

## Discussion topic: voting power when voters' independence is not assumed

This note outlines the conceptual problem of measuring voting power when voters cannot be assumed to act independently of one another – as is the case where a posteriori or actual (as opposed to a priori) voting power is concerned.

By ‘voting power’ we mean here what has been called ‘I-power’, which is supposed to quantify the influence of a voter over the outcomes of divisions of the voters (rather than a voter’s expected payoff).

We assume a given decision rule (aka ‘simple voting game’)  $\mathcal{W}$  with assembly (set of voters)  $N$ . We put  $n =: |N|$ .

By a *division* of a set of voters we mean its partition into two subsets: a *positive* (‘yes’) *camp* and a *negative* (‘no’) *camp*.

$\mathcal{W}$  is a mapping of the set of all  $2^n$  divisions of  $N$  to the set of two possible outcomes: *positive* (the proposed bill is passed) and *negative* (the proposed bill is blocked) – subject to the well-known monotonicity condition.

We also assume a probability distribution  $P$  on the set of all  $2^n$  divisions of  $N$ .  $P$  induces, in an obvious way, a probability distribution on the set of all divisions of any  $S \subseteq N$ .

For any  $a \in N$ :

- we denote by  $p_a$  the probability of the event that  $a$  votes ‘yes’: ie, belongs to the positive camp of a division of  $N$ ;
- we denote by  $\psi_a$  the probability of the event that  $a$  is *critical*: ie, that the set of all other voters,  $N - \{a\}$ , is so divided, that the outcome under  $\mathcal{W}$  is positive or negative according as  $a$  joins the positive or negative camp of  $N - \{a\}$ ;
- we denote by  $\rho_a$  the probability of the event that  $a$  is *successful*: ie, that  $N$  is so divided, that the outcome under  $\mathcal{W}$  has the same sign as the camp of  $N$  containing  $a$ ;
- we say that  $a$  is *independent* if the way  $a$  votes is probabilistically independent of any division  $D$  of  $N - \{a\}$ ; ie, for every such  $D$ ,

$$P(a \in \text{positive camp of a division of } N \mid D) = p_a,$$

and

$$P(D \mid a \in \text{positive camp of a division of } N) = P(D).$$

In the absence of any information about the voters and the bills that are to be decided, or when any such information is deliberately hidden behind

the proverbial ‘veil of ignorance’, one assumes a priori that all voters are independent, and that  $p_x = \frac{1}{2}$  for all  $x \in N$ . The conjunction of these two assumptions is equivalent to assuming that  $P$  is uniform: all  $2^n$  divisions of  $N$  are equiprobable.

Under this a priori assumption, Penrose proposed  $\psi_a$  as the measure of the voting power of  $a$ . (This is also known as the *absolute Banzhaf index* of a priori voting power.)

Moreover, if  $a$  is independent and  $p_a = \frac{1}{2}$ , then it is easy to prove the following identity (due to Penrose):<sup>1</sup>

$$\psi_a = 2\rho_a - 1. \quad (1)$$

Of course – as pointed out by Laruelle and Valenciano<sup>2</sup> – the above definition of  $\psi_a$  makes formal sense for arbitrary probability distribution  $P$ . However, the question is whether it is always reasonable to regard  $\psi_a$  as a measure of the influence of  $a$  on the outcome of divisions.

The answer is certainly affirmative under the classical a priori assumption of uniform  $P$ . Moreover, a good case can be made also for extending this affirmative answer to any  $P$  for which  $a$  is independent. An argument in support of this is that if  $a$  is independent then

$$\psi_a = \frac{\partial A}{\partial p_a}, \quad (2)$$

where  $A$  is the probability of the event that the outcome of a division of  $N$  is positive (which Coleman called ‘the ability of the collectivity to act’).<sup>3</sup> Thus  $\psi_a$  is the marginal contribution of the propensity of  $a$  to vote for a bill to the probability of the bill’s passing.

Indeed, in their well-known 1979 paper, ‘Mathematical properties of the Banzhaf power index’ (*Mathematics of Operations Research*, **4** pp. 99–131) Dubey and Shapley generalize the Penrose measure  $\psi$  (which they call the ‘Banzhaf probability index’ and denote by  $\beta'$ ) to the case where *all* voters are independent, but the  $p_x$  ( $x \in N$ ) are arbitrary (and not necessarily all equal). Under this assumption, the expression they give for  $\psi_a$  – *ibid.* p. 122, equations (47) and (48) – coincides with our definition of  $\psi_a$  stated above.

---

<sup>1</sup>Note that (1) does not require any assumption regarding the values of  $p_x$  or the independence of  $x$  for any  $x \in N$  other than  $a$ .

<sup>2</sup>See, for example, ‘Assessing success and decisiveness in voting situations’, *Social Choice and Welfare*, 2005, **24**:171–197.

<sup>3</sup>This follows at once from the fact that  $A = \psi_a p_a + B$ , where  $B$  is the probability of the event that  $N - \{a\}$  is so divided, that the outcome is positive irrespective of how  $a$  votes. Note that (2) does not require any assumption regarding the independence of any  $x \in N$  other than  $a$ .

However, things become quite problematic when the assumption of voters' independence is dropped. It is generally recognized that when we are concerned with a posteriori or actual (rather than a priori) voting power, one can no longer assume that all the  $p_x$  ( $x \in N$ ) are equal to one another. *But surely also the assumption that all (or indeed any) voters are independent cannot be taken for granted.*

Here it may be useful to distinguish between a posteriori and actual voting power. We refer to *a posteriori voting power* where the probability distribution  $P$  is obtained empirically from statistics of how the set of voters  $N$  divided on (sufficiently many) past occasions. On the other hand, in addressing *actual voting power*, several authors represent voters' preferences as 'ideal points', which are in general random variables in some Euclidean state space. The proposed bill as well as the status quo are also represented in a similar way in this space.<sup>4</sup> The power of a given voter  $a$  is then assessed by looking at the marginal effect of a shift of the preferences of  $a$  on the outcome of a division.

In what follows, we shall concentrate on the a posteriori approach. But it seems to us that the actual-power approach presents analogous problems.

In the absence of independence,  $\psi$  behaves in a strange way, which indicates that it cannot be taken as a reasonable measure of a voter's influence over the outcomes. At least,  $\psi$  doesn't tell the whole story about that influence.

As a simple example, consider  $\mathcal{W} = \mathcal{M}_5$ ; ie, the canonical simple majority decision rule with an assembly of 5 voters:  $I_5 = \{1, 2, 3, 4, 5\}$ . Let  $P$  be the probability distribution that assigns probability 0 to the 20 divisions of  $I_5$  in which the positive camp contains exactly two or exactly three voters; and equal probability of  $\frac{1}{12}$  to each of the remaining 12 divisions. Then no voter is ever critical, so  $\psi_i = 0$  for all  $i \in I_5$ . Yet it would be absurd to claim that the voters here are powerless, in the sense of having no influence over the outcome of divisions.

One may be tempted to base the measurement of voting power on  $\rho$  rather than  $\psi$ . In the classical a priori mode, this is essentially equivalent to using  $\psi$ . Indeed, that was Penrose's approach: he started with  $\rho$ , and arrived at  $\psi$  by way of (1). However, (1) need not hold when  $p_a \neq \frac{1}{2}$  or  $a$  is not independent, so  $\rho$  provides a distinct approach to measuring power.

But this approach also does not seem to work in the absence of independence. To take an extreme example, suppose the voting behaviour of a voter who is a priori very weak – say a dummy – is highly correlated with that

---

<sup>4</sup>We leave aside the thorny but quite separate question as to how sufficiently accurate information for such representation may be obtained.

of a voter who is a priori very strong – say a dictator. The two voters may just happen to have very similar preferences. Then the two will have similar values of  $\rho$ . Surely, this cannot mean that they have similar influence.

All this suggests that the ways of measuring voting power of independent voters are inadequate when independence cannot be assumed. This may be because the very concept of power that is used in the a priori mode is in some sense too individualistic. Thus we are led to raise questions that are not merely technical but also conceptual and perhaps philosophical.