, the value of  $x_1$  is given in a unique way by  $x_1 = \lfloor \frac{N-1}{2} \rfloor - x_2 - x_3 - \dots - x_K$ . The number of possible decompositions is then given by

$$\pi\left(\left\lfloor \frac{N-1}{2} \right\rfloor, N_1, N_2, \dots, N_k, \dots, N_K\right) = (N_2 + 1)(N_3 + 1)\dots(N_K + 1)$$

and the result follows from relations (1.a) and (1.b).  $\square$

## 4.2 The Case of Two Groups

Let us consider the case where  $K = 2$  i.e. the situation where the voters are partitioned into two groups. This setting has been examined by various authors in the literature including

<sup>15</sup>For a general background on Ehrhart theory and on the general problem of counting integer points in polytopes, see for example Beck and Robins (2007).

<sup>16</sup>For a rigorous description of this algorithm and for implementation details, see Verdoolage et al. (2004, 2005).

Beck (1975), Kleiner (1980), Chamberlain and Rothschild (1981) and Le Breton and Lepelley (2014).

In such a case, if  $N$  is odd, then  $N_1 > N_2$  as the two integers don't have the same parity. It is easily seen that:

$$\pi\left(\frac{N-1}{2}, N_1-1, N_2\right) = N_2 + 1 \text{ and } \pi\left(\frac{N-1}{2}, N_1, N_2-1\right) = N_2$$

and therefore:

$$Piv_1(\lambda, N) = \frac{1}{N_1} \text{ and } Piv_2(\lambda, N) = \frac{1}{N_1+1}$$

in accordance with Proposition 2.

### 4.3 Three groups of voters

In this section, we consider the case where the population is divided into three groups of voters i.e.  $K = 3$ :  $N_1 \geq N_2 \geq N_3$  and  $N_1 + N_2 + N_3 = \widehat{N} + 1$ , with  $\widehat{N}$  even.

**Proposition 3:** 1. If  $N_1 \geq \frac{\widehat{N}}{2} + 1$  (in accordance with our preliminary result),

$$Piv_1(\lambda, N) = \frac{1}{N_1} \text{ and } Piv_2(\lambda, N) = Piv_3(\lambda, N) = \frac{1}{N_1+1}$$

2. If  $N_1 \leq \frac{\widehat{N}}{2}$ ,

$$\begin{aligned} Piv_1(\lambda, N) &= \frac{4N_1^2 + 4N_1(N_2 - \widehat{N} - 2) + 4N_2^2 - 4N_2(\widehat{N} + 1) + \widehat{N}(\widehat{N} + 2)}{4N_1(N_2 + 1)(N_1 + N_2 - \widehat{N} - 2)}, \\ Piv_2(\lambda, N) &= \frac{4N_1^2 + 4N_1(N_2 - \widehat{N} - 1) + 4N_2^2 - 4N_2(\widehat{N} + 2) + \widehat{N}(\widehat{N} + 2)}{4(N_1 + 1)N_2(N_1 + N_2 - \widehat{N} - 2)}, \\ Piv_3(\lambda, N) &= \frac{4N_1^2 + 4N_1(N_2 - \widehat{N}) + 4N_2^2 - 4N_2\widehat{N} + \widehat{N}^2 - 2\widehat{N} - 4}{4(N_1 + 1)(N_2 + 1)(N_1 + N_2 - \widehat{N} - 1)}. \end{aligned}$$

*Proof.* The value of  $\pi(\frac{\widehat{N}}{2}, N_1 - 1, N_2, N_3)$  is given by the number of integer solutions of the following set of (in)equalities, where  $x_k$  can be interpreted as the number of voters voting for alternative 0 in group  $k$ ,  $k = 1, 2, 3$ :

$$\begin{aligned} 0 &\leq x_1 \leq N_1 - 1 \\ 0 &\leq x_2 \leq N_2 \\ 0 &\leq x_3 \leq N_3 \\ x_1 + x_2 + x_3 &= \frac{\widehat{N}}{2} \end{aligned}$$

Given the last equality,  $N_3 = N - N_1 - N_2$  and the above set of inequalities reduces to:

$$\begin{aligned} 0 &\leq x_1 \leq N_1 - 1 \\ 0 &\leq x_2 \leq N_2 \\ 0 &\leq x_3 \leq N - N_1 - N_2 \\ x_1 + x_2 + x_3 &= \frac{\widehat{N}}{2} \end{aligned}$$

where the parameters satisfy:

$$\begin{aligned} N_1 &\geq N_2 \\ 2N_2 + N_1 - N - 1 &\geq 0 \text{ and} \\ N_1 + N_2 &\leq N + 1 \end{aligned}$$

A representation for the number of integer solutions of this set of inequalities with three variables and three parameters ( $N_1$ ,  $N_2$  and  $N$ ) can be derived by using the multiparameter version of the Barvinok's algorithm (see Lepelley et al. (2008)). We obtain:

$$\pi\left(\frac{\widehat{N}}{2}, N_1 - 1, N_2, N_3\right) = (\widehat{N} - N_1 - N_2 + 2)(N_2 + 1) = (N_3 + 1)(N_2 + 1)$$

if  $N_1 \geq \frac{\widehat{N}}{2} + 1$  and

$$\pi\left(\frac{\widehat{N}}{2}, N_1 - 1, N_2, N_3\right) = (-\widehat{N}^2 + 2\widehat{N}(2N_1 + 2N_2 - 1) - 4(N_1^2 + N_1(N_2 - 2) + N_2(N_2 - 1)))/4$$

if  $N_1 \leq \frac{\widehat{N}}{2}$ .

Representations for  $\pi(\frac{\widehat{N}}{2}, N_1, N_2 - 1, N_3)$  and  $\pi(\frac{\widehat{N}}{2}, N_1, N_2, N_3 - 1)$  can be derived in a similar way to obtain:

$$\pi\left(\frac{\widehat{N}}{2}, N_1, N_2 - 1, N_3\right) = (\widehat{N} - N_1 - N_2 + 2)N_2 = (N_3 + 1)N_2$$

if  $N_1 \geq \frac{\widehat{N}}{2} + 1$  and

$$\pi\left(\frac{\widehat{N}}{2}, N_1, N_2 - 1, N_3\right) = (-\widehat{N}^2 + 2\widehat{N}(2N_1 + 2N_2 - 1) - 4(N_1^2 + N_1(N_2 - 1) + N_2(N_2 - 1)))/4$$

if  $N_1 \leq \frac{\widehat{N}}{2}$  ;

$$\pi\left(\frac{\widehat{N}}{2}, N_1, N_2, N_3 - 1\right) = (\widehat{N} - N_1 - N_2 + 1)(N_2 + 1) = N_3(N_2 + 1)$$

if  $N_1 \geq \frac{\widehat{N}}{2} + 1$  and

$$\pi\left(\frac{\widehat{N}}{2}, N_1, N_2, N_3 - 1\right) = (-\widehat{N}^2 + 2\widehat{N}(2N_1 + 2N_2 + 1) - 4(N_1^2 + N_1N_2 + N_2^2 - 1))/4$$

if  $N_1 \leq \frac{\widehat{N}}{2}$ .

Observe that we recover the results we have mentioned for two groups by taking  $N_3 = 0$ . From the above results, we can now derive the probability of casting a decisive vote for a voter belonging to each of the three groups. We obtain :

$$\begin{aligned} Piv_1(\lambda, N) &= \frac{(N_3 + 1)(N_2 + 1)}{N_1(N_2 + 1)(N_3 + 1)} = \frac{1}{N_1} \\ Piv_2(\lambda, N) &= \frac{(N_3 + 1)N_2}{(N_1 + 1)N_2(N_3 + 1)} = \frac{1}{N_1 + 1} \\ Piv_3(\lambda, N) &= \frac{N_3(N_2 + 1)}{(N_1 + 1)(N_2 + 1)N_3} = \frac{1}{N_1 + 1} \end{aligned}$$

if  $N_1 \geq \frac{\widehat{N}}{2} + 1$  and

$$\begin{aligned} Piv_1(\lambda, N) &= \frac{4N_1^2 + 4N_1(N_2 - \widehat{N} - 2) + 4N_2^2 - 4N_2(\widehat{N} + 1) + \widehat{N}(\widehat{N} + 2)}{4N_1(N_2 + 1)(N_1 + N_2 - \widehat{N} - 2)} \\ Piv_2(\lambda, N) &= \frac{4N_1^2 + 4N_1(N_2 - \widehat{N} - 1) + 4N_2^2 - 4N_2(\widehat{N} + 2) + \widehat{N}(\widehat{N} + 2)}{4(N_1 + 1)N_2(N_1 + N_2 - \widehat{N} - 2)} \\ Piv_3(\lambda, N) &= \frac{4N_1^2 + 4N_1(N_2 - \widehat{N}) + 4N_2^2 - 4N_2\widehat{N} + \widehat{N}^2 - 2\widehat{N} - 4}{4(N_1 + 1)(N_2 + 1)(N_1 + N_2 - \widehat{N} - 1)} \end{aligned}$$

if  $N_1 \leq \frac{\widehat{N}}{2}$ .  $\square$

In order to simplify the above representations, let  $\alpha_1 = N_1/\widehat{N}$  and  $\alpha_2 = N_2/\widehat{N}$  denote the proportion of voters in the first and the second group. Replacing  $N_1$  by  $\alpha_1\widehat{N}$  and  $N_2$  by  $\alpha_2\widehat{N}$  and assuming that  $\widehat{N}$  is large give, we obtain the following approximation.

**Corollary 1:** Let  $c_3(\alpha_1, \alpha_2) = \frac{4\alpha_1^2 + 4\alpha_1\alpha_2 - 4\alpha_1 + 4\alpha_2^2 - 4\alpha_2 + 1}{4\alpha_1\alpha_2(\alpha_1 + \alpha_2 - 1)}$  if  $\alpha_1 \leq 0.50$  and  $c_3(\alpha_1, \alpha_2) = 1/\alpha_1$  if  $\alpha_1 > 0.50$ . Then for  $k = 1, 2, 3$ ,  $Piv_k(\lambda, N) \simeq c_3(\alpha_1, \alpha_2)\frac{1}{N}$ .

We finally obtain that, for  $N$  large, the probability of casting a decisive vote for a voter belonging to an electorate divided in three groups is approximately equal to the Shapley-Shubik index multiplied by  $c_3(\alpha_1, \alpha_2)$ . We give in Table 3 some computed values of  $c_3(\alpha_1, \alpha_2)$  for various values of  $\alpha_1$  and  $\alpha_2$ .

$\alpha_1 / \alpha_2$	1/3	0.35	0.40	0.45	0.50
1/3	2.250	-	-	-	-
0.35	2.248	2.245	-	-	-
0.40	2.219	2.214	2.188	-	-
0.45	2.145	2.143	2.130	2.099	-
0.50	2	2	2	2	2
> 0.50	$1/\alpha_1$	$1/\alpha_1$	$1/\alpha_1$	$1/\alpha_1$	$1/\alpha_1$

Table 3 : Values of  $c_3(\alpha_1, \alpha_2)$

These values show that the probability of casting a decisive vote is maximum when  $\alpha_1 = \alpha_2 = 1/3$ , i.e. when each of the three groups has the same size.

#### 4.4 The Symmetric Case

We consider here the case with  $N_1 = N_2 = \dots = N_K = \frac{\widehat{N}+1}{K}$  and we assume that  $N = \widehat{N} + 1$  is a multiple of  $K$ , which implies that  $K$  is odd. In this symmetric case, the value of  $\pi(\frac{\widehat{N}}{2}, \frac{\widehat{N}+1}{K} - 1, \frac{\widehat{N}+1}{K}, \dots, \frac{\widehat{N}+1}{K})$  is given as the number of integer solutions of the following set of (in)equalities:

$$\begin{aligned} 0 &\leq x_1 \leq \frac{\widehat{N}+1}{K} - 1 \\ 0 &\leq x_2 \leq \frac{\widehat{N}+1}{K} \\ &\dots \\ 0 &\leq x_K \leq \frac{\widehat{N}+1}{K} \\ x_1 + x_2 + \dots + x_K &= \frac{\widehat{N}}{2} \end{aligned}$$

For specific small values of  $K$ , it is fairly easy to obtain close forms for the probability of being pivotal as a function of the parameter  $N$ . Let us consider the first values of  $K$ .

- $K = 3$ . To compute  $\pi(\frac{\widehat{N}}{2}, \frac{\widehat{N}+1}{3} - 1, \frac{\widehat{N}+1}{3}, \frac{\widehat{N}+1}{3})$ , we make use of the Barvinok's algorithm to obtain:

$$\pi(\widehat{K} - 1, \widehat{K}, \widehat{K}) = \frac{(\widehat{N} + 2)(\widehat{N} + 4)}{12}$$

for  $\widehat{N} = 2$  modulo 6. From this result, we derive:

$$Piv(\lambda, N) = \frac{9(\widehat{N} + 2)}{4(\widehat{N} + 1)(\widehat{N} + 4)} \quad (2)$$

for  $\widehat{N} = 2$  modulo 6. Notice that this result is consistent with the representations given in Proposition 3: replacing  $N_1$  and  $N_2$  by  $(N + 1)/3$  in the formulas given in this proposition leads to (2).

Hence, we get for  $N$  large:

$$Piv(\lambda, N) \simeq c_3 \frac{1}{N}$$

with  $c_3 = \frac{9}{4} = 2.25$ , in accordance with the result obtained in the preceding subsection for  $\alpha_1 = \alpha_2 = 1/3$ .

- $K = 5$ . We obtain via Barvinok's algorithm:

$$\pi\left(\frac{\widehat{N}}{2}, \frac{\widehat{N}+1}{5} - 1, \frac{\widehat{N}+1}{5}, \frac{\widehat{N}+1}{5}, \frac{\widehat{N}+1}{5}, \frac{\widehat{N}+1}{5}\right) = \frac{(\widehat{N} + 2)(\widehat{N} + 6)(23\widehat{N}^2 + 276\widehat{N} + 928)}{24000}$$

from which we deduce

$$Piv(\lambda, N) = \frac{25(\widehat{N} + 2)(23\widehat{N}^2 + 276\widehat{N} + 928)}{192(\widehat{N} + 1)(\widehat{N} + 6)^3}$$

for  $\widehat{N} = 4$  modulo 10. In this case, the limiting value of the probability of casting a decisive vote is given as:

$$Piv(\lambda, N) \simeq c_5 \frac{1}{N}$$

with  $c_5 = \frac{575}{192} = 2.995$ .

•  $K = 7$ ,  $K = 9$  and  $K = 11$ . Although we have been able to obtain the complete polynomials associated with  $\pi(\frac{\widehat{N}}{2}, \frac{\widehat{N}+1}{K} - 1, \frac{\widehat{N}+1}{K}, \dots, \frac{\widehat{N}+1}{K})$  for  $K = 7, 9, 11$ , we only give here the values of  $c_K$ :

$$c_7 = \frac{41209}{11520} = 3.577$$

$$c_9 = \frac{2337507}{573440} = 4.076$$

and

$$c_{11} = \frac{4199504287}{928972800} = 4.521.$$

For values of  $K$  higher than 11, the implementation of the Barvinok's algorithm demands a very long computation time that prevents from obtaining some numerical results. The following proposition describes the asymptotic behavior of  $\pi(\frac{\widehat{N}}{2}, \frac{\widehat{N}+1}{K} - 1, \frac{\widehat{N}+1}{K}, \dots, \frac{\widehat{N}+1}{K})$  when  $N$  gets large.

**Proposition 4:** Let  $K$  be an odd number ( $K \geq 3$ ).

Let  $\varphi(K) = \lim_{N \rightarrow +\infty} [\frac{1}{N^{K-1}} \pi(\frac{\widehat{N}}{2}, \frac{\widehat{N}+1}{K} - 1, \frac{\widehat{N}+1}{K}, \dots, \frac{\widehat{N}+1}{K})]$ . Then, for each fixed value of  $K$ , we have:

$$\varphi(K) = \frac{1}{(K-1)!} \sum_{m=0}^{\frac{K-1}{2}} (-1)^m \binom{K}{m} \left(\frac{K-2m}{2K}\right)^{K-1}.$$

*Proof.* By definition,  $\pi(\frac{\widehat{N}}{2}, \frac{\widehat{N}+1}{K} - 1, \frac{\widehat{N}+1}{K}, \dots, \frac{\widehat{N}+1}{K})$  is the number of integer solutions of the following parametric linear system:

$$\left\{ \begin{array}{l} 0 \leq x_1 \leq \frac{N}{K} - 1 \\ 0 \leq x_k \leq \frac{N}{K} \quad \text{for all } k = 2, \dots, K \\ \sum_{k=1}^K x_k = \frac{N-1}{2} \end{array} \right.$$

We know by Ehrhart's theorem (Ehrhart, 1967) that this number is a quasi-polynomial of degree  $K - 1$  on the variable  $N$ . Hence,  $\varphi(K)$  is equal to the leading coefficient of this quasi-polynomial<sup>17</sup>. As the additive constants in the second member of the constraints do not affect this coefficient,  $\varphi(K)$  is also the leading coefficient of the quasi-polynomial computing the number of integer solutions of the system

$$\begin{cases} 0 \leq x_k \leq \frac{N}{K} & \text{for all } k = 1, 2, \dots, K \\ \sum_{k=1}^K x_k = \frac{N}{2} \end{cases}$$

The system represents the dilatation by the factor  $N$  of the rational  $(K - 1)$ -dimensional polytope  $\mathbf{Q}$  defined by:

$$\begin{cases} 0 \leq x_k \leq \frac{1}{K} & \text{for all } k = 1, \dots, K \\ \sum_{k=1}^K x_k = \frac{1}{2} \end{cases}$$

By Ehrhart's theorem, and by definition of  $\varphi(K)$ , we know that  $\varphi(K)$  is equal to the relative volume of  $\mathbf{Q}$ , which is the (normalized) volume in  $\mathbb{R}^{K-1}$  of the full-dimensional polytope  $\mathbf{P}$  defined by:

$$\begin{cases} 0 \leq x_k \leq \frac{1}{K} & \text{for all } k = 1, \dots, K - 1 \\ \frac{K - 2}{2K} \leq \sum_{k=1}^{K-1} x_k \leq \frac{1}{2} \end{cases}$$

Let  $\text{Vol}(\mathbf{P})$  be the volume of  $\mathbf{P}$ . To compute this volume, we consider some particular subsets of  $\mathbb{R}^{K-1}$ . Let  $\Delta$  and  $\Delta'$  be the  $K - 1$ -dimensional simplices defined by:

$$\Delta = \{x \in \mathbb{R}^{K-1} : x_k \geq 0 \text{ for all } k = 1, \dots, K - 1 \text{ and } x_1 + \dots + x_{K-1} \leq 1/2\}$$

$$\Delta' = \{x \in \mathbb{R}^{K-1} : x_k \geq 0 \text{ for all } k = 1, \dots, K - 1 \text{ and } x_1 + \dots + x_{K-1} \leq (K - 2)/2K\}$$

It is easy to see that  $\text{Vol}(\mathbf{P}) = \text{Vol}(A) - \text{Vol}(B)$ , where:

$$A = \{x \in \Delta : x_k \leq 1/K, \forall k = 1, \dots, K - 1\}$$

$$B = \{x \in \Delta' : x_k \leq 1/K, \forall k = 1, \dots, K - 1\}$$

---

<sup>17</sup>A degree- $d$  quasi-polynomial on the variable  $n$  is a polynomial expression  $f(n) = \sum_{i=0}^d c_i(n)n^i$ , where the coefficients  $c_i(n)$  are rational periodic numbers on  $n$ . A rational periodic number, of period  $q$ , on the integer variable  $n$  is a function  $U : \mathbb{Z} \rightarrow \mathbb{Q}$  such that  $U(n) = U(n')$  whenever  $n \equiv n' \pmod{q}$ .



We only show how to compute  $\text{Vol}(A)$ , the same method will be applied to obtain  $\text{Vol}(B)$ . For each  $i$  in  $\{1, \dots, K-1\}$  let  $\Delta_i = \{x \in \Delta : x_i \geq 1/K\}$ . More generally, for each non empty subset  $S$  of  $\{1, \dots, K-1\}$ , we define  $\Delta_S$  by  $\Delta_S = \cap_{i \in S} \Delta_i$ . Note that  $\Delta_S = \emptyset$  for  $|S| > \frac{K-1}{2}$ .

For  $S$  such that  $\#S \leq \frac{K-1}{2}$ , let  $\#S = m$  and let  $t_u$  be the translation of vector  $u$ , where  $u$  is the vector of  $\mathbb{R}^{K-1}$  defined by  $u_i = -\frac{1}{K}$  if  $i \in S$  and  $u_i = 0$  if not. It is obvious that  $t_u(\Delta_S) = \Delta(m)$ , where  $\Delta(m) = \{x \in \mathbb{R}^{K-1} : x_k \geq 0 \text{ for all } k = 1, \dots, K-1 \text{ and } x_1 + \dots + x_{K-1} \leq (K-2m)/2K\}$ . Since translations conserve volumes, and applying the formula giving the volume of a simplex, we obtain:

$$\text{Vol}(\Delta_S) = \text{Vol}(\Delta(m)) = \frac{1}{(K-1)!} \left( \frac{K-2m}{2K} \right)^{K-1}$$

On the other hand, we can write  $\text{Vol}(A) = \text{Vol}(\Delta) - \text{Vol}(\cup_{i=1}^{K-1} \Delta_i)$ . Applying the inclusion-exclusion principle, we get:

$$\begin{aligned} \text{Vol}(\cup_{i=1}^{K-1} \Delta_i) &= \sum_{m=1}^{\frac{K-1}{2}} (-1)^{m-1} \sum_{S, |S|=m} \text{Vol}(\Delta_S) \\ &= \sum_{m=1}^{\frac{K-1}{2}} (-1)^m \binom{K-1}{m} \frac{1}{(K-1)!} \left( \frac{K-2m}{2K} \right)^{K-1} \end{aligned}$$

Since  $\text{Vol}(\Delta) = \frac{1}{(K-1)!} \left( \frac{1}{2} \right)^{K-1}$ , we obtain:

$$\text{Vol}(A) = \frac{1}{(K-1)!} \sum_{m=0}^{\frac{K-1}{2}} (-1)^m \binom{K-1}{m} \left( \frac{K-2m}{2K} \right)^{K-1}$$

Now,  $\text{Vol}(B)$  can be computed in a similar way and we can easily establish that:

$$\text{Vol}(B) = \frac{1}{(K-1)!} \sum_{m=0}^{\frac{K-3}{2}} (-1)^m \binom{K-1}{m} \left( \frac{K-2-2m}{2K} \right)^{K-1}.$$

Finally, the following simple calculus gives the result:

$$\begin{aligned}
\text{Vol}(\mathbf{P}) &= \frac{1}{(K-1)!} \left( \sum_{m=0}^{\frac{K-1}{2}} (-1)^m \binom{K-1}{m} \left(\frac{K-2m}{2K}\right)^{K-1} - \sum_{m=0}^{\frac{K-3}{2}} (-1)^m \binom{K-1}{m} \left(\frac{K-2-2m}{2K}\right)^{K-1} \right) \\
&= \frac{1}{(K-1)!} \left( \left(\frac{1}{2}\right)^{K-1} + \sum_{m=1}^{\frac{K-1}{2}} (-1)^m \left[ \binom{K-1}{m} + \binom{K-1}{m-1} \right] \left(\frac{K-2m}{2K}\right)^{K-1} \right) \\
&= \frac{1}{(K-1)!} \left( \left(\frac{1}{2}\right)^{K-1} + \sum_{m=1}^{\frac{K-1}{2}} (-1)^m \binom{K}{m} \left(\frac{K-2m}{2K}\right)^{K-1} \right) \\
&= \frac{1}{(K-1)!} \sum_{m=0}^{\frac{K-1}{2}} (-1)^m \binom{K}{m} \left(\frac{K-2m}{2K}\right)^{K-1} \quad . \quad \square
\end{aligned}$$

Using the analytical expression obtained in Proposition 3, we can extend the calculation of  $c_K = K^K \varphi(K)$  to larger values of  $K$ . The following Table gives the exact value of  $c_K$  for  $K = 5$  to 49 ( $K$  odd).

$K$	5	7	9	11	13	15	17	19	21	23	25	27
$c_K$	2.995	3.577	4.076	4.521	4.925	5.298	5.647	5.976	6.288	6.584	6.870	7.143
$K$	29	31	33	35	37	39	41	43	45	47	49	51
$c_K$	7.408	7.657	7.903	8.141	8.372	8.597	8.817	9.031	9.240	9.444	9.644	9.840

Table 4 : Exact values of  $c_K$

Notice that the limiting result obtained in this subsection can be easily extended to the case where  $N$  is even and the population is divided into  $K$  groups of size  $\frac{N}{K}$ . The integer  $K$  can be odd or even and the unique assumption is that  $N$  is an even multiple of  $K$ . Let  $\psi(K) = \lim_{N \rightarrow +\infty} \left[ \frac{1}{N^{K-1}} \pi \left( \frac{N-2}{2}, \frac{N}{K} - 1, \frac{N}{K}, \dots, \frac{N}{K} \right) \right]$ . With slight modifications in the proof of Proposition 3, we obtain:

$$\psi(K) = \frac{1}{(K-1)!} \sum_{m=0}^{\frac{K-1}{2}} (-1)^m \binom{K}{m} \left(\frac{K-2m}{2K}\right)^{K-1}$$

if  $K$  is odd, and

$$\psi(K) = \frac{1}{(K-1)!} \sum_{m=0}^{\frac{K}{2}} (-1)^m \binom{K}{m} \left(\frac{K-2m}{2K}\right)^{K-1}$$

if  $K$  is even.

As the relation in Proposition 4 is not easy to implement when  $K$  becomes large, we have developed a probabilistic argument to conjecture that:

$$Piv(\lambda, N) \simeq \frac{1}{N} \sqrt{\frac{6K}{\pi}} \text{ when } K \text{ is large.}$$

Our conjecture on the asymptotic behavior of  $c_K$  when  $K$  tends to  $\infty$  is based upon the following heuristic probabilistic argument. Proceeding as in remark 2 of section 3, the probability of being pivotal for a small group of size  $\epsilon N$  where  $\epsilon > 0$  is fixed can be expressed as the probability of the event :

$$\frac{N}{2} - \frac{\epsilon N}{2} \leq S_N \leq \frac{N}{2} + \frac{\epsilon N}{2}$$

where:

$$S_N = \sum_{k=1}^K S_N^k \text{ where } S_N^k = \sum_{i=1}^{N_k} X_i^k \text{ and } N_k = \frac{N}{K}$$

The random variables  $S_N^1, S_N^2, \dots, S_N^K$  are independent and identically distributed. Following the argument used in Proposition 4 of Chamberlain and Rothschild (1981), we deduce that for all  $k = 1, \dots, K$ ,  $\frac{S_N^k}{N_k}$  converges weakly to the uniform law on the interval  $[0, 1]$  when  $N_k \rightarrow \infty$ . Since the  $S_N^k$  are independent, this implies that  $\frac{S_N}{N}$  converges weakly to  $Z = \frac{1}{K} \sum_{k=1}^K U^k$  where the random variables  $U_k$  are independent and identically distributed, with common distribution the uniform distribution on  $[0, 1]$ . From the central limit theorem, we deduce that if  $K$  is large then:

$$\frac{\sum_{k=1}^K U_k}{K} - \frac{1}{2} \simeq N\left(0, \frac{1}{\sqrt{12K}}\right)$$

since  $\sqrt{\frac{1}{12}}$  is the standard deviation of the uniform variable on  $[0, 1]$ . We deduce then that the probability of a group of relative size  $\epsilon$  to be pivotal denoted  $Piv(\epsilon, N)$  is approximatively equal to

$$\begin{aligned} \Pr \left\{ -\frac{\epsilon}{2} \leq \frac{\sum_{k=1}^K U_k}{K} - \frac{1}{2} \leq \frac{\epsilon}{2} \right\} &\simeq \Pr \left\{ -\frac{\epsilon}{2} \leq N\left(0, \frac{1}{\sqrt{12K}}\right) \leq \frac{\epsilon}{2} \right\} \\ &\simeq \epsilon \sqrt{\frac{6K}{\pi}} \end{aligned}$$

Some values of  $\sqrt{\frac{6K}{\pi}}$  are tabulated below:

$K$	3	5	7	9	11	...	49	...	99
$\sqrt{\frac{6K}{\pi}}$	2.3937	3.0902	3.6564	4.1459	4.5835	...	9.6738	...	13.7505

Table 5 : Approximate values of  $c_K$ 

## 5 Concluding Remarks

In this paper, we have studied the impact of correlation across preferences and votes on the probability of being pivotal under the majority rule and we have shown how increasing correlation reduces this probability. To illustrate our contribution, consider 1000 voters divided into  $K$  independent groups of equal size. In each group, preferences are generated according to the IAC assumption. When  $K = 1000$  we obtain the IC model with a probability of being pivotal equal to  $\sqrt{\frac{1}{500\pi}} = .0252$ ; the case with  $K = 1$  corresponds to the usual IAC model, with a probability of being pivotal equal to  $\frac{1}{1000}$ . The tools offered in this paper allow to consider all the possible intermediate situations between these two polar cases. The following Table displays the probability of being pivotal for various values of  $K$ : these probabilities are computed with the help of Proposition 1 for large values of  $K$  ( $K > 100$ ) and are deduced from Proposition 4 for small values of  $K$ .

$K$	1	2	4	5	...	50	...	100	...	200	250	500	1000
$Piv(\lambda)$	.0010	.0020	.0027	.0030	...	.0096	...	.0138	...	.0195	.0178	.0219	.0252

Table 6 : Probability of being pivotal, 1000 voters,  $K$  groups

When the group sizes are not equal, our results suggest that, for large electorates, the probability of casting a decisive vote does not depend on the size of the group to which the voter belongs and is only governed by the distribution of the group sizes.

Let us conclude with two remarks.

In this paper, we have mostly focused on a specific pattern of correlation that we call the IAC partitioning model. It is important to recall that this model is specific on two grounds. First, it is based on a partition of the individuals such that individuals belonging to two different groups in that partition have independent preferences. Second, it has been assumed that in each group the correlations among the preferences in the group were resulting from the IAC model. In this remark, we keep the partitioning assumption but examine a particular generalization of the existing IAC version.

In the IAC setting, the correlation coefficient between the votes of two voters from the same group is equal to  $\frac{1}{3}$ . Let us consider instead the case where the correlation coefficient between the votes of two voters is positive but arbitrary<sup>18</sup> and denoted by  $\rho$ :  $Cov(X_{ik}, X_{jk})$ ,

<sup>18</sup>In Appendix 2, we prove that this construction is always possible. An alternative construction in the spirit of the definition of IAC due to Berg (1990) is also possible through the family of Beta densities.

the covariance between the votes of  $i$  and  $j$  when they belong to the same group is then equal to  $\frac{\rho}{4}$ . As before, as long as  $\rho \neq 1$  we obtain:

$$\lim_{R \rightarrow \infty} \sqrt{N} Piv_{(s)}(\lambda_R, N) = \frac{1}{\left( \sqrt{\frac{1-\rho+\rho\gamma^1}{4} + \frac{\rho}{4} \sum_{l=2}^S \gamma^l l} \right) \sqrt{2\pi}}$$

In particular, in the case where  $N$  is a multiple of  $s$  and all groups are of size  $s$ , we obtain:

$$\lim_{R \rightarrow \infty} \sqrt{N} Piv_{(s)}(\lambda_R, N) = \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{1 + \rho(s-1)}}$$

We observe that  $\sqrt{N} Piv_{(s)}(\lambda_R, N)$  decreases with  $s$  and with  $\rho$ . This is consistent with intuition as an increase in  $s$  or an increase in  $\rho$  leads to more correlation among the votes and less room for pivotality.

Our second remark is about another key assumption, namely the neutrality among the two alternatives. We have assumed that the two alternatives were similar ex ante. One interesting generalization could consist in assuming that there is a partition of the population into groups where in each group the preferences are as here correlated but also possibly biased towards one candidate. The bias could of course vary from group to another. In such a setting a group could be defined as a subset of individuals displaying some homogeneity defined through a vector of characteristics.

We are not aware of an ambitious attempt to generalize the current theory to a setting that would allow for differences across alternatives. To the best of our knowledge, the only<sup>19</sup> model along these lines is due to Beck (1975). He considers a population divided into two groups of equal size. In the first group, the votes are independent and people vote left with probability  $p \geq \frac{1}{2}$ . In the second group, votes are also independent and people vote left with probability  $1 - p$ . Beck estimates numerically the probability for a voter to be pivotal for several values of the parameter  $p$ . Modulo a simple adjustment of the proof of Proposition 1, we obtain an asymptotic exact value of the probability of being pivotal in Beck's model. Precisely, we obtain :

$$\lim_{N \rightarrow \infty} \sqrt{N} Piv(\lambda, N) = \frac{1}{\sqrt{2\pi p(1-p)}}$$

When  $p = \frac{1}{2}$ , we obtain the traditional constant  $\sqrt{\frac{2}{\pi}} = 0.79788$ . When  $p = \frac{3}{4}$ , we obtain  $\frac{1}{\sqrt{2\pi \times \frac{3}{16}}} = 0.92132$  and when  $p = \frac{4}{5}$ , we obtain  $\frac{1}{\sqrt{2\pi \times \frac{4}{25}}} = 0.99736$ . Moving towards polarization increases drastically the probability of being pivotal!

<sup>19</sup>See also Berg (1990) for another illustration.

## 6 Appendix

### 6.1 Berk's Theorem

For each  $k = 1, 2, \dots$  let  $n = n(k)$  and  $m = m(k)$  be specified and suppose that  $y_1^k, y_2^k, \dots, y_n^k$  is an  $m$ -dependent triangular array of random variables with zero means<sup>20</sup>. Assume the following conditions hold. For some  $\delta > 0$  and some constants  $M$  and  $K$ :

- (i)  $E |y_i^k|^{2+\delta} \leq M$  for all  $i$  and all  $k$ .
- (ii)  $Var(y_{i+1}^k + \dots + y_j^k) \leq (j - i) K$  for all  $i, j$ , and  $k$ .
- (iii)  $\lim_{k \rightarrow \infty} \frac{Var(y_1^k + \dots + y_n^k)}{n}$  exists and is nonzero. Denote  $v$  the limit.
- (iv)  $\lim_{k \rightarrow \infty} \frac{m^{2+\frac{\delta}{2}}}{n} = 0$

Then  $\frac{y_1^k + \dots + y_n^k}{\sqrt{n}}$  is asymptotically normal with mean 0 and variance  $v$ .

### 6.2 Positively Correlated Bernoulli Random Variables

In this appendix, we present a procedure to generate correlation patterns for vectors of Bernoulli random variables with parameter  $\frac{1}{2}$ . Consider a Gaussian vector  $z \equiv (z_1, z_2, \dots, z_n)$  such that the random variables  $z_i$  are identically distributed with first moment equal to 0 and let  $\Omega = (\sigma_{ij})_{1 \leq i, j \leq n} = \sigma^2 (\rho_{ij})_{1 \leq i, j \leq n}$  to denote its variance-covariance matrix with  $\sigma^2 = E(z_i^2)$  and  $\rho_{ii} = 1$  for all  $i = 1, \dots, n$ . We derive from  $z$  a vector  $x \equiv (x_1, x_2, \dots, x_n)$  of Bernoulli random variables such that  $\Pr(x_i = 1) = \frac{1}{2}$  for all  $i = 1, 2, \dots, n$  as follows. Let:

$$x_i = \begin{cases} 1 & \text{if } z_i \geq 0 \\ 0 & \text{if } z_i < 0 \end{cases}$$

From Stieljes's formula (Gupta (1963)), we obtain that<sup>21</sup>

$$Cov(x_i, x_j) = \Pr(z_i \geq 0 \text{ and } z_j \geq 0) - \frac{1}{4} = \frac{\arcsin \rho_{ij}}{2\pi}.$$

In particular when  $\Omega = I + \lambda \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}$  where  $\lambda > 0$ , we obtain  $\sigma^2 = 1 + \lambda$  and  $\sigma_{ij} = \lambda$

for all  $i \neq j$ . Note that  $\Omega$  is positive definite and is therefore the variance-covariance matrix of a gaussian vector. We derive  $\rho_{ij} = \frac{\lambda}{1+\lambda}$  for all  $i \neq j$ . In such setting, all the correlation coefficients are equal and span all the values from 0 to 1 when  $\lambda$  spans the range  $[0, +\infty[$ .

<sup>20</sup>The triangular array  $(y_{n(k)}^k)_{k \geq 1}$  is  $m$ -dependent if  $(y_1^k, y_2^k, \dots, y_j^k)$  and  $(y_{j+n}^k, y_{j+1+n}^k, \dots, y_{j+n+l}^k)$  are independent whenever  $n > m$  (Billingsley(1995)).

<sup>21</sup>Note that since  $\arcsin x = x + \frac{x^3}{6} + \frac{3x^5}{40} + O(x^5)$ , this implies that  $Cov(x_i, x_j) \simeq \frac{\rho_{ij}}{2\pi}$  in the neighborhood of 0.

### 6.3 Intragroup Perfect Correlation

In the case of perfect correlation i.e.  $\rho = 1$ , we need to be more careful as we cannot use Petrov's theorem. The reason is easy to see in the case where all the groups are of same size  $s$ . In such case the variables  $Z_k^R$  and  $W_i^R$  introduced in the proof of Proposition 1 have respectively a span of  $s$  and a span of  $s - 1$ . Only the  $Z_k^R$  variables appear infinitely often. To see what is going on, consider the case where  $s = 2$  i.e. the case where the  $N$  random variables are grouped into  $M \equiv \frac{N}{2}$  packs of size 2. Let us focus on the case where  $M$  is odd. In such case:

$$Piv(\lambda_R, 2) = \text{Prob}(S_i^R - W_i^R = \frac{N-2}{2})$$

Since all the variables in the sum are independent, identically distributed with a maximal span of 2 and a variance equal to 1, we deduce from the standard Moivre-Laplace's local theorem<sup>22</sup> that:

$$\lim_{R \rightarrow \infty} \frac{1}{2} \sqrt{\frac{N-2}{2}} \text{Prob}(S_i^R - W_i^R = \frac{N-2}{2}) = \frac{1}{\sqrt{2\pi}}$$

and therefore:

$$\lim_{R \rightarrow \infty} \sqrt{N} Piv(\lambda_R, 2) = \frac{2}{\sqrt{\pi}} \simeq 1.1284$$

which is different from corresponds to the value of  $\sqrt{\frac{2}{\pi} \frac{1}{\sqrt{1+\rho(s-1)}}} = \frac{1}{\sqrt{\pi}} \simeq 0.5642$  when  $\rho = 1$  and  $s = 2$ . More generally, consider the case of an arbitrary value of  $s$  i.e. the case where the  $N$  random variables are grouped into  $M \equiv \frac{N}{s}$  of size  $s$  which correspond to  $M = \frac{N}{s}$  independent and identically distributed Bernoulli random variables. Let us focus on the case where  $M$  is odd. As above, we deduce that:

$$Piv(\lambda_R, s) = \text{Prob}(S_i^R - W_i^R = \frac{N-s}{2}) \simeq \sqrt{\frac{2}{\pi} \frac{1}{\sqrt{M}}} = \sqrt{\frac{2s}{\pi} \frac{1}{\sqrt{N}}}$$

and therefore:

$$\lim_{R \rightarrow \infty} \sqrt{N} Piv(\lambda_R, s) = \sqrt{\frac{2s}{\pi}}$$

---

<sup>22</sup>Alternatively, it is also a direct consequence of Gnedenko's theorem (1948) (theorem 1 in chapter 7 of Petrov (1975)) as all the variables in the sum are independent, identically distributed with a maximal span of 2 and variance equal to 1.

For the case where  $s = 3$ , we obtain  $\sqrt{\frac{6}{\pi}} = 1.382$  which is larger than  $\sqrt{\frac{2}{\pi} \frac{1}{\sqrt{1+\rho(s-1)}}} = \sqrt{\frac{2}{\pi} \frac{1}{\sqrt{3}}} = 0.46066$ . This discontinuity (we jump from  $\sqrt{\frac{2}{\pi s}}$  to  $\sqrt{\frac{2s}{\pi}}$ ) in the neighborhood of  $\rho = 1$  is rather peculiar but corresponds to the fact that when  $\rho = 1$ , the voters belonging to the same group vote as a block. Everything is as if we had a population of  $\frac{N}{s}$  independent voters.

## 6.4 The Le Breton and Lepelley 's Historical case Study

An interesting setting where the correlations are not the same across all voters appears in the Le Breton and Lepelley (2011) study of the French electoral law of June 29 1820. This electoral law, known as the "law of double vote", has been used in France to elect the deputies from 1820 to 1830. France was divided into a number of electoral districts (the so called French "départements") and each district sent a number of deputies to the chamber. Each district was divided itself into subdistricts (the so called "arrondissements"). Each arrondissement elected one deputy and to be voter in an arrondissement, your amount of tax had to be above some fixed level (called the "cens"). In addition, the voters in the top quartile of the income distribution of the voters in the département were members of an additional electoral college which elected  $D$  additional deputies. These "rich" voters had a double vote: they voted in their arrondissement and also in the electoral college constituted at the level of the département. This explains the name which was given to this law. It was decided that  $\frac{3}{5}$  of the deputies was elected by the arrondissements and  $\frac{2}{5}$  by the voters in the top colleges. Le Breton and Lepelley (2011) study a symmetric version of that problem where there are  $K$  départements, with  $A$  arrondissements in each département and  $4r + 1$  voters in each arrondissement where  $r$  is an odd integer denoting the number of voters with two votes in that arrondissement. The size  $N$  of the chamber is therefore  $K(A + D)$ <sup>23</sup>. A good approximation of the French data at that time is given by  $K = 86$ ,  $A = 3$  and  $D = 2$  leading to  $N = 430$ : 258 being elected in arrondissements and 172 elected by the top colleges. Hereafter, we will limit however our attention to the case where  $K$  is odd. In the case where  $A = 3$  and  $D = 2$ , the  $5K$  deputies are partitioned into groups of size 5. These legislators have in common to be elected from the same territory. Even if we assume that the preferences of the  $A(4r + 1)$  voters across the  $A$  districts are independent, the preferences of the deputies are not independent because some voters have a double vote. Let  $(S_j^1, S_j^2, S_j^3, S_j^4, S_j^5)$  be the profile of the five votes in the  $j^{th}$  département where the first three coordinates denote the votes in the three arrondissements and the last two the votes in the top college. When  $r$  is

---

<sup>23</sup>They assumed that  $A$  is an odd integer and that  $D$  is an even integer.



large this random vector is approximatively Gaussian with (after normalization) the matrix of variances-covariances:

$$\Omega = \begin{pmatrix} \frac{\sqrt{4r+1}}{2} & 0 & 0 & \frac{\sqrt{r}}{2} & \frac{\sqrt{r}}{2} \\ 0 & \frac{\sqrt{4r+1}}{2} & 0 & \frac{\sqrt{r}}{2} & \frac{\sqrt{r}}{2} \\ 0 & 0 & \frac{\sqrt{4r+1}}{2} & \frac{\sqrt{r}}{2} & \frac{\sqrt{r}}{2} \\ \frac{\sqrt{r}}{2} & \frac{\sqrt{r}}{2} & \frac{\sqrt{r}}{2} & \frac{\sqrt{3r}}{2} & \frac{\sqrt{3r}}{2} \\ \frac{\sqrt{r}}{2} & \frac{\sqrt{r}}{2} & \frac{\sqrt{r}}{2} & \frac{\sqrt{3r}}{2} & \frac{\sqrt{3r}}{2} \end{pmatrix}$$

We note that the coefficient of correlation  $\rho$  between any of the first three variables and any of the last two ones is equal to  $\sqrt{\frac{1}{12}}$ . Consider now the 5-dimensional vector of Bernoulli variables  $(X_j^1, X_j^2, X_j^3, X_j^4, X_j^5)$  where  $X_j^l = 1$  if  $S_j^l \geq 2r + 1$  for  $l = 1, 2, 3$  and  $X_j^l = 1$  if  $S_j^l \geq \frac{Ar+1}{2}$  for  $l = 4, 5$ . Based on the Gaussian orthant probabilities, the matrix of variances-covariances of this vector is:

$$\begin{pmatrix} \frac{1}{4} & 0 & 0 & \frac{1}{4} + \frac{\arcsin \rho}{2\pi} & \frac{1}{4} + \frac{\arcsin \rho}{2\pi} \\ 0 & \frac{1}{4} & 0 & \frac{1}{4} + \frac{\arcsin \rho}{2\pi} & \frac{1}{4} + \frac{\arcsin \rho}{2\pi} \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} + \frac{\arcsin \rho}{2\pi} & \frac{1}{4} + \frac{\arcsin \rho}{2\pi} \\ \frac{1}{4} + \frac{\arcsin \rho}{2\pi} & \frac{1}{4} + \frac{\arcsin \rho}{2\pi} & \frac{1}{4} + \frac{\arcsin \rho}{2\pi} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} + \frac{\arcsin \rho}{2\pi} & \frac{1}{4} + \frac{\arcsin \rho}{2\pi} & \frac{1}{4} + \frac{\arcsin \rho}{2\pi} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

In the specific case where  $\rho = \sqrt{\frac{1}{12}}$ , we obtain that  $\frac{1}{4} + \frac{\arcsin \rho}{2\pi} = 0.29849$ . Since the random variables  $X_j^1 + X_j^2 + X_j^3 + X_j^4 + X_j^5$  are independent, identically distributed and have a span equal to 1, by using the same argument as in the proof of Proposition 1, we deduce that the probability for a deputy to be pivotal if both  $r$  and  $K$  are large integers is approximatively equal to:

$$\frac{1}{\left(\frac{1}{4} + \frac{1}{\sqrt{5}} \left[\frac{1}{2} + 6 \times 0.29849\right]\right) \sqrt{2\pi N}} \simeq \frac{0.31301}{\sqrt{N}}$$

## 6.5 Petrov Local Limit Central Theorem

Let  $k$  be an arbitrary fixed positive integer. A sequence of random variables  $(y_n)_{n \geq 1}$  is said to be a  $k$ -sequence if the number of different distribution functions in the sequence of the distribution functions corresponding to  $(y_n)_{n \geq 1}$  is equal to  $k$ . Consider a  $k$ -sequence of independent integer-valued random variables  $(y_n)_{n \geq 1}$  each having finite variance. We denote by  $F^1, \dots, F^l$  the  $l$  distributions which are non-degenerate and occur infinitely often in the sequence  $(F^i)_{1 \leq i \leq k}$ . We denote by  $H^r$  the maximal span of  $F^r$  for  $r = 1, \dots, l$ . Let

$S_n = \sum_{j=1}^n y_j$ ,  $M_n = \sum_{j=1}^n E(y_j)$ ,  $B_n = \sum_{j=1}^n E(y_j - E(y_j))^2$  and  $\Pr_n(N) = \Pr(S_n = N)$ .  
Then:

If g.c.d.  $(H^1, H^2, \dots, H^l) = 1$ , then  $\text{Sup}_N \left| \sqrt{B_n} \Pr_n(N) - \frac{1}{\sqrt{2\pi}} e^{-\frac{(N-M_n)^2}{2B_n}} \right| \xrightarrow{n \rightarrow \infty} 0$

## 7 References

Banzhaf, J.F. (1965) "Weighted voting does not work : a mathematical analysis", *Rutgers Law Review*, 19, 317-343.

Barberà, S. and M.O. Jackson (2006) "On the Weights of Nations : Assigning Voting Weights in a Heterogeneous Union", *Journal of Political Economy*, 114, 317-339.

Barvinok, A. (1994) "Polynomial time algorithm for counting integral points in polyhedra when the dimension is fixed". *Math. Oper. Res.*, 19, 769-779.

Barvinok, A. and J. Pommersheim (1999) "An algorithmic theory of lattice points in polyhedra". in: New perspectives in algebraic combinatorics, Berkeley, Ca, 1996-1997. *Math. Sci. Res. Inst. Publ.*, 38, 91-147.

Beck, M. and S. Robins (2007) *Computing the continuous discretely: Integer-point enumeration in polyhedra* Undergraduate Texts in Mathematics, Springer-Verlag, New York.

Beck, N. (1975) "A Note on the Probability of a Tied Election", *Public Choice*, 23, 75-79.

Berg, S. (1990) "The Probability of Casting a Decisive Vote : The Effect of a Caucus", *Public Choice*, 64, 73-92.

Berk, K.N. (1973) "A Central Limit's Theorem for  $m$ -Dependent Random Variables with Unbounded  $m$ ", *Annals of Probability*, 1, 352-354.

Berg, S. (1990) "The Probability of Casting a Decisive Vote: The Effects of a Caucus", *Public Choice*, 64, 73-92.

Billingsley, P. (1995) *Probability and Measure*, Third Edition, New York, John Wiley and Sons.

Chamberlain, G. and M. Rothschild (1981) "A Note on the Probability of Casting a Decisive Vote", *Journal of Economic Theory*, 25, 152-162.

Ehrhart, E. (1962) "Sur les Polyèdres Rationnels Homothétiques à  $n$  Dimensions", *C. R. Acad. Sci. Paris*, 254, 616-618.

Ehrhart, E. (1967) "Sur un problème de Géométrie Diophantienne Linéaire", *Journal für die R. und A. Math.*, 226, 1-49.

Ehrhart E. (1977) "Polynômes Arithmétiques et Méthodes des Polyèdres en Combinatoire" in *International Series of Numerical Mathematics*, Basles/Stuttgart: Birkhauser

- Fishburn, P.C. and Gehrlein, W.V. (1976) "Borda's Rule, Positional Voting, and Condorcet's Simple Majority Principle", *Public Choice*, 28, 79-88.
- Flajolet, P. and Sedgewick R. (2009) *Analytic Combinatorics*, Cambridge University Press
- Good, I.J. and L.S. Mayer (1975) "Estimating the Efficacy of a Vote", *Behavioral Science*, 20, 25 -33.
- Gnedenko, B.V. (1948) "On a Local Limit Theorem of the Theory of Probability", *Upsekhi Mat Nauk*, 3, 187-194.
- Gupta, S.S. (1963) "Probability Integrals of Multivariate Normal and Multivariate  $t$ ", *Annals of Mathematical Statistics*, 34, 792-828.
- Kleiner, A. (1980) "A Weighted Voting Model ", *Mathematics Magazine*, 53, 28-32.
- Kuga, K. and H. Nagatani (1974) "Voter Antagonism and the Paradox of Voting", *Econometrica*, 42, 1045-1067.
- LattE homepage at <http://www.math.ucdavis.edu/latte>.
- Le Breton, M. and D. Lepelley (2014) "Une Analyse de La Loi Electorale du 29 Juin 1820", *Revue Economique*, 65, 469-518.
- Lepelley, D, Louichi, A. et H. Smaoui (2008) "On Ehrhart's Polynomials and Probability Calculations in Voting Theory", *Social Choice and Welfare*, 30, 363-383.
- McDonald, D. (1979) "On Local Limit Theorems for Integer Valued Random Variables", *Theory of Probability and Applications*, 24, 613-619.
- Owen, G. (2001) *Game Theory*, Third Edition, Academic Press, New York.
- Penrose, L.S. (1946) "The elementary statistics of majority voting", *Journal of the Royal Statistical Society*, 109, 53-57.
- Penrose, L.S. (1952) *On the Objective Study of Crowd Behavior*, Lewis and Co, London.
- Petrov, V.V. (1975) *Sums of Independent Random Variables*, Springer-Verlag, Heidelberg.
- Shapley, L.S. (1962) "Simple Games: An Outline of the Descriptive Theory", *Behavioral Science*, 7, 59-66.
- Shapley, L.S. and M. Shubik (1954) "A method for evaluating the distribution of power in a committee system", *American Political Science Review*, 48, 787-792.
- Straffin, P.D. (1977) "Homogeneity, Independence, and Power Indices", *Public Choice*, 30, 107-118.
- Straffin P.D. (1988) "The Shapley-Shubik and Banzhaf Power Indices as Probabilities", in A.E. Roth (eds.) *The Shapley Value : Essays in Honor of Lloyd S. Shapley*, Cambridge University Press
- Taylor, A.D. and W.S. Zwicker (1999) *Simple Games*, Princeton, Princeton University Press.

Verdoolaege, S., Seghir, R. Beyls, K., Loechner, V. and M. Bruynooghe (2004) "Analytical computation of Ehrhart polynomials: Enabling more compiler analysis and optimizations". *In Proceedings of International Conference on Compilers, Architecture, and Synthesis for Embedded Systems*, 248–258. Washington D.C.

Verdoolaege, S., Woods, K., Bruynooghe, M. and R. Cools (2005) "Computation and manipulation of enumerators of integer projections of parametric polytopes". Technical Report CW 392, Katholieke Universiteit Leuven, Department of Computer Sciences, Celestijnenlaan 200A–B-3001 Heverlee.

Wilson, M.C. and G. Pritchard (2007) "Probability Calculations under the IAC Hypothesis", *Mathematical Social Sciences*, 54, 244-256.