Measuring Influence for Dependent Voters: A Generalization of the Banzhaf Measure

Luc Bovens (LSE London, U.K.) and Claus Beisbart (Technische Universität Dortmund, Germany)
LOFT Preproceedings
2008

Abstract: We construct a new measure of voting power that is a reasonable measure of influence even if the votes are not cast independently. The crucial building blocks of our measure are probabilities of counterfactuals, such as the probability that the outcome of the vote would have been yes, had a voter voted yes rather than no as she did in the real world. The probabilities of such counterfactuals are calculated on the base of causal information, following the approach by Balke and Pearl. Opinion leaders who have causal influence on other people's votes can have significantly more voting power under our measure than the latter. We provide several examples in which our measure yields intuitively plausible results and show that the measure reduces to Banzhaf voting power in the limiting case of independent and equiprobable votes.

1. Introduction

The Banzhaf measure of voting power quantifies the degree to which political agents can influence the outcome of a vote (Felsenthal and Machover 1998, p. 36). It equals the chance of pivotality in a binary vote under the assumptions that (i) the voters cast their votes independently and (ii) each voter is equally likely to vote yes or no. These probabilistic assumptions define the so-called Bernoulli model (Felsenthal and Machover 1998, Def. 3.1.1, p. 37). However, one might argue, if sufficient empirical data on past voting patterns are available, then one should base the measurement of influence on a probability model that fits the data, and of course the Bernoulli model will seldom be a fitting model (see Gelman et al. 2002 and Gelman et al. 2004 for a related discussion). A simple suggestion to generalize Banzhaf voting power for such cases would be to calculate the probability of being pivotal under the appropriate probability model—whatever it is like. This suggestion has been put forward by e.g. Morriss (1987/2002), p. 169. But unfortunately this suggestion does not work if the independence assumption does not hold. It leads to counterintuitive assessments of political influence. The following example is a simplification of what has come to be known as Wilmers' example in Machover (2007). Five voters take decisions under simple majority vote. The 5-0 split and the 0-5 split have a probability of one half each. The other voting profiles have probability zero. Clearly, the probability of being pivotal is zero for each voter. Accordingly, under the simple suggestion, no voter would have political influence in the voting process. But, intuitively, it seems odd to say that no voter has any influence on the outcome of the vote (*ibid.*, p. 3). A different measure of voting power is needed.

How can we construct such a measure? In the example above, the obvious question to ask is: How did the probability distribution arise? Suppose, first, that voter A is an opinion leader, i.e. if A votes yes (or no), then all other voters simply copy her vote. What one would say intuitively is that A has more voting power than the other voters—if A had voted differently, then the others would have followed suit and the outcome would have been different. But we cannot say the same for the other voters. However, if, second, a different voter B is an opinion leader, clearly B should have more voting power than the other voters. Or, third, if the voting pattern comes about only due to political views shared by the voters, then we expect equal voting power in the example. Thus, in order to provide an assessment of voting power, one should take into account the *causal relations* between the votes that bring about the distribution over profiles.

We will go beyond the simple suggestion and develop a measure of voting power that does not clash with our intuitions because it is sensitive to causal information. The measure relies on causal models and counterfactual probability distributions. We define the measure in section 2. In section 3, we develop a simple example to show how the measure deals with opinion leaders, shared political views and dictatorial voting procedures. In section 4, we relate the measure to other measures; we show that the measure reduces to Banzhaf voting power under the Bernoulli model. Section 5 provides an analysis of Wilmers' example (Machover 2007). Section 6 concludes, and technical details are provided in the appendix.

Since we work with causal models and probabilities, one may distinguish between causal dependence and stochastic dependence. If not stated differently, "(in)dependence" means stochastic (in)dependence. If two random variables are stochastically independent, there can not be a causal dependence, either.

Recently, Laruelle and Valenciano (2005) have suggested a general framework in which various voting power measures – the Banzhaf measure as well as other measures – can be embedded. The framework has two layers: first, a voting rule that is modelled as a simple voting game (p. 174); second, a probability model over profiles (pp. 175–6). Various measures of voting power turn out to be unconditional or conditional probabilities of either success or decisiveness in this framework. In Laruelle and Valenciano's terms, we are interested in decisiveness (or, in our terms, pivotality) rather than success. But the measure that we propose cannot be fitted into their framework. In the general case, it is neither an unconditional nor a conditional probability. Rather, our measure is a weighted sum of differences between certain probabilities. Roughly, what we are interested in is the extent to which probabilities of acceptance (respectively rejection) change, if a voter were to switch her vote from no to yes (respectively from yes to no). These changes are traced on the basis of a causal model. We will discuss the connection between our work and Laruelle and Valenciano (2005) in section 4.

2. **Generalized Voting Power**

The voting power of a voter *i* is the extent to which she is able to make a difference as to whether a bill passes or not (Felsenthal and Machover 1998, p. 36). How can that extent be measured? One way to measure it is to calculate the probability that that voter's vote is pivotal (cf. ibid.). But this proposal runs into the problem plain from Wilmers' example. So an alternative is required.

Suppose that with person *i* voting no on a proposal in the actual world, the chance of acceptance is quite low. But if *i* were to vote yes instead, then the chance of acceptance would be much greater. Similarly, suppose, with *i* voting yes on a proposal in the actual world, the chance of rejection is quite low. But if *i* were to vote no instead, then the chance of rejection would be much greater. Under these assumptions, intuitively, *i* has more influence than a person *j* for whom these respective effects would be notably smaller. The idea is thus to take differences of chances in order to measure voting power. Let us make this idea precise by constructing a measure.

We introduce the following notation. We assume that there are N voters. The i^{th} vote is modeled as a random variable V_i . We set $V_i = 1$, if *i* votes yes, and $V_i = 0$, if she votes no. The outcome of the vote is described by the random variable V. $V = 1$ means that the proposal is accepted. $V = 0$ means that the proposal is rejected.

We assume that we have a full probability model for the votes. The model provides us with probabilities for each possible voting profile (i.e. the joint probabilities of the V_i s). We assume for now that all conditional probabilities such as $P(V_j = 0| V_i = 1)$ are defined. (In Section 5, we will consider a case in which this assumption is violated.) Once we know the decision rule, we can calculate conditional probabilities for acceptance given some single vote or given a combination of votes, such as $P(V = 1|\ V_i = 1)$ or $P(V = 1|\ V_i = 1, V_j = 0)$. It is not our concern in this paper how to obtain a realistic probability model from empirical data.

We will also assume that we have causal information about the votes. In our models a person's votes can be influenced by other people's votes and by her political views. This is not inconsistent with taking voting to be an instance of free agency. How the related causal information can be obtained is not our concern in this paper.

For calculating our measure of influence, we first assess the chance that a proposal is accepted given that *i* voted no, i.e. $P(V = 1|\ V_i = 0)$ and the chance that a proposal is rejected given that *i* voted yes, i.e. $P(V = 0|\ V_i = 1)$. Subsequently, we construct probability distributions over certain counterfactuals. We ask what the chance is that a proposal would have been accepted, had *i* voted yes (rather than no, as *i* did in the actual world), call it $Q_i^0(V = 1|\ V_i = 1)$. And we ask what the chance is that a proposal would have been rejected, had *i* voted no (rather than yes, as *i* did in the actual world), i.e. $Q_i^1(V = 0|\ V_i = 0)$.

These chances are calculated following the approach by Balke and Pearl (1994). The basic idea is very simple. For calculating $Q_i^0(V = 1|\ V_i = 1)$, e.g., we first assume that *i* votes no, as she does in the actual world. We infer the probabilities of the other votes being one way or another. Then *i*'s vote is switched to yes. We trace the causal effects of *i*'s voting yes and recalculate the probabilities that the votes that are causally affected by *i*'s vote are one way or another. Finally, the probability of acceptance is calculated on this base. A general algorithm for calculating the probabilities for counterfactuals is given in the Appendix.

Let D_i^0 be the difference between the chance that the proposal would have been accepted had *i* voted yes and the chance that the proposal is accepted conditional on *i* having voted no:

(1) $$D_i^0 = Q_i^0(V = 1|\ V_i = 1) - P(V = 1|\ V_i = 0)\ .$$

Let D_i^1 be the difference between the chance that the proposal would have been rejected had *i* voted no and the chance that the proposal is rejected conditional on *i* having voted yes:

(2) $$D_i^1 = Q_i^1(V = 0|V_i = 0) - P(V = 0|\ V_i = 1)\ .$$

Because $Q_i^1(V = 0|V_i = 0) = 1 - Q_i^1(V = 1|V_i = 0)$ and $P(V = 0|\ V_i = 1) = 1 - P(V = 1|\ V_i = 1)$, D_i^1 can also be written as follows:

(2') $$D_i^1 = P(V = 1|\ V_i = 1) - Q_i^1(V = 1|V_i = 0)\ .$$

We can now construct the measure:

(3) $$D_i = D_i^0\ P(V_i = 0) + D_i^1\ P(V_i = 1)\ .$$

It can be shown that the *D* is always in the interval [0,1].

D is a measure of the impact of a person's vote that takes into account the causal relationships between the various people casting their votes. This measure is analogous to the measure of the *average treatment effect* in epidemiology and causal analysis more generally. Suppose that there is a particular treatment for a particular disease. We want to quantify the potential impact of that treatment on survival in the subpopulation of people with that disease. As a matter of fact, some people chose to take the treatment, whereas others did not. It is no use comparing the survival rates between people who chose the treatment and people who did not, since choosing the treatment may be connected with social strata membership, which, in turn, may influence the chances of recovering from the disease. Rather, we can quantify the impact of the treatment on survival in the subpopulation as follows. First, we try to assess how the chance of survival among people who did not undergo the treatment would have increased had they undergone the treatment. To do so, we subtract the probability that a person in the non-treatment group survived from the probability of the counterfactual that she would have survived had she undergone the treatment. We expect that, typically, the result is positive. Second, we try to assess how the chance of survival among people who underwent the treatment increased compared to the counterfactual case in which they had not undergone the treatment. To do so, we subtract the probability of the counterfactual that a person in the treatment group would have survived had she not undergone the treatment from the probability that she did survive. Again, we expect a positive difference in the typical case. The impact of the treatment in the subpopulation is measured by the sum of these differences, weighted respectively by the proportion of people who underwent the treatment and the proportion of people who did not undergo the treatment. This measurement is discussed by Winship and Morgan (1999: 665, Eq. 4) and Morgan and Winship (2007: 45, Eq. 2.8).

Consider now voting again. We want to measure the influence of a person's vote on the outcome given a particular voting procedure within a particular society. As a person might survive or not as an outcome, the proposal to be voted on might be accepted or rejected as an outcome. And just as a person might have undergone the treatment or not, a person might vote yes or vote no. Let us use this analogy in order to understand our *D*-measure more closely.

First, typically, for people who did not undergo the treatment, their chance of survival would have increased had they undergone the treatment. Similarly, given that a person votes no, typically, the chance of acceptance would increase were she to vote yes. The extent to which it increases, is the difference in probabilities, $Q_i^0(V = 1|\ V_i = 1) - P(V = 1|\ V_i = 0)$, which equals D_i^0. Second, typically, for people who underwent the treatment, their chance of survival is greater than in the counterfactual case in which they had not undergone treatment. Similarly, given that a person votes yes, the chance of acceptance is greater than in the counterfactual case in which she were to vote no. The extent to which it is greater, is $P(V = 1|\ V_i = 1) - Q_i^1(V = 1|V_i = 0)$ or D_i^1 (according to Eq. 2'). The impact of the treatment on survival of people in the subpopulation who have the disease is measured by the weighted sum of these differences. Similarly, the influence of a person's vote in a particular society on acceptance is measured by the sum of these differences weighted by the chance of her voting no and the chance of her voting yes, respectively – which precisely yields our *D*-measure. So one can think of the *D*-measure as the average treatment effect of a vote on the outcome of the vote. We will show that this measure can handle Wilmers' example and reduces to the Banzhaf measure under suitable assumptions.

The definition of the *D*-measure need not only be motivated by this analogy. There is also an ex post justification for quantifying influence in terms of *D*: *D* yields intuitively plausible results if applied to a range of examples. We now turn to such examples.

3. **A simple example**

We start with a simple three-person example featuring a Supreme Court with Scalia, Thomas and Ginsburg as justices.[1] This example will illustrate how the *D*-values are calculated. We will first assume that proposals are decided on by a simple majority vote. Rather than using the terminology that is fitting for the Supreme Court, we will conduct our presentation in terms of *voters* and *proposals*.

We consider different models for the votes. Under every model, each voter votes yes with a probability of .5.

[1] These names are just mnemonic aids. For actual correlation coefficients between Supreme Courts votes, see Kaniovski and Leech (2007), Table 3.

3.a Opinion leader

Ginsburg's vote is stochastically independent from the other votes, and there are no causal relations between her and the others' votes. Thomas, however, keeps a close eye on Scalia and there is .9 chance that he will vote yes, given that Scalia votes yes, and there is a .9 chance that he will vote no, given that Scalia votes no. One can thus say that Thomas's vote causally depends on Scalia's vote. The causal influences in the model are represented in Fig. 1. Let us now assess the influence of each voter by calculating her *D*-value.

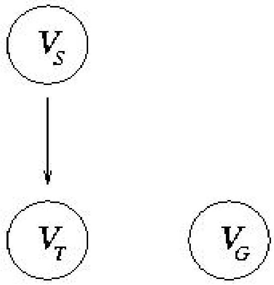

Figure 1. The causal network for the opinion leader model

We start with the *D*-value of Thomas. We consider the first addend, viz. $D_T^0 \times P(V_T = 0)$. Clearly, $P(V_T = 0) = .5$.

For calculating D_T^0 we have to assume that Thomas votes no in the real world. We first turn to $P(V = 1| V_T = 0)$. If Thomas votes no, then the conditional chance that Scalia voted no is .9. Given that Thomas votes no, we only get acceptance in the real world, if both Scalia and Ginsburg vote yes. And so the conditional chance that the proposal is accepted is $P(V = 1| V_T = 0) = .1 \times .5 = .05$. We now turn to $Q_T^0(V = 1| V_T = 1) = 1 – Q_T^0(V = 0| V_T = 1)$. If Thomas were to vote for the proposal, this would not affect the chance that Scalia voted yes—that chance is still .9—since the causal link does not go from Thomas to Scalia. The only profile under which the proposal would be rejected, if Thomas voted yes, is a profile with Scalia and Ginsburg voting no. That chance is $.9 \times .5 = .45$. So $Q_T^0(V = 1| V_T = 1) = 1 - .45 = .55$. Hence, $D_T^0 = .55 - .05 = .50$. The argument for D_T^1 runs parallel and so $D_T = .50$.

Let us now calculate the *D*-value for Scalia. We first consider Q_S^0. If Scalia votes no, then the chance that Thomas votes no is .9. Given that Scalia votes no, we only get acceptance in the real world, if both Thomas and Ginsburg vote yes. And so the conditional chance that the proposal is accepted is $P(V = 1| V_S = 0) = .1 \times .5 = .05$. We now turn to $Q_S^0(V = 1| V_S = 1) = 1 – Q_S^0(V = 0| V_S = 1)$. If Scalia were to vote yes, this would affect the chance that Thomas votes no—that chance is now .1—since the causal link goes from Scalia to Thomas. The only profile under which the proposal would be rejected is a profile with Thomas and Ginsburg voting no. The chance is $.1 \times .5 = .05$. So $Q_S^0(V = 1| V_T = 1) = .95$ and $D_S^0 = .9$. The argument for D_S^1 runs parallel and so $D_S = .9$.

Finally, we assess the influence of Ginsburg. If Ginsburg votes no, then this does not affect the chance that Thomas or Scalia vote no. Given that Ginsburg votes no, the chance that both Thomas and Scalia vote yes is $.9 \times .5 = .45$. So $P(V = 1| V_G = 0) = .45$. Suppose that Ginsburg asks herself what the chance would be that the motion had been accepted had she voted yes. The chance that both Thomas and Scalia would have voted no is .9 × .5 = .45. So $Q_G^0(V = 1| V_G = 1) = .55$. Hence, $D_G^0 = .1$. The argument for D_G^1 runs parallel and so $D_G = .1$.

This is not unreasonable. Scalia does have more influence than Thomas, because he takes Thomas along with him as he changes votes, but not vice versa. Ginsburg has very little influence because she faces a quasi-block vote of the two other voters.

Voter	Scalia	Thomas	Ginsburg
D-value	0.9	0.5	0.1

Table 1. Results under the opinion leader model

3.b Common causes

Let us now change our assumptions in the following way:

1. We introduce a parameter ε which permits us to vary the extent to which the votes of Scalia and Thomas are correlated (correlations always imply stochastic dependency). The parameter ranges from –1 for full negative correlations over 0 for independence to +1 for full positive correlation:

$$P(V_T = 1 \mid V_S = 1) = P(V_T = 0 \mid V_S = 0) = .5*(1+ \varepsilon) .$$

Hence,

$$P(V_T = 1 \mid V_S = 0) = P(V_T = 1 \mid V_S = 0) = .5*(1 - \varepsilon) .$$

2. Correlations do not arise due to a direct causal influence from Scalia to Thomas or vice versa as in Subsection 3.a. Rather, they are due to a common cause (see Figure 2). Positive correlations are due to shared political views. Negative correlations are due to diverging political views. We can model this in the following way. We introduce a random variable which captures the nature of the proposal (*cf*. Balke and Pearl 1994). If $C = 0$, then the proposal is such that both Scalia and Thomas vote no; if $C = 1$, then the proposal is such that Scalia votes no and Thomas votes yes; if $C = 2$, then the proposal is such that Scalia votes yes and Thomas votes no; if $C = 3$, then the proposal is such that both vote yes. C models, in Dretske's terms (Dretske 1988, pp. 42–4), a triggering cause for the voting behaviour of Thomas and Scalia. The features of the proposal trigger votes that match the political views of Thomas and Scalia. By specifying the probability values in Table 2 we can fix the degree to which Thomas and Scalia's votes are correlated or anti-correlated.

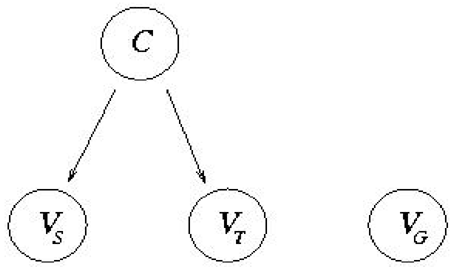

Figure 2. The causal network for the common cause model

$C = i$	$P(V_S = 1 \| C = j)$	$P(V_T = 1 \| C = j)$	$P(C = i)$
$C = 0$	$P(V_S = 1 \| C = 0) = 0$	$P(V_T = 1 \| C = 0) = 0$	$.25*(1+ \varepsilon)$
$C = 1$	$P(V_S = 1 \| C = 1) = 0$	$P(V_T = 1 \| C = 1) = 1$	$.25*(1 - \varepsilon)$
$C = 2$	$P(V_S = 1 \| C = 2) = 1$	$P(V_T = 1 \| C = 2) = 0$	$.25*(1 - \varepsilon)$
$C = 3$	$P(V_S = 1 \| C = 3) = 1$	$P(V_T = 1 \| C = 3) = 1$	$.25*(1+ \varepsilon)$

Table 2. The probability model for the common cause model. The variable V_G is not included – it is independent from the other variables V_i and takes values of 0 and 1 with a probability of .5 each

In the following, we will omit the details of our calculations. A general algorithm for calculating the Q_i^0s and Q_i^1s is provided in Appendix A.

Applying this algorithm, we obtain the following results:

Voter	Scalia	Thomas	Ginsburg
D-value	0.5	0.5	$.5 * (1 - \varepsilon)$

Table 3. Results under the common cause model

Thus, under the common cause model, the *D*-values for Scalia and Thomas do not depend on the strength of the correlations. On our model, the influence of a voter is the same regardless whether there is a shared political view with another voter or not. However, the influence of Ginsburg depends on whether and how the votes of the other voters are correlated. If Scalia and Thomas always vote the same, then Ginsburg has no influence. If, on the other hand, Scalia and Thomas always cast opposing votes, then Ginsburg has maximal influence. Clearly, if the votes of Scalia and Thomas are independent, then every voter has a *D*-value of .5. Note that this *D*-value for independent votes coincides with the Banzhaf measure for this simple voting game. As we will show in Section 4, this is due to a more general connection between Banzhaf voting power and the *D*-value.

But one might object that, intuitively, the influence of Scalia (and Thomas) is greater in a court in which their votes are correlated than in a court in which they vote independently. This is indeed not an unreasonable interpretation of influence. Let us call this the block interpretation. It is not the interpretation that is captured by the *D*-measure, though. On the *D*-measure, Scalia's influence is the same with correlated or independent votes, because if he had voted differently, then this would have had no effect on Thomas's vote in either case, and so the chance of rejection or acceptance would have been equally affected.[2] But we can also construct a measure D^* which is in line with the block interpretation of influence. Let D^* be the normalized measure of D, i.e. every *D*-value is divided by the sum of the *D*-values for all voters. Then D^* is monotonically increasing in ε and so according to this measure the influence of Scalia and Thomas increases as we move from anti-correlated votes over independence to correlated votes. D^* stands to D in the same way as the Banzhaf index β stands to the non-normalized Banzhaf measure β' (*cf*. Felsenthal and Machover 1998, Def. 3.2.2, p. 39).

3.c Causal influence on multiple votes

Let us now change our assumptions once more. Start with the opinion leader model from Section 3.a and suppose that Scalia's causal influence extends to Ginsburg's vote as well. This leads to a causal model as depicted in Fig.3. We define a family of probability models parameterized by ε. As in Section 3.b, ε can take values in the interval [0,1]. For $\varepsilon = 1$, there are full positive correlations between Thomas's (or Ginsburg's) and Scalia's votes. In this case, Scalia is a very influential opinion leader and his vote will be copied by any other voter in the court. For $\varepsilon = -1$, there are full negative correlations between Thomas's (or Ginsburg's) and Scalia's votes. For instance, both Thomas and Ginsburg might dislike Scalia and try to outvote him on every issue.

In mathematical terms, the model looks as follows: There is still equiprobability for yes and no votes in the marginal probabilities of the different voters. Given Scalia's votes, Thomas's and Ginsburg's votes are independent. Furthermore,

$$P(\, V_T = 1 \mid V_S = 1) = P(\, V_G = 1 \mid V_S = 1) = .5^*(1+ \varepsilon)$$

and

$$P(\, V_T = 0 \mid V_S = 0) = P(\, V_G = 0 \mid V_S = 0) = .5^*(1+ \varepsilon)\ .$$

This fixes the probability model.

[2]Following Lewis (1979), we take it to be the case that truth-value assignments to counterfactuals under a default interpretation do not permit backtracking.

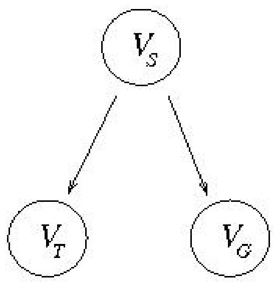

Figure 3. The causal network with causal influence on multiple votes

Results can be calculated following the method from Section 3.a. They are shown in Table 4. For $\varepsilon = 0$, the votes of the voters are independent and each voter's *D*-measure is .5. For full positive correlations ($\varepsilon = 1$), Scalia has a *D*-measure of 1, whereas the other have a zero *D*-measure. This is a plausible thing to say, since, in this case, Thomas and Ginsburg lack the ability to affect the outcome of the vote by unilaterally casting a different vote.

Voter	Scalia	Thomas	Ginsburg
D-value	$1 - .5 * (1 - \varepsilon)^2$	$.5 * (1 - \varepsilon)$	$.5 * (1 - \varepsilon)$

Table 4. Results for a model in which Scalia's vote causally influences both other votes.

For full negative correlations ($\varepsilon = -1$), we obtain: $D_S = -1$, $D_T = 1$, $D_G = 1$. The question thus arises what negative *D*-values mean.

Negative values of the *D*-measure mean that a voter is able to affect the outcome of a vote, but in an unexpected way. If she were to switch her vote from yes to no (from no to yes), this would make a difference, but it is not the probability of rejection (acceptance) that is increased, as one might expect, but rather the probability of acceptance (rejection).[3] We take it to be a strength of our measure that it captures the direction of the influence. If one is only interested in whether a voter is able to affect the outcome in any way, one can look at the absolute value of the *D*-measure. Thus, in our example, Scalia has in some way influence according to our *D*-measure, and this is as it should be.

Of course, this unexpected way of influencing the outcome of a vote invites strategic voting. If a person were to know that she has negative influence, then she could vote strategically and vote against her actual preferences. Suppose that her strategy is successful. One might then want to say that she has positive influence – she is able to get what she wants. One can turn this into an objection against our *D*-measure, since the *D*-measure is maximally negative. However, this objection does not work. For measuring influence, what a voter wants is not relevant. Our measure is not about the probability that a voter is able to get what she wants. Rather, the question is to what extent a voter is able to influence the outcome of a vote by switching her vote.

Let us now turn to the *D*-measure of Thomas and Ginsburg – which is 1. At first sight, this seems strange. The votes of Thomas and Ginsburg are fully causally determined by the vote of Scalia, but

[3] Since the *D*-measure is the sum of two addends, it is sufficient for a voter's *D*-measure to be negative that both possible switches change the probability of acceptance in an unexpected way, and it is necessary that one of the switches changes this probability in an unexpected way.

still have a maximal D-value. But on second reflection this is as it should be. For our measure is about what voters *could* do, not about what they do as a matter of fact. As a matter of fact, Thomas and Ginsburg will always vote differently from Scalia. But what matters for our measure is that each of them *could* always switch the outcome single-handedly, if he or she were to vote differently. We conclude that our measure deals with the example appropriately.

3.d Dictator

So far we have assumed simple majority voting with equal weights. Let us now change the weights as follows: Scalia has a block vote of three votes, whereas Thomas and Ginsburg have only one vote each. The Supreme Court issues a yes (no) vote if and only if there are at least three yes (no) votes. Thus, Scalia is a dictator, whereas Thomas and Ginsburg are dummies (*cf*. Def. 2.3.4 on p. 24 in Felsenthal and Machover 1998).

Whether we calculate these results under the model under which Scalia is an opinion leader (Section 3.a) or under the common cause model with any value of ε (Section 3.b) or under the model under which Scalia has influence on multiple votes with any value of ε (Section 3.c), we obtain the following results:

Voter	Scalia	Thomas	Ginsburg
D-value	1	0	0

Table 5. Results for a voting rule in which Scalia is a dictator.

These results are very plausible. Only the dictator has influence, whereas the dummies do not have influence. This gives rise to the following observations:

1. The *D*-value significantly depends on the voting rule. As we change the voting rule and keep the model for the voting profile fixed, the value of *D* changes. This is important, since, in voting theory, we are particularly interested in how different voting rules affect the influences of the voters.

2. The results also show that the *D*-value does not suffer from a problem that affects other generalizations of Banzhaf voting power for probability models different from the Bernoulli model. One way of generalizing Banzhaf voting power starts from the observation that standard voting power is a linear transform of the probability of your vote coinciding with the outcome of the vote. One can then quantify influence as the probability of the coincidence of your vote and the outcome of the vote under the probability model that is adopted. As Machover (2007) argues, this will not work, because, in the dictator model, every dummy vote that is perfectly correlated with the vote of the dictator will obtain the same value of the measure as the dictator. Our *D*-value does not have this problem. Even if Thomas blindly follows Scalia as an opinion leader or even if Scalia and Thomas's political views completely overlap (model from Section 3.b with $\varepsilon = 1$), *D* assigns non-identical values to the dictator Scalia and the dummy Thomas. Thomas will only have the same *D*-value as Scalia, if Scalia's vote is fully causally determined by Thomas's vote – and this again is a very plausible assignment of power.

4. The relation to measures of (conditional) success and decisiveness and to Banzhaf Voting Power

Let us now discuss the connection between the *D*-measure and the framework that Laruelle and Valenciano (2005) have set up for measuring decisiveness and success.

For calculating the *D*-measure for voter *i*, we need to know the chances of the counterfactuals $Q_i^0(V = 1| V_i = 1)$ and $Q_i^1(V = 0| V_i = 0)$. These quantities are not present in the framework of Laruelle and Valenciano. Hence, in the general case, our *D*-measure does not fit this framework. However, if there are no causal connection between the votes, then the chances of the counterfactuals coincide with conditional probabilities, i.e.

(4) $Q_i^0(V = 1 \mid V_i = 1) = P\,(V = 1 \mid V_i = 1)$ and $Q_i^1(V = 0 \mid V_i = 0) = P\,(V = 0 \mid V_i = 0)$ for all i

and we can place the *D*-measure in the Laruelle and Valenciano framework. Furthermore, there are also cases in which there are causal connections between the voters in which Eq. (4) holds, e.g. in case all causal influence between the voters is fully mutual.[4] Finally, the *D*-measure for a single voter *i* can be cast in the terms of Laruelle and Valenciano in case the equalities in Eq. (4) hold for that voter only. This is so in case voter *i* is an opinion leader who is not influenced by other people's votes.

Let us now take Eq. (4) for granted and consider the *D*-measure. We obtain

(5) $$D_i = (P(V = 1 \mid V_i = 1) - P(V = 1 \mid V_i = 0))\ \mathrm{P}\,(V_i = 0) \\ + (P(V = 0 \mid V_i = 0) - P(V = 0 \mid V_i = 1))\ \mathrm{P}(V_i = 1)\ .$$

Following Laruelle and Valenciano (2005), p. 175, we will say that voter *i* is successful, if her vote coincides with the outcome of the collective vote – if either ($V = 1$ and $V_i = 1$) or ($V = 0$ and $V_i = 0$). So $P(V = 1 \mid V_i = 1)$ is the conditional probability that *i* is successful given that she votes yes (for the only way that she can be successful given that she votes yes is if the collective vote is acceptance). Following Laruelle and Valenciano, we denote this conditional probability by $\Omega_i^{\,i+}$ (ibid., p. 178). Likewise, $P(V = 0 \mid V_i = 0)$ is the conditional probability that *i* is successful, given that she votes no. We abbreviate this by $\Omega_i^{\,i-}$. Consider now the "mixed" conditional probability $P(V = 1 \mid V_i = 0)$. This equals $1 - P(V = 0 \mid V_i = 0) = 1 - \Omega_i^{\,i-}$. Likewise, $P(V = 0 \mid V_i = 1) = 1 - \Omega_i^{\,i+}$. Putting this all together, we obtain:

(6) $$D_i = (\Omega_i^{\,i+} + \Omega_i^{\,i-} - 1)\,(\mathrm{P}\,(V_i = 0) + \mathrm{P}(V_i = 1)) = \Omega_i^{\,i+} + \Omega_i^{\,i-} - 1\ .$$

Thus, in case Q_i^0 and Q_i^1 equal *P* the *D*-measure can be expressed in terms of quantities from the framework in Laruelle and Valenciano (2005).

Let us now additionally assume that all votes are stochastically independent and that, for each voter *i*, the probability of her voting yes is .5. That is, we adopt the Bernoulli model. Given equiprobility, $\Omega_i^{\,i+} + \Omega_i^{\,i-}$ can further be simplified:

(7) $$\Omega_i^{\,i+} + \Omega_i^{\,i-} = 2\,(.5\,\Omega_i^{\,i+} + .5\,\Omega_i^{\,i-}) = 2\,(P\,(V = 1)\,\Omega_i^{\,i+} + P\,(V = 0)\,\Omega_i^{\,i-}) = \\ = 2\,(P\,(V = 1 \text{ and } V_i = 1) + P\,(V = 0 \text{ and } V_i = 0)) = \\ = 2\,P(\,i \text{ successful}\,)\ .$$

Thus,

[4] To set up a model of full mutual influence requires some background in Bayesian Networks. To represent full mutual influence in Bayesian Networks we can insert v-structures between the nodes with converging arrows into child nodes with dummy variables that are instantiated. Since there are no incoming arrows into the nodes representing the votes, there are no arrows to be erased in the construction of the counterfactual probability model and hence the probability model remains unaffected (*cf.* the Appendix for the terminology).

(8) $$D_i = 2\,P(\,i \text{ successful}\,) - 1.$$

Hence, in the terminology of Laruelle and Valenciano, D_i equals twice Ω_i (the probability of success) minus 1 – it is a linear transform of that quantity.

But, under the full Bernoulli model, the r.h.s. of this equation equals Banzhaf voting power of voter i (see Felsenthal and Machover 1998, Theorem 3.2.16, p. 45). Hence, D_i coincides with Banzhaf voting power in the special case of equiprobability and independence. This result is very pleasing; it tells us that our measure coincides with a standard measure of voting power in case suitable assumptions are made.

5. Wilmers' Example

Let us now turn to a more complex example. Suppose that we have a five person Supreme Court with simple majority voting and equal weights for each voter. The 12 profiles in which at least four voters cast the same vote each occur with probability 1/12 – ε for $0 \leq \varepsilon \leq 1/12$. The 20 remaining profiles each occur with probability $.6\varepsilon$. At $\varepsilon = 0$, we reproduce Wilmers' example, as specified by Machover et al. (2007, p. 3): All 12 profiles with 4 or 5 voters casting the same vote are equiprobable and all other 20 profiles occur with probability zero. Probability models with a finite ε correspond to generalizations of Wilmers' example.

If ε is set at zero, the probability of pivotality is zero for every voter. This shows that we cannot measure influence by means of the probability of pivotality, because, as Machover (2007, p. 3) writes, "it would be absurd to claim that every voter here is powerless, in the sense of having no influence over the outcome of divisions." So let us examine whether our *D*-measure yields more fitting values.

For $0 < \varepsilon \leq 1/12$, we specify a causal interpretation that is consistent with the probability model. We first calculate the conditional probabilities $P(V_A = 1)$, $P(V_B = 1|\ V_A = 1)$, $P(V_B = 1|\ V_A = 0)$, ..., $P(V_E = 1|\ V_A = 0, V_B = 0, V_C = 0, V_D = 0, V_E = 0)$. We then impose the following causal model. A is not influenced by any other voter. B's vote is influenced by and only by A's vote. B votes yes with probability $P(V_B = 1|\ V_A = 1)$ if A votes yes, and with probability $P(V_B = 1|\ V_A = 0)$ if A votes no. C is influenced by both A and B's vote and so on. So A's vote has a causal bearing on B through E's votes, B's vote has a causal bearing on C through E's votes, … and E's vote has no causal bearing. This is illustrated in Figure 4.

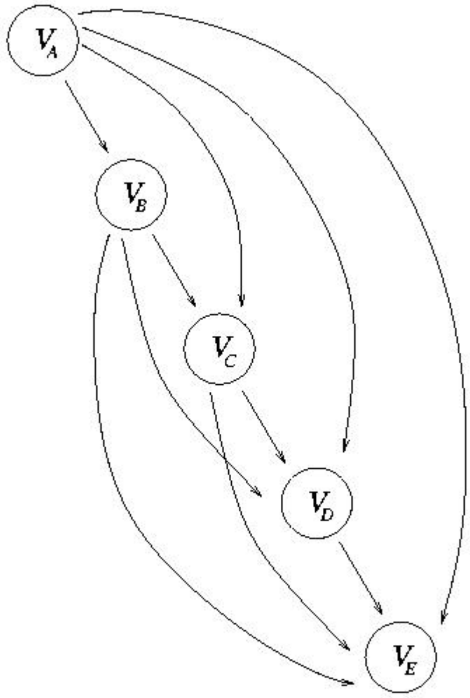

Figure 4. The causal relations that we assume for the generalization of Wilmers' example

Of course there are many other causal models that are consistent with the probability model. E.g. a permutation of A,…, E would also yield a causal model compatible with the probabilities. Furthermore, many more common cause models could be spelled out that are consistent with the probability model. But we will assume that our particular causal model appropriately represents the influences in the real world.

We calculate the *D*-values for this causal model following our methodology. As an example, we obtain $D_A = 2/3 – 9.2\varepsilon$. Subsequently we calculate the limits as ε goes to 0 for all voters. In Table 6, we see that the *D*-values cascade downwards as we move from voters A to E. This squares very nicely with the fact that A is an opinion leader to more voters than B, B is an opinion leader to more voters than C etc.

i	$\mathrm{Lim}_{\varepsilon\to 0}Q_i^0(V = 1 \mid V_i = 1)$ $\mathrm{Lim}_{\varepsilon\to 0}Q_i^1(V = 0 \mid V_i = 0)$	$\mathrm{Lim}_{\varepsilon\to 0}P(V = 0 \mid V_i = 1)$ $\mathrm{Lim}_{\varepsilon\to 0}P(V = 1 \mid V_i = 0)$	$\mathrm{Lim}_{\varepsilon\to 0}D_i$
A	5/6	1/6	4/6
B	2/3	1/6	3/6
C	½	1/6	2/6
D	5/12	1/6	3/12
E	1/6	1/6	0

Table 6. Results for the extension of Wilmers' example

What happens when we set ε at 0 – as is the case of Wilmers' example (2007, p. 3)? When $\varepsilon = 0$ we face a problem in calculating the *D*-value of voter *D*. Suppose that $V_D = 0$ in the real world. We ask what the chance of acceptance would be, if *D* were to have voted yes, i.e. $Q_D^0(V = 1 \mid V_D = 1)$. If we follow the algorithm from Appendix A, we need the probability $P(V_E = 1 \mid V_A = 1, V_B = 0, V_C =$

0, $V_D = 1$). But this probability is undefined since $P(V_A = 1, V_B = 0, V_C = 0, V_D = 1) = 0$. Hence we cannot calculate the $Q_D^0(V = 1| V_D = 1)$ for a probability model with extreme values. For this reason, we stipulated a non-extreme ε-model and calculated the limiting value of the *D*-measure.

But one might object that there are other families of models that approach Wilmers' example in some limit. For example, we could set the probability of one profile with three yes-votes and two no-votes at $.4\varepsilon$ and set the probability of another such profile at $.8\varepsilon$. Again, we will recover Wilmers' example, as we set ε at zero. However, if we take the limit $\varepsilon \rightarrow 0$, we might obtain a slightly different limit for the *D*-measure for D. But it can be shown that for all such families of models, the *D*-values for A, B, C and E are unaffected in the limit $\varepsilon \rightarrow 0$ and the *D*-value for D ranges from 0 to 1/3, i.e. it takes the *D*-values of C and E as its bounds.

6. Discussion

Our assessment of a voter's influence is based upon a causal model. But how do we know whether correlations in the votes are due to opinion leaders or to shared political views, e.g.? And if they are due to opinion leaders, how do we know whether Scalia is an opinion leader for Thomas or vice versa? The standard line in Causal Learning is that a probability model yields conditional independence structures that define a class of causal models. For this class, it may be possible to obtain bounds on the *D*-values of the voters. But if we want to identify a unique causal model we need additional information. One may look at the temporal structure: If Thomas always casts his vote after Scalia, then the causal direction is clear. Or one may appeal to experiment: For instance, one could toggle Scalia's vote and see whether Thomas follows suit. Once a unique causal model is specified, we can assess the *D*-value of each voter. But how this is done and whether this can be done is not the subject of our inquiry.

In this paper we have only dealt with very simple causal models. It is sometimes appropriate to switch to models under which the causal pattern of votes depends on the issue under consideration. Suppose for instance that Scalia has very positive influence on the other voters for economic issues – the other voters are very likely to copy Scalia's vote (model from Section 3.c with $\varepsilon > 0$). But at the same time, Scalia has a negative influence on the other voters for issues of social morality – the other voters are very likely to vote differently from Scalia (model from Section 3.c with $\varepsilon < 0$). In such a case, it is appropriate to construct two causal models and to calculate a *D*-measure for each type of issue that one wishes to consider.

Let us finally deal with an objection against the proposed measurement of voting power. One might object that voting theorists are interested in voting power, i.e. the power that a voter has in virtue of her vote. Measures of this power are supposed to be specifically about a voting rule. The objection is that we mix in different kinds of influence, e.g. the influence that a person has as an opinion leader. This objection motivates an alternative treatment of Wilmers' example: One can say that our intuitions regarding the example are confused, because no distinction is drawn between the extent to which a person can influence the outcome of the vote in virtue of her vote, and in virtue of other things. In virtue of her vote, nobody has voting power, and this is exactly what is captured by the simple suggestion to generalize the Banzhaf measure.

In order to deal with this objection, we want first to point out an ambiguity. "In virtue of her vote" can mean (i) in virtue of her casting a vote - i.e. the real life event of casting a vote - or (ii) in virtue of her vote cast - i.e. the vote as it is registered in the voting profile. Let us then deal with the objection under both readings.

On reading (i), really no objection is left. For it is perfectly conceivable that the event of my voting yes causally bears on how other voters vote and this causal influence should then be incorporated into voting power. But our opponent could then come back and object that opinion leaders typically do not exercise influence through actually casting a vote, since this is in many settings done privately. But nonetheless, we respond, opinion leaders express in more or less uncertain terms what votes they will cast and other people follow their lead. One could have a measure of power that takes into account this type of influence as well. One would not just measure the influence of casting a vote in this case, but rather the influence of having an intention to vote in one way or another. The idea is that this intention has an influence on the outcome of the vote by shaping one's own vote as well as by influencing the votes of others.

On reading (ii), we could indeed say that none of the actual votes cast in Wilmers' example have any influence. But it is odd to say that none of the voters have any influence. The default interpretation of the influence of a voter due to her vote seems to involve some complex interaction of the social processes that brings about the votes as well as the actual voting rule and these two components cannot just be isolated. Our *D*-value measures the influence that a voter has in virtue of a decision rule and her influence on others, and this measure is a worthwhile thing to have precisely because it matches a very natural interpretation of political influence.

Appendix A. The algorithm for calculating the $Q_i^{0/1}$-values

We will follow Balke and Pearl (1994), though our notation diverges. Let us take $Q_i^0(V = 1 | V_i = 1)$ as an example. This is the probability of the counterfactual that the proposal would have been accepted, had *i* voted yes, though *i* votes no in the actual world. *i*'s vote is an event that is embedded in a causal structure. It is caused by certain events and it causes certain effects. Isolate all the non-effects of *i*'s vote. If we learn that *i* votes no in the actual world, then this teaches us something about some of these non-effects of *i*'s vote. We determine a joint probability model for the non-effects of *i*'s vote, conditional on *i* voting no. Subsequently we set *i*'s vote at yes, as if this came about, in Lewis's terms, by a miracle, that is, as if some exogenous force interfered in the course of nature and changed the event from voting no to voting yes (Lewis 1979). We evaluate how the effects of *i*'s vote would be affected by the probability model over the non-effects of *i*'s vote conjoint with *i* voting yes.

Formally, for evaluating $Q_i^0(V = 1 \mid V_i = 1)$ we consider the actual world in which $V_i = 0$. Let the random variables $C_1, .., C_n$ be the non-effects of V_i. The C_js may include other votes and variables representing common causes. We calculate the joint probabilities $P(C_1 = c_1, ..., C_n = c_n \mid V_i = 0)$ where the c_j s range over the possible values for C_j for each *j*. For each combination of the $C_1 = c_1, ..., C_n = c_n$, we then multiply $P(C_1 = c_1, ..., C_n = c_n \mid V_i = 0)$ with $P(V = 1 \mid C_1 = c_1, ..., C_n = c_n , V_i = 1)$. That is, we ask: What is the probability of acceptance, if $V_i = 1$, but if the non-effects of V_i are as they are in the actual world. By summing the products

$$P(C_1 = c_1, ..., C_n = c_n \mid V_i = 0) \times P(V = 1 \mid C_1 = c_1, ..., C_n = c_n , V_i = 1)$$

for every possible combination $C_1 = c_1, ..., C_n = c_n$, we obtain $Q_i^0(V = 1 \mid V_i = 1)$. To calculate $P(V = 1 \mid C_1 = c_1, ..., C_n = c_n , V_i = 1)$, we average over all the effect variables of V_i, call them $E_1 , ..., E_m$:

$$\begin{aligned} P(V = 1 | C_1 = c_1,...,C_n = c_n, V_i = 1) = \\ = \sum_{e_1} ... \sum_{e_m} P(V = 1 | C_1 = c_1,...,C_n = c_n, V_i = 1, E_i = e_i,...,E_m = e_m) \\ \times P(E_i = e_i,...,E_m = e_m | C_1 = c_1,...,C_n = c_n, V_i = 1). \end{aligned}$$

In terms of Bayesian Networks, we can characterize the algorithm as follows:

1. Construct a Bayesian Network with variables (1) for the votes of each voter, (2) for the common causes and (3) for the outcome of the vote. Insert arrows for opinion leaders as in Figure 1, for common cause political views as in Figure 2, and arrows from each voter into the outcome of the vote modelling the decision rule.

2. Read off the prior probabilities $P(V_i = 1)$ and $P(V_i = 0)$ from this network.

3. Set the value of the variable for voter i at no and read off the probability of acceptance, i.e. $P(V = 1| V_i = 0)$.

4. Determine the joint probability distribution over the non-effect variables of V_i conditional on $V_i = 0$.

5. Construct a node for the combination of all non-effect variables $C_1,..., C_n$ and insert this joint probability distribution as a new prior.

6. Erase the nodes for the individual non-effect variables along with their incoming and outgoing arrows.

7. Insert the requisite arrows from the combined non-effect variable to the effect variables of V_i, including the node for the outcome of the vote, and put in the concomitant conditional probability distributions. Note that V_i is a root node in this new network.

8. Set the value of the variable for voter i at yes in this new network. Read off the probability of acceptance. This is $Q_i^0(V = 1| V_i = 1)$.

9. A similar procedure yields $P(V = 0| V_i = 1)$ and $Q_i^1(V = 0|V_i = 0)$.

Acknowledgements

to be inserted

References

Balke, A. and Pearl, J., Probabilistic Evaluation of Counterfactual Queries, Proceedings of the Twelfth National Conference on Artificial Intelligence, AAAII-94, 1994, pp. 230–237

Beisbart, C. and Bovens, L., A Power Measure Analysis of Amendment 36 in Colorado, Public Choice 134 (2008), pp. 231–46

Dretske, F., Explaining Behaviour. Reasons in a World of Causes, MIT Press, Cambridge (MA), 1988

Dubey, P. and Shapley, L. S., Mathematical Properties of the Banzhaf Power Index, Mathematics of Operations Research 4, pp. 99–131

Felsenthal, D. S. and Machover, M., The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes. Cheltenham: Edward Elgar 1998

Gelman, A., Katz, J. N. and Tuerlinckx, F., The Mathematics and Statistics of Voting Power, Statistical Science 17 (2002), pp. 420–35

Gelman, A., Katz, J. N. and Bafumi, J., Standard Voting Power Indexes don't Work: An Empirical Analysis, British Journal of Political Science 34 (2004), pp. 657–74

Kaniovski, S. and Leech, D., A Behavioural Power Index, Mimeo (2007), online under http://serguei.kaniovski.wifo.ac.at/fileadmin/files_kaniovski/pdf/bpower.pdf, checked 12/2007

Laruelle, A. and Valenciano, F., Assessing Success and Decisiveness in Voting Situations: Social Choice and Welfare 24 (2005), pp. 171–97

Lewis, D., Counterfactual Dependence and Time's Arrow, Noûs 13, No. 4, Special Issue on Counterfactuals and Laws (1979), pp. 455–76

Machover, M., Discussion Topic: Voting Power when Voters' Independence is not Assumed, mimeo 2007, online under http://eprints.lse.ac.uk/2966/, checked 12/2007

Morgan, S. L. and Winship, C., Counterfactuals and Causal Inference, Cambridge: Cambridge University Press 2007

Morriss, P., Power. A Philosophical Analysis, Manchester: Manchester University Press 1987 (second edition 2002)

Winship, C. and Morgan, S. L., The Estimation of Causal Effects from Observed Data, Annu. Rev. Sociol. 25 (1999), pp. 659–707