View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Toulouse Capitole Publications

WORKING PAPERS

N° TSE-511

July 2014

# "Influence Vs. Utility in the Evaluation of Voting Rules: A New Look at the Penrose Formula"

Michel Le Breton and Karine Van der Straeten

Toulouse School of Economics

# Influence Vs. Utility in the Evaluation of Voting Rules: A New Look at the Penrose Formula

Michel Le Breton* Karine Van der Straeten†

July 2014

**Abstract**

In this paper, we clarify the relationship between *influence/power measurement* and *utility measurement*, the most popular two social objective criteria used when evaluating voting mechanisms. For one particular probabilistic model describing the preferences of the electorate, the so-called *Impartial Culture* ($IC$) model used by Banzhaf, the Penrose formula show that the two objectives coincide. The $IC$ probabilistic model assumes that voter preferences are independent. In this note, we prove a general version of the Penrose formula, allowing for correlations in the electorate, and show that in that case, the two social objectives no longer coincide and qualitative conclusions can be very different.

## 1 Introduction

The purpose of this note is to clarify the relationship between *power measurement* and *utility measurement/voting design* in the case where a group (society, assembly, committee,...) must decide among two alternatives. Power measurement is a developed and popular area in applied political science which has already a long history and has received a great deal of attention while voting design (which we view as the application of mechanism design to the normative analysis of political institutions) is newer. One possible explanation of this discrepancy is the implicit belief that if a player is influential (powerful, pivotal,...) in an institution then her utility will also be large as she will be in position to reduce the gap between the collective decision and her favorite one. While intuitive, this assertion is not true in full generality and the relationship between the two notions will depend upon the *probabilistic model* describing the preferences of the electorate.

*Institut Universitaire de France and Toulouse School of Economics, France.
†Toulouse School of Economics, CNRS and Institute for Advanced Study in Toulouse (IAST), France.

The most popular two answers to the first question (what is the probability that my vote will make a difference in the group decision under a given mechanism?) are the *power indices* due to *Banzhaf* (1965) and *Shapley and Shubik* (1954); they are defined as the probability for a player to be influential for two different probabilistic models. *Sensitivity* is defined[1] by Felsenthal and Machover (1998) as the sum of the power indices over the individuals. To defend this quantity as a reasonable social objective, they write "the sum of power indices can be regarded as a measure of the sensitivity of the decision rule : the ease with which it responds to fluctuations in the voters' wishes".

In contrast to the first question, the second question (what is my expected utility under a given mechanism?) has been somehow neglected and the unique attempt to measure the satisfaction or utility of a player in a voting body is due to Rae (1969) who has proposed an *index of satisfaction*. Likely, one reason for this neglect is the nice relationship between satisfaction and power that was established earlier[2] by Penrose (1946) for the particular probabilistic model used by Banzhaf, the so-called *Impartial Culture* ($IC$) model. He shows that in such a case the satisfaction index is an affine transform of the power index[3]. Both Rae and Penrose suppose that the utility of any of the two alternatives can only take two values, which are the same for all the individuals. Note that in such a case, the utility of any player coincides with the probability that the group decision will agree with the player decision.[4]

In this note, we prove a general (while still assuming that the utility takes only two values) version of the Penrose formula which highlights the differences between power and utility measurement in a general probabilistic model. It calls the attention on the role of correlation as the main explanation of the gap between the two notions. After deriving a general formula for any random electorate (section 3), we offer a detailed calculation of total expected utility attached to any mechanism in the probabilistic model underlying the Shapley-Shubik power index, the so-called *Impartial Anonymous Culture* ($IAC$) model

[1]Definition 3.3.1. on page 52.

[2]Felsenthal and Machover (1998) report on the history of the result. The formula is also proved in Brams and Lake (1978) and Dubey and Shapley (1979). Dubey and Shapley comment that the connection between satisfaction and the Banzhaf's index was not noticed for several years after 1969. In fact, the connection has been noticed a good many years before, in 1946; it is stated in Penrose without proof, something the readers are expected to work out for themselves.

[3]Penrose's formula is quoted as theorem 3.2.16 in Felsenthal and Machover (1998). They reproduce the following nice observation by Penrose: " In general, the power of the individual vote can be measured by the amount by which his chance of being on the winning side exceeds one half. The power, thus defined, is the same as half the likelihood of a situation in which an individual can be decisive...".

[4]The extension of the Banzhaf's probabilistic model to more than two alternatives (known also as the $IC$ model) has been used by several authors including for instance Myerson (1998, 2000) and Weber (1995).

(section 4). We conclude that while the results obtained through IC are insightful, we should be careful about evaluating the qualities of alternative social mechanisms and deriving the optimal one on the sole basis of the IC model, since qualitative conclusions can be very different when allowing for some correlations in the electorate.

# 2 Random Electorates and Voting Mechanisms

We consider a society $N = \{1, 2, ..., n\}$ of voters facing a choice between two alternatives $A$ and $B$. The state space is the set of profiles of strict preferences. A profile of strict preferences is a vector $X \in \{0,1\}^N$ where $X_i = 1$ (respectively 0) means that voter $i$ prefers $A$ to $B$ (prefers $B$ to $A$). We will assume hereafter that the utility can only take two values: 1 for the best alternative and 0 for the worst.

A *random electorate* is a joint probability distribution $\lambda$ over $\{0,1\}^N$. Hereafter, we will denote $P_i = \Pr_\lambda [X_i = 1]$ the probability that voter $i$ prefers $A$ to $B$.

A random electorate $\lambda$ is *uniform* if $P_i = P_j = P$ for all $i, j \in N$.

It is *neutral* if $\lambda(X) = \lambda(1 - X)$ for all $X \in \{0,1\}^N$. Note that if $\lambda$ is neutral, then $P_i = \frac{1}{2}$ for all $i \in N$.

A random electorate $\lambda$ is *independent* if for all $X \in \{0,1\}^N$, $\lambda(X) = \prod_{i=1}^{n} (P_i)^{X_i} (1 - P_i)^{1-X_i}$.

The *Banzhaf random electorate* is the unique neutral and independent random electorate $\lambda$ defined by $\lambda(X) = \frac{1}{2^n}$ for all $X \in \{0,1\}^N$. In the social choice literature this model is referred to as Impartial Culture (IC).

The *Shapley-Shubik random electorate* $\lambda$ is defined by $\lambda(X) = \frac{1}{(n+1)\binom{n}{k}}$ for all $X \in \{0,1\}^N$ such that $\mid \{i \in N : X_i = 1\} \mid = k$. In the social choice literature, this model is referred to as Impartial Anonymous Culture (IAC).[5]

A *voting mechanism* is a monotonic mapping $\mathfrak{C}$ from $\{0,1\}^N$ into $\{0,1\}$.[6]

A mechanism $\mathfrak{C}$ is *anonymous* if for all $X \in \{0,1\}^N$ and all permutation $\sigma$ over $N$, $\mathfrak{C}(X) = \mathfrak{C}(X^\sigma)$ where $X^\sigma$ is the vector $\left(X_{\sigma(1)}, ..., X_{\sigma(n)}\right)$.

[5] As shown by Straffin (1978), the Shapley-Shubik random electorate is equivalently defined as the mixture according to the uniform density over $[0,1]$ of uniform and independent random electorates. This model has been pioneered independently in voting theory by Chamberlain and Rothschild (1981), Fishburn and Gehrlein (1976), Good and Mayer (1975) and Kuga and Nagatani (1974). His connection with the Shapley-Shubik power index, i.e. when vote takes place between two alternatives, was discovered by Straffin (1978).

[6] It is equivalently described by the inverse image $\mathfrak{C}^{-1}(1)$ of 1 by $\mathfrak{C}$, i.e. the list $\mathcal{W}$ of coalitions such that if $\{i \in N : X_i = 1\} \in W$, then $\mathfrak{C}(X) = 1$. The set $\mathcal{W}$ is refered to as the set of *winning coalitions*. Such a mechanism is also called a simple game (Shapley (1962), Taylor and Zwicker (1999)).

A mechanism $\mathfrak{C}$ is *neutral* if for all $X \in \{0,1\}^N$, $\mathfrak{C}(1-X) = 1-\mathfrak{C}(X)$. If $\lambda$ is neutral and $\mathfrak{C}$ is neutral then $\lambda(\mathfrak{C}) \equiv \lambda[X : \mathfrak{C}(X)=1] = \frac{1}{2}$.

Among the neutral voting mechanisms, we will consider $\mathfrak{Dic}$ and $\mathfrak{Maj}$: $\mathfrak{Dic}_i(X) = X_i$ (individual $i$ dictates his choice) and the ordinary majority mechanism $\mathfrak{Maj}$ defined[7] by:

$$\mathfrak{Maj}(X) = \begin{cases} 1 \text{ if } \sum_{1\leq i\leq n} X_i > \frac{n}{2}, \\ 0 \text{ if } \sum_{1\leq i\leq n} X_i < \frac{n}{2}. \end{cases}$$

Given a random electorate $\lambda$ and a voting mechanism $\mathfrak{C}$, *the influence (power) of individual* $i \in N$ is defined as the probability that $i$ is pivotal under mechanism $\mathfrak{C}$, that is:

$$Influence\ (i,\lambda,\mathfrak{C}) = \lambda[X : \mathfrak{C}(X_{-i}\cup\{0\}) = 0 \text{ and } \mathfrak{C}(X_{-i}\cup\{1\}) = 1], \tag{1}$$

and the expected utility of individual $i$, given our assumption that the utility can only take two values: 1 for the best alternative and 0 for the worst, is the probability that the collective decision coincides with individual $i$'s preferred alternative; it writes as:

$$Utility\ (i,\lambda,\mathfrak{C}) = \lambda[\mathfrak{C}(X) = X_i]. \tag{2}$$

On the aggregate side, the sensitivity[8] (total influence) of the mechanism $\mathfrak{C}$ is defined as:

$$Total\ Influence\ (\lambda,\mathfrak{C}) = \sum_{i\in N} Influence\ (i,\lambda,\mathfrak{C}),$$

and the expected aggregate utility writes as:

$$Total\ Utility\ (\lambda,\mathfrak{C}) = \sum_{i\in N} Utility\ (i,\lambda,\mathfrak{C}).$$

For any given random electorate $\lambda$, these two numerical evaluations define two orderings $\succsim_\lambda^{TI}$ and $\succsim_\lambda^{TU}$ over the set of mechanisms.

Note that since $Total\ Utility\ (\lambda,\mathfrak{C})$ coincides with the average number of people who agree with the collective decision, for all $\lambda$, the best mechanism according to $\succsim_\lambda^{TU}$is the ordinary majority mechanism. This means that an utilitarian mechanism designer facing no constraints in the choice of a mechanism will pick up the ordinary majority mechanism.

Obtained when $\lambda$ is the Banzhaf electorate (the $IC$ case), the celebrated Penrose formula (1946) writes as[9]:

$$Utility\ (i,IC,\mathfrak{C}) = \frac{1}{2} + \frac{1}{2} Influence\ (i,IC,\mathfrak{C}). \tag{3}$$

[7] When $n$ is odd.

[8] Since the word sensitivity can receive many different interpretations, we prefer to use the word total influence which is used in the mathematical literature on Boolean functions.

[9] A proof for this formula will be provided in section 3 when we establish a generalized Penrose formula.

The Penrose formula implies that if $\lambda$ is $IC$ then $\gtrsim_\lambda^{TI} = \gtrsim_\lambda^{TU}$ .

In the next section, we demonstrate a generalized version of the Penrose's formula which show that the affine relationship between influence and utility does not hold in general when the electorate is not $IC$. A third term, that we call a *correction term*, appears in the relationship.

Before turning to this general formula describing the link between influence and utility when we allow for some correlations in the electorate, let us first note that there are neutral and uniform random electorates $\lambda$ such that the strict components of the orderings $\gtrsim_\lambda^{TI}$ and $\gtrsim_\lambda^{TU}$ disagree. As a very simple, motivating example, consider the case where $n = 7$ and $\lambda$ is the random electorate where the mass is distributed uniformly on all the profiles where 5 voters are on one side and 2 voters are on the other side.

Consider first the ordinary majority mechanism. The probability that an individual is pivotal is equal to 0. Therefore, the total influence of this mechanism is equal to 0. On the other hand, the social utility attached to each profile is equal to 5. There the expected total utility is equal to 5.

Consider now the mechanism where 1 dictates his choice. The total influence of this mechanism is equal to 1. The ex post social utility is either equal to 5 if 1 is on the majority side or to 2 if 1 is on the minority side. Therefore, the expected social utility is equal to $\frac{5}{7} \times 5 + \frac{2}{7} \times 2 = \frac{29}{7}$. Without surprise given our observation that for all $\lambda$, the best mechanism according to $\gtrsim_\lambda^{TU}$is the ordinary majority mechanism, we have $\frac{29}{7} < 5$.

With this random electorate $\lambda$, we have $\mathfrak{Dic}_1 \succ_\lambda^{TI} \mathfrak{Maj}$ and $\mathfrak{Maj} \succ_\lambda^{TU} \mathfrak{Dic}_1$.

*Remark:* We will focus in this note on the most popular two criteria used to evaluate mechanisms: sensitivity and social utility/utilitarianism. Felsenthal and Machover (1999) add to these two notions, the notion of majority deficit. It is defined as follows.[10] Given a profile of preferences $X$ and a voting mechanism $\mathfrak{C}$, define the majority deficit of $\mathfrak{C}$ at $X$ as the difference between the size of the majority group at $X$ and the size of the winning group if this difference is positive and 0 otherwise. This number measures the gap, if any, between the majority and the minority. Given a random electorate $\lambda$ , the majority deficit of $\mathfrak{C}$ is the expected value (over $X$) of the majority deficit of $\mathfrak{C}$ at $X$.

Note that if $\lambda$ is $IC$, then[11], the majority deficit of $\mathfrak{C}$ is equal to:

$$\frac{Total\ Influence\,(IC, \mathfrak{Maj}) - Total\ Influence(IC, \mathfrak{C})}{2},$$

and in that case, all three criteria agree on the ranking of mechanisms.

[10]See e.g. page 60 in Felsenthal and Machover (1998).
[11]See theorem 3.3.17 in Felsenthal and Machover (1998).

# 3 A Generalized Penrose Formula

In this section, we propose a generalization of the Penrose formula (3) to any random electorate $\lambda$.

By definition (2), $Utility\ (i, \lambda, \mathfrak{C}) = \lambda[\mathfrak{C}(X) = X_i]$. We use the fact that the two sets $\{X : \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = 1 - X_i\}$ and $\{X : \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i\}$ form a partition of the state space, to decompose this expected utility as follows:

$$\begin{aligned} Utility\ (i, \lambda, \mathfrak{C}) &= \lambda\left[X : \mathfrak{C}(X) = X_i \text{ and } \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = 1 - X_i\right] \\ &\quad + \lambda\left[X : \mathfrak{C}(X) = X_i \text{ and } \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i\right]. \end{aligned} \tag{4}$$

Note that the first term on the right hand side of (4) is $Influence\ (i, \lambda, \mathfrak{C})$ (see definition (1)).[12] Besides, since the voting mechanism $\mathfrak{C}$ is monotonic, if $\mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i$, then $\mathfrak{C}(X) = X_i$, and therefore the second term on the right hand side of (4) can be simplified as:

$$\lambda\left[X : \mathfrak{C}(X) = X_i \text{ and } \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i\right] = \lambda\left[X : \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i\right].$$

Substituting in (4), one gets:

$$Utility\ (i, \lambda, \mathfrak{C}) = Influence\ (i, \lambda, \mathfrak{C}) + \lambda[X : \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i]. \tag{5}$$

To get a formula closer to Penrose's formula, note that:

$$Utility\ (i, \lambda, \mathfrak{C}) = 1 - \lambda[X : \mathfrak{C}(X) = 1 - X_i]. \tag{6}$$

Summing (5) and (6) and dividing by 2, one gets our generalized Penrose formula:

$$Utility\ (i, \lambda, \mathfrak{C}) = \frac{1}{2} + \frac{1}{2} Influence\ (i, \lambda, \mathfrak{C}) + \frac{1}{2}\left[\begin{array}{c} \lambda[X : \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i] \\ -\lambda[X : \mathfrak{C}(X_{-i} \cup \{X_i\}) = 1 - X_i] \end{array}\right]. \tag{7}$$

Formula (7) is a generalization of Penrose's formula (3) to any random electorate. Note that the third term of the right hand side of (7), which can been seen as a *correction term*

[12] Indeed,

$$\begin{aligned} & \lambda\left[X : \mathfrak{C}(X) = X_i \text{ and } \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = 1 - X_i\right] \\ = \; & \lambda\left[X_i = 0\right] \,.\, \lambda\left[X : \mathfrak{C}(X) = X_i \text{ and } \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = 1 - X_i \mid X_i = 0\right] \\ & + \lambda\left[X_i = 1\right] \,.\, \lambda\left[X : \mathfrak{C}(X) = X_i \text{ and } \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = 1 - X_i \mid X_i = 1\right] \\ = \; & (\lambda\left[X_i = 0\right] + + \lambda\left[X_i = 1\right]) \,.\, \lambda\left[X : \mathfrak{C}(X_{-i} \cup \{0\}) = 0 \text{ and } \mathfrak{C}(X_{-i} \cup 1) = 1\right] \\ = \; & \lambda\left[X : \mathfrak{C}(X_{-i} \cup \{0\}) = 0 \text{ and } \mathfrak{C}(X_{-i} \cup 1) = 1\right]. \end{aligned}$$

compared to the Penrose formula, receives a simple interpretation. It is (half) the difference between the probability that the collective decision coincides with individual $i$'s wish even if $i$ were to vote against his interest, and the probability that the collective decision does not coincide with individual $i$'s wish when $i$ votes for his preferred alternative.

To check that we get Penrose's formula when $\lambda$ is the $IC$ electorate, note that the correction term on the right hand side of (7) can write as:

$$\frac{1}{2}\left(\lambda[X : \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i] - \lambda[X : \mathfrak{C}(X_{-i} \cup \{X_i\}) = 1 - X_i]\right)$$
$$= \frac{1}{2}\lambda[X_i = 1] * \left(\lambda[X : \mathfrak{C}(X_{-i} \cup \{0\}) = 1 \mid X_i = 1] - \lambda[X : \mathfrak{C}(X_{-i} \cup \{1\}) = 0 \mid X_i = 1]\right)$$
$$+ \frac{1}{2}\lambda[X_i = 0] * \left(\lambda[X : \mathfrak{C}(X_{-i} \cup \{1\}) = 0 \mid X_i = 0] - \lambda[X : \mathfrak{C}(X_{-i} \cup \{0\}) = 1 \mid X_i = 0]\right).$$

When $\lambda$ is independent, one can drop all the conditioning in the equality above, and the correction term is:

$$\left(P_i - \frac{1}{2}\right) * \left(\lambda[X : \mathfrak{C}(X_{-i} \cup \{0\}) = 1] - \lambda[X : \mathfrak{C}(X_{-i} \cup \{1\}) = 0]\right).$$

When $\lambda$ is both independent and neutral (remember that neutrality implies that $P_i = 1/2$), this term is equal to zero and formula (7) boils down to Penrose's formula (3).

# 4 An application to the Shapley-Shubik $IAC$ Random Electorate

In this section, we apply our generalized formula to the popular Shapley-Shubik $IAC$ random electorate. In the first subsection we derive a general formula for the calculation of $Utility$ $(i, IAC, \mathfrak{C})$ for any voting mechanism $\mathfrak{C}$. In a second subsection, we use this formula to compute this expected utility in the case of an anonymous mechanism. In a last subsection, we compare the Banzhaf $IC$ setting and the Shapley-Shubik $IAC$ setting, highlighting the differences between the two models.

## 4.1 A general formula for the $IAC$ case

Consider our generalized Penrose formula (7). The term $Influence$ $(i, \lambda, \mathfrak{C})$ is the probability that $i$ is pivot under mechanism $\mathfrak{C}$. When $\lambda$ is the $IAC$ electorate, the probability that out of the $n-1$ individuals in $N\backslash\{i\}$, some $k$ predetermined voters vote for $A$ is the same independently of the identity of these $k$ voters, and this probability is $\frac{1}{n}\frac{1}{\binom{n-1}{k}}$. Denoting by

$\gamma_i(k,\mathfrak{C})$ the number of coalitions in $N\backslash\{i\}$ with $k$ votes for $A$ for which $i$ is pivotal, $Influence$ $(i, IAC, \mathfrak{C})$ writes as:

$$Influence\ (i, IAC, \mathfrak{C}) = \sum_{k=0}^{k=n-1} \frac{\gamma_i(k,\mathfrak{C})}{n\binom{n-1}{k}}. \tag{8}$$

Consider now the correction term $\frac{1}{2}\lambda[X : \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i] - \frac{1}{2}\lambda[X : \mathfrak{C}(X) = 1 - X_i]$ in (7), it is equal to:

$$\frac{1}{2}\left(\begin{array}{c} \lambda[X : X_i = 0 \text{ and } \mathfrak{C}(X_{-i} \cup \{1\}) = 0]+ \\ \lambda[X : X_i = 1 \text{ and } \mathfrak{C}(X_{-i} \cup \{0\}) = 1\ ] \end{array}\right) - \frac{1}{2}\left(\begin{array}{c} \lambda[X : X_i = 0 \text{ and } \mathfrak{C}(X_{-i} \cup \{0\}) = 1]+ \\ \lambda[X : X_i = 1 \text{ and } \mathfrak{C}(X_{-i} \cup \{1\}) = 0] \end{array}\right),$$

that is:

$$\frac{1}{2}\left(\begin{array}{c} \lambda[X : \mathfrak{C}(X_{-i} \cup \{0\}) = 1 \text{ and } X_i = 1]- \\ \lambda[X : \mathfrak{C}(X_{-i} \cup \{0\}) = 1 \text{ and } X_i = 0] \end{array}\right) + \frac{1}{2}\left(\begin{array}{c} \lambda[X : \mathfrak{C}(X_{-i} \cup \{1\}) = 0 \text{ and } X_i = 0]- \\ \lambda[X : \mathfrak{C}(X_{-i} \cup \{1\}) = 0 \text{ and } X_i = 1] \end{array}\right).$$

Denoting by $\alpha_i(k,\mathfrak{C})$ the number of coalitions in $N\backslash\{i\}$ with $k$ votes for $A$, for which $A$ wins even if $i$ votes $B$, and by $\beta_i(k,\mathfrak{C})$ the number of coalitions in $N\backslash\{i\}$ with $k$ votes for $A$, for which $B$ wins even if $i$ votes $A$, and using the property of the IAC electorate, one gets that the correction term is equal to:

$$\frac{1}{2}\sum_{k=0}^{k=n-1} \alpha_i(k,\mathfrak{C}) \frac{1}{n+1}\left(\frac{1}{\binom{n}{k+1}} - \frac{1}{\binom{n}{k}}\right) + \frac{1}{2}\sum_{k=0}^{k=n-1} \beta_i(k,\mathfrak{C}) \frac{1}{n+1}\left(\frac{1}{\binom{n}{k}} - \frac{1}{\binom{n}{k+1}}\right).$$

Since

$$\frac{1}{\binom{n}{k+1}} - \frac{1}{\binom{n}{k}} = \frac{1}{n}\frac{1}{\binom{n-1}{k}}\left[2k - (n-1)\right],$$

when $\lambda = IAC$, one gets for the correction term:

$$\begin{aligned} \frac{1}{2}\lambda[X \quad &: \quad \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i] - \frac{1}{2}\lambda[X : \mathfrak{C}(X) = 1 - X_i] \\ &= \quad \frac{1}{2}\sum_{k=0}^{k=n-1} \frac{\alpha_i(k,\mathfrak{C}) - \beta_i(k,\mathfrak{C})}{n\binom{n-1}{k}} \times \frac{2k - (n-1)}{n+1}. \end{aligned} \tag{9}$$

## 4.2 A focus on anonymous mechanisms in the IAC case

For the sake of illustration, let us focus on anonymous mechanisms[13]. Consider an arbitrary integer quota $q \in [0, n+1]$, such that option $A$ is chosen if and only if at least $q$ individuals vote for $A$. The case $q = 0$ (resp. $q = n+1$) corresponds to the mechanism which always

[13]The formulas are much more cumbersome for non anonymous mechanisms.

selects option $A$ (resp. option $B$). Let us first compute, for any $k \in [0, n-1]$, $\alpha_i(k,q)$, $\beta_i(k,q)$ and $\gamma_i(k,q)$. Since the mechanisms are anonymous, we can drop the $i$ subscript. Given quota $q$:

$$\begin{aligned}
\alpha(k,q) &= \begin{cases} \binom{n-1}{k} \text{ if } k = q, ..., n-1 \text{ and } q \leq n-1, \\ 0 \text{ otherwise,} \end{cases} \\
\beta(k,q) &= \begin{cases} \binom{n-1}{k} \text{ if } k = 0, 1, ..., q-2 \text{ and } q \geq 2, \\ 0 \text{ otherwise,} \end{cases} \\
\gamma(k,q) &= \begin{cases} \binom{n-1}{k} \text{ if } k = q-1 \text{ and } 1 \leq q \leq n, \\ 0 \text{ otherwise.} \end{cases}
\end{aligned} \tag{10}$$

We deduce from (8) that:

$$Influence\ (i, IAC, q) = \begin{cases} \frac{1}{n} \text{ for } 1 \leq q \leq n, \\ 0 \text{ for } q \in \{0, n+1\}. \end{cases}$$

Remark: Note that this provides a simple proof for the well known claim[14] that, for any non constant anonymous mechanism $\mathfrak{C}$, $Total\ Influence(IAC, \mathfrak{C}) = 1$. In the $IAC$ case, any difference in Total Utility between mechanisms therefore stems from the correction term.

Let us now compute the correction term (9).

Consider first the case $2 \leq q \leq n-1$. One gets by substituting the values for $\alpha(k,q)$ and $\beta(k,q)$:

$$\begin{aligned}
\frac{1}{2}\lambda[X &: \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i] - \frac{1}{2}\lambda[X : \mathfrak{C}(X) = 1 - X_i] \\
&= \frac{1}{2}\frac{1}{n\,(n+1)}\left(\sum_{k=q}^{k=n-1} [2k - (n-1)] - \sum_{k=0}^{k=q-2} [2k - (n-1)]\right).
\end{aligned}$$

Simple computations yield:

$$\sum_{k=q}^{k=n-1} [2k - (n-1)] = q(n-q) \text{ and } \sum_{k=0}^{k=q-2} [2k - (n-1)] = -(q-1)(n+1-q),$$

therefore

$$\frac{1}{2}\lambda[X : \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i] - \frac{1}{2}\lambda[X : \mathfrak{C}(X) = 1 - X_i] = -\frac{1}{2n} + \frac{(n+1-q)\,q}{n\,(n+1)}. \tag{11}$$

[14] Indeed as already pointed out, $Influence\ (i, IAC, \mathfrak{C})$ is the Shapley imputation of player $i$ in the simple game attached to $\mathfrak{C}$. If the game is symmetric, then all the players receive the same payoff in the Shapley solution (as it is a symmetric solution) and since the total payoff is 1 (as the game is not constant), the claim follows.

Consider now the case $q = 1$ (the mechanism selects option $A$ as soon as at least one voter votes for $A$) or the symmetric case $q = n$ (the mechanism selects option $B$ as soon as at least one voter votes for $B$).

When $q = 1$, $\alpha(0,1) = 0$ and $\alpha(k,1) = \binom{n-1}{k}$ for all $k \in [1, n-1]$, and $\beta(k,1) = 0$ for all $k \in [0, n-1]$. Therefore, substituting in (9), one gets:

$$\begin{aligned}\frac{1}{2}\lambda[X &: \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i] - \frac{1}{2}\lambda[X : \mathfrak{C}(X) = 1 - X_i] \\ &= \frac{1}{2}\sum_{k=1}^{k=n-1} \frac{2k - (n-1)}{n\,(n+1)} = \frac{(n-1)}{2n\,(n+1)},\end{aligned}$$

and therefore formula (11) is still valid. The same results can be shown to hold when $q = n$, by noting that in that case, $\alpha(k,n) = 0$ for all $k \in [0, n-1]$ and $\beta(k,n) = \binom{n-1}{k}$ for all $k \in [0, n-2]$, $\beta(n-1,n) = 0$.

Consider last the case of a constant mechanism ($q = 0$ or $q = n+1$). It is straightforward to check directly that $\lambda[X : \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i] - \lambda[X : \mathfrak{C}(X) = 1 - X_i] = 0$. Indeed, when for example $q = 0$, $\lambda[X : \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i] = \lambda[X : X_i = 1] = \frac{1}{2}$ and $\lambda[X : \mathfrak{C}(X) = 1 - X_i] = \lambda[X : X_i = 0] = \frac{1}{2}$.

To summarize, we therefore get for the correction term when $\lambda = IAC$:

$$\frac{1}{2}\begin{pmatrix} \lambda[X : \mathfrak{C}(X_{-i} \cup \{1 - X_i\}) = X_i] \\ -\lambda[X : \mathfrak{C}(X) = 1 - X_i] \end{pmatrix} = \begin{cases} -\frac{1}{2n} + \frac{(n+1-q)q}{n(n+1)} \text{ for } 1 \le q \le n, \\ 0 \text{ for } q \in \{0, n+1\}. \end{cases}$$

Combining the pivot term and the correction term, one gets that for all $q$, $0 \le q \le n+1$:

$$Utility\ (i, IAC, q) = \frac{1}{2} + \frac{(n+1-q)\,q}{n(n+1)},$$

where which can be decomposed for $1 \le q \le n$ as

$$Utility\ (i, IAC, q) = \frac{1}{2} + \frac{1}{2}\left(\frac{1}{n}\right) + \frac{1}{2}\left[-\frac{1}{n} + \frac{2}{n}\frac{(n+1-q)\,q}{n+1}\right],$$

where the second term in the sum on the right-hand side is the influence term and the third term is the correction term; when $q \in \{0, n+1\}$, both the influence and the correction term are zero. This shows that all (non-constant) anonymous mechanisms yield the same power $(\frac{1}{n})$, but that yield different expected utilities.

Denoting $q$ as $\lceil \theta n \rceil$ where $\theta \in \,]0,1[$ and $\lceil x \rceil$ is the smallest integer strictly greater than $x$, the correction term is approximately, for a large electorate:

$$-\frac{1}{2}\frac{1}{n} + \frac{(n+1-\theta n)\,\theta n}{n\,(n+1)} \simeq -\frac{1}{2}\left(\frac{1}{n}\right) + \theta\left[(1-\theta) + \frac{\theta}{n}\right] = \theta(1-\theta) + \left(\theta^2 - \frac{1}{2}\right)\frac{1}{n}.$$

Note that for a large electorate, the correction term is much larger than the pivot term ($\frac{1}{2n}$), since the order of magnitude is $\theta(1-\theta)$. For a large electorate:

$$Utility\ (i, IAC, q) \simeq \frac{1}{2} + \theta(1-\theta) + \frac{\theta^2}{n} \text{ where } q = \lceil \theta n \rceil. \tag{12}$$

While simple, let us highlight the fact that:

$$Utility\ (i, IAC, q) - Utility\ (i, IAC, \text{Pure Randomization}) \simeq \theta(1-\theta) + \frac{\theta^2}{n} \text{ where } q = \lceil \theta n \rceil,$$

i.e., that the per capita gain of optimization compared to randomization is 0(1).

### 4.3 Comparison with the IC setting

By definition, under $IC$ the correction term is absent. Using the same notations as in the above subsections, and using the fact that when $\lambda$ is the $IC$ electorate, the probability that out of the $n-1$ individuals in $N\backslash\{i\}$, some $k$ predetermined voters vote for $A$ is the same independently of the identity of these $k$ voters, and this probability is $\frac{1}{2^{n-1}}$, we obtain that for any voting mechanism $\mathfrak{C}$:

$$Influence\ (i, IC, \mathfrak{C}) = \sum_{k=0}^{k=n-1} \frac{\gamma_i(k, \mathfrak{C})}{2^{n-1}}.$$

Not much can be said on the probability of being pivotal at this level of generality besides the following inequality known as the edge-isoperimetric inequality[15]:

$$Total\ Influence\ (IC, \mathfrak{C}) \geq 2\lambda\left(\mathfrak{C}\right) Log_2 \left(\frac{1}{\lambda\left(\mathfrak{C}\right)}\right),$$

where we have defined $\lambda\left(\mathfrak{C}\right)$ as $\lambda\left(\mathfrak{C}\right) = \lambda\left[X : \mathfrak{C}(X) = 1\right]$ (see section 2). If the mechanism is neutral, then $\lambda\left(\mathfrak{C}\right) = \frac{1}{2}$ and the edge-isoperimetric inequality implies that $Total\ Influence$ $(IC, \mathfrak{C}) \geq 1$. Equality holds when $\mathfrak{C} = \mathfrak{Dic}$.

We are not aware of any characterization of the influence ranking $\gtrsim_{\lambda}^{TI}$ when $\lambda = IC$. Let us focus on the subclass of anonymous mechanisms, and still denote by $q$ the quota above which option $A$ is elected. Given the values for $\gamma_i(k, q)$ (see (10)), we obtain that:

$$Influence\ (i, IC, q) = \frac{1}{2^{n-1}} \binom{n-1}{q-1} \text{ for } q \notin \{0, n+1\}.$$

If we assume that $n$ and $q$ are large, then from Stirling's formula, we derive:

[15] See Kalai and Srafa (2006).

$$Influence\ (i, IC, q) \simeq \frac{1}{2^{n-1}} \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n-1}{(q-1)(n-q)}} \frac{(n-1)^n}{(q-1)^{q-1}(n-q)^{n-q}}.$$

Denoting $q$ as $\lceil \theta n \rceil$ where $\theta \in ]0,1[$ and $\lceil x \rceil$ is the smallest integer strictly greater than $x$, the above formula writes as:

$$Influence\ (i, IC, q) \simeq \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{1}{\theta(1-\theta)}} \frac{1}{(2\theta^\theta(1-\theta)^{1-\theta})^n}.$$

When $\theta \neq \frac{1}{2}$, $\frac{1}{(2\theta^\theta(1-\theta)^{1-\theta})^n}$ behaves as $e^{-\gamma n}$, where $\gamma = \ln\left(2\theta^\theta(1-\theta)^{1-\theta}\right)$ is a positive constant (since the function $\theta \ln \theta + (1-\theta)\ln(1-\theta)$ has a maximal value of $-\ln 2$ (on the interval $[0,1]$ ) at $\theta = \frac{1}{2}$). In contrast, when $\theta = \frac{1}{2}$, we obtain:

$$Influence\ \left(i, IC, \frac{n}{2}\right) \simeq \sqrt{\frac{2}{\pi n}}.$$

From Penrose's formula, we deduce:

$$Utility\ (i, IC, q) \simeq \begin{cases} \frac{1}{2} + \sqrt{\frac{1}{2\pi n}} \text{ if } q = \lceil \frac{n}{2} \rceil, \\ \frac{1}{2} + \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{1}{\theta(1-\theta)}} e^{-\gamma n} \text{ if } q = \lceil \theta n \rceil \text{ with } \theta \neq \frac{1}{2}. \end{cases} \quad (13)$$

It is important to note from (13) that if $\lambda$ is $IC$, then the social improvement over pure randomization of any non constant anonymous mechanism $q$, *Total Utility* $(IC, q)$ – *Total Utility* ($IC$,Pure Randomization), *is* $0(\sqrt{n})$.

More generally, take any voting mechanism $\mathfrak{C}$. Using the fact that for all $\lambda$, the best mechanism according to $\gtrsim_\lambda^{TU}$is the ordinary majority mechanism $\mathfrak{Maj}$, one gets that

$$\begin{aligned} & Total\ Utility\ (IC, \mathfrak{C}) - Total\ Utility\ (IC, \text{Pure Randomization}) \\ \leq\ & Total\ Utility\ (IC, \mathfrak{Maj}) - Total\ Utility\ (IC, \text{Pure Randomization}) \end{aligned}$$

and for any mechanisms $\mathfrak{C}$, the social improvement over pure randomization is $0(\sqrt{n})$: the per capita gain of optimization compared to randomization is of second order with respect to the size of the population.

We have seen in the previous subsection that the picture looks very different when the random electorate is not IC: in particular, if $\lambda$ is $IAC$, then there are mechanisms $\mathfrak{C}$ such that the per capita improvement over pure randomization *is* $0(1)$.

In an environment with correlation, the mechanism design exercise is useful as it has a first order effect when compared to pure randomization. In such case, any departure from the optimal mechanism has an impact. In contrast, in the $IC$ case, it has only second order effects.

## 5 Concluding Remarks

This note has offered some insights on the differences between two popular social objective functions in the area of voting mechanism design: $Total\ Influence$ and $Total\ Utility$. In the case of the $IC$ model, the Penrose formula shows that the two social objective functions coincide. While conclusions obtained with the $IC$ model are insightful, this note show that we should be careful about evaluating the qualities of alternative social mechanisms and deriving the optimal one on the sole basis of the $IC$ model. Indeed, with the Penrose formula, we have seen that in the $IC$ model, the exercise amounts to comparing second order effects among themselves and that the gain with respect to randomness is questionable. Second, and more importantly, we have shown that as soon as we introduce correlations across individual preferences, the correction term dominates the pivot term in the evaluation of social utility.

Of course many important questions remains to be solved.

First, can we derive a full characterization of the two social objectives attached to the orderings $\succsim_{\lambda}^{TI}$ and $\succsim_{\lambda}^{TU}$ for a large class of random electorates $\lambda$? Is it easy to recognize for any pair of mechanisms and any of these two orderings which mechanism in the pair is the best one?

Second, what happens if we depart from the assumption that the individual utilities can only take two values, assumed to be the same across all individuals. In real world applications where, often, the players are not individuals but population of individuals (regions, countries,...) this assumption is not relevant and should be replaced by other assumptions. This is what is done in Barbera and Jackson (2004) and Beisbart, Bovens and Hartman (2005). In such a context, it is no longer true that the best mechanism according to $\succsim_{\lambda}^{TU}$is the ordinary majority mechanism. The characterization of the top element of $\succsim_{\lambda}^{TU}$ for a large class of independent random electorates is solved in Barbera and Jackson, and Beisbart *et al.*.

Third, optimization over $\mathfrak{C}$ may be subject to constraints. The designer task is then to find the best mechanism (according to any of the two objectives studied above), under some set of constraints. One classical set of constraints consists in imposing to the mechanism to be indirect or two-tier on the basis of a given partition of the individuals into areas (countries, regions, districts,..): this means that the mechanism defines first for each of this group a representative or temporary winner (A or B) for the group based upon the preferences in the group and then a second mechanism to map the profile of temporary group winners into the ultimate choice. In this literature, the first mechanism is often postulated (for instance it is assumed to be the ordinary majority mechanism) and the designer only optimizes with respect to the second mechanism. The contributions of Barbera and Jackson (2004) and Beisbart, Bovens and Hartman (2005) can also be interpreted in such way: derivation of the optimal second best mechanism. Given the second best nature of the problem, the ordinary majority mechanism (direct democracy) is not a feasible choice.[16]

# 6 References


Banzhaf, J.F. (1965) "Weighted Voting does not Work : A Mathematical Analysis", *Rutgers Law Review*, 19, 317-343.

Barbera, 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.

Beisbart, C. and L. Bovens (2013) "Minimizing the threat of a positive majority deficit in two-tier voting systems with equipopulous units", Public Choice, 145, 75-94.

Beisbart, C., Bovens L, and S. Hartmann (2005) "A Utilitarian Assessment of Alternative Decision Rules in the Council of Ministers", *European Union Politics*, 6, 395–418.

Brams, S.J. and M. Lake (1978) "Power and Satisfaction in a Representative Democracy" in *Game Theory and Political Science*, Ordeshook (Ed), New York.

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

Curtis, R.B. (1972) "Decision Rules and Collective Values in Constitutional Choice" in *Probability Models of Collective Decision Making*, Niemi, R. and H. Weisberg (Eds), Merrill, Columbus, Ohio.

Dubey, P. and L.S. Shapley (1979) "Mathematical Properties of the Banzhaf Power Index", *Mathematics of Operations Research*, 4, 99-131.

Feix, M., Lepelley, D., Merlin, V. and J.L. Rouet (2004) "The probability of conflicts in a U.S. presidential type election," *Economic Theory,* 23, 227-257.

Felsenthal, D.S. and M. Machover (1998) *The Measurement of Voting Power*, Edward Edgar, Northampton.

Felsenthal, D. S. and Machover, M. (1999) Minimizing the mean majority deficit : the second square-root rule. *Mathematical Social Sciences,* 37, 25-37.


[16]This literature contains several results under $IC$. When the utilities takes two values, it has been demonstrated by Felsentahl and Machover (1999) that the maximal $\succsim_{\lambda}^{TU}$ mechanism is a weighted majority mechanism where each group receives a weight proportional to the square root of its population. Then, if groups are equipopulous, the optimal second tier mechanism is the ordinary majority mechanism. It has been further demonstrated by Beisbart and Bovens (2013) that when the groups are equipopulous, *Total Utility* $(\lambda, \mathfrak{C})$ where $\mathfrak{C}$ is the majority two-tier mechanism reaches it minimum when the partition is such that the population in each group is about the square root of the total population.

Fishburn, P.C. and Gehrlein, W.V. (1976) “Borda’s Rule, Positional Voting, and Condorcet’s Simple Majority Principle”, *Public Choice*, 28, 79-88.

Good, I.J. and L.S. Mayer (1975) “Estimating the Efficacy of a Vote”, *Behavioral Science*, 20, 25 -33.

Kalai, G. and S. Safra (2006) ”Threshold Phenomena and Influence with Some Perspectives from Mathematics, Computer Science, and Economics”, *Computational Complexity and Statistical Physics*, Percus, A., Istrate G., and C.Moore (Eds), Santa Fe Institute Studies on the Sciences of Complexity, Oxford University Press.

Kuga, K. and H. Nagatani (1974) “Voter Antagonism and the Paradox of Voting”, *Econometrica*, 42, 1045-1067.

Myerson, R.B.(1998) “Population Uncertainty and Poisson Games”, *International Journal of Game Theory*, 27, 375-392

Myerson, R.B.(2000) “Large Poisson Games”, *Journal of Economic Theory*, 94, 7-45.

Penrose, L.S. (1946) ”The Elementary Statistics of Majority Voting”, *Journal of the Royal Statistical Society*, 109, 53-57.

Rae, D.W. (1969) ”Decision Rules and Individual Values in Constitutional Choice”, *American Political Science Review*, 63, 40-56.

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. (1978) ”Probability Models for Power Indices” in *Game Theory and Political Science*, Ordeshook (Ed), New York.

Taylor, A.D. and W.S. Zwicker (1999) *Simple Games*, Princeton University Press, New Jersey.

Weber, R.J (1995) “Approval Voting”, *Journal of Economic Perspectives*, 9, 39-49.

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