Notions of A Priori Voting Power: Critique of Holler and Widgrén

by
Moshé Machover*

Abstract: These are critical remarks on Holler and Widgrén (1999) and in particular on Holler's Public Good Index. The central idea is a distinction between two underlying pre-formal notions of a priori voting power: power as influence over the outcome, and power as expected relative share in a payoff.

1. Introduction

I have read the preprint of Holler and Widgrén (1999)—henceforth briefly, VCP— and have found it very interesting and stimulating, but I have a number of critical comments to make.

In some of my comments I will refer for further elaboration to the recently published book, Felsenthal and Machover (1998)—henceforth briefly, MVP.

There is one general matter which I must touch upon at the outset, because without it most of my subsequent comments would be totally unintelligible.

Any mathematical measure of voting power must be based on some intuitive pre-formal notion of what voting power is. A given measure can be considered reasonable only if it passes two kinds of test. First, the underlying pre-formal notion must be reasonable and coherent. Second, the mathematical properties of the measure must be reasonable for the intuitive notion it is supposed to formalize.

* Department of Philosophy, King's College, Strand, London WC2R 2LS, England. I am indebted to Manfred Holler for his stimulating comments on earlier drafts, and for his encouragement to submit for publication what started life as an informal letter. I have preserved some of the original informal flavour of my text.

An explicit mathematical formula for measuring voting power can be deduced from a particular model or from a set of axioms (which is just a way of specifying implicitly a class of models). But the model or set of axioms cannot be a substitute for a pre-formal notion. On the contrary, they must provide a bridge from that notion to the formula. For the formula to be justified, some model or set of axioms from which it is deducible must itself be justified in terms of the pre-formal notion.

A corner-stone of MVP is the thesis that there are in fact two quite distinct underlying (intuitive and pre-formal) notions of a priori voting power, which the various measures of voting power attempt to explicate and formalize.

The first notion is that of power as *influence*: the voter's ability to affect the outcome of a division of a voting body—whether the bill in question will be passed or defeated. We call this notion of power "I-power".

The second notion is that of power as a voter's expected *relative share in some prize*, which a winning coalition can put its hands on by the very act of winning. We call this notion of power "P-power".

In his critique of the Shapley–Shubik index, Coleman (1971) drew attention to this distinction (without of course using our terminology), but his insight was largely ignored by most writers on voting power, who tended to conflate the two notions and were thereby led to all sorts of error. Dan Felsenthal and I were among the sinners—most lamentably in our (1995), which is cited in VCP. Only in Felsenthal, Machover and Zwicker (1998) did we begin to rid ourselves of this confusion, and the process of clarification led to the position we advocate in MVP.

In fact, I-power and P-power are fundamentally different. Their explication and formalization lead in rather different directions, and there is no reason to expect, in general, that a property or piece of behaviour that is essential for an acceptable index of P-power should apply also to a valid measure of I-power; or vice versa.

2. I-Power versus P-Power

The notion of I-power has essentially nothing to do with cooperative game theory or, for that matter, with game theory generally. You can see this in several ways. Under this conception, voting behaviour is motivated by *policy seeking* and does not depend on bargaining. In fact, it applies equally well to decision-making bodies in which voting is secret, so that bargaining among the voters is pointless if not impossible. Each voter simply votes for or against a given bill on what s/he considers to be the merit of this bill;

and the way s/he votes is independent of the decision rule. The passage or failure of a bill is here best regarded as a public good, which affects all voters, irrespective of how they have voted on that bill.

Since no bargaining is assumed in connection with I-power, it is misleading to talk here about the "formation" of a coalition in any conscious sense. Even the very term "coalition", as referring to an arbitrary set of voters, is perhaps somewhat misleading, as it seems to imply conscious coordination. But as far as I-power is concerned no such coordination is envisaged. In a division of the voters on a given bill, a set of voters happen to find themselves voting on the same side; this is all. Unfortunately, the use of terminology borrowed from game theory has contributed to the widespread confusion by creating the false impression that voting power is necessarily a game-theoretic notion.

Another way of seeing that the notion of I-power is not fundamentally game-theoretic is the observation that, under this notion, the voting *power* of a voter has nothing whatsoever to do with payoffs. Rather, a voter's I-power depends only on the structure of the voting "game" itself, which contains no information about any payoffs. (So from the viewpoint of I-power, it is not really a game in the true game-theoretic sense, which requires payoffs to be specified.) Of course, we may assume that payoffs do affect voting *behaviour*: they enter the calculations of voters when making up their minds how to vote on a specific given bill. A rational voter will vote for or against a bill by comparing the expected payoff of the passage of the given bill with the expected payoff of its defeat. But the point is that these payoffs are individually determined: they can vary from voter to voter and from bill to bill; and they are completely exogenous to the structure of the voting "game" itself.

Note also that in the case of I-power one can talk meaningfully not only about relative voting power but also about voting power in an *absolute* sense. In fact, absolute I-power is the primary notion, whereas relative I-power is derived from it by normalization. Consider for example two voting games, each having seven voters: the unanimity game $\mathcal{B}_7$, in which passage of a bill requires the "yes" votes of all seven voters; and the simple majority game $\mathcal{M}_7$, in which the "yes" votes of four voters suffice to pass a bill. In each of these games, for reasons of symmetry, every voter must have exactly $\frac{1}{7}$ of the total power. But clearly a voter in $\mathcal{M}_7$ has, in an absolute sense, more influence over the outcome than a voter in $\mathcal{B}_7$. In $\mathcal{B}_7$ a given voter can make a difference in one situation only: when all other six voters vote 'yes'. But in $\mathcal{M}_7$ there are many more situations (in fact, 20) in which a given voter can decide the issue.[1] A voter who expects a priori to be as often in agreement with other voters as in disagreement with them will prefer to be in $\mathcal{M}_7$ rather than in $\mathcal{B}_7$. A voter who is afraid of other voters' powers and expects a priori to be opposed to them—the Thatcher-in-Europe syndrome— will prefer $\mathcal{B}_7$.

Note that in the case of I-power the total amount of absolute power in a game is given endogenously: we can tell how big it is by looking at the game itself.

All serious attempts to explicate and formalize the notion of I-power have led—and in my view must lead—in one direction: to the Banzhaf measure. This has happened several times to people who [re]-invented essentially the same measure independently of one another. They include Lionel Penrose (1946), who as far as I know was the original inventor; Banzhaf (1965); Rae (1969), whose measure is the Banzhaf measure in thin disguise; Coleman (1971), whose two measures are slightly more sophisticated variations on the same theme; and Barry (1980).[2]

By *Banzhaf measure* I mean what is often called the "*absolute* Banzhaf index" and denoted by β'. I reserve the term "Banzhaf index" to the *relative* quantity, usually denoted by β, obtained from β' by normalization, so that its values for all voters always add up to 1.

The Banzhaf measure has a clear probabilistic meaning: *the power of a given voter in a given voting game is the a priori probability of that voter being decisive, tipping the balance between passage and failure of the bill in question.* For example, it assigns to each voter in $\mathcal{B}_7$ absolute power $\frac{1}{64}$, and to each voter in $\mathcal{M}_7$ absolute power $\frac{5}{16}$.

P-power, on the other hand, is a thoroughly game-theoretic notion. It presupposes *office-seeking* voting behaviour aimed at winning, for the sake of obtaining part of the prize, which is available *only* to the winners (and therefore cannot be a public good in the true sense!). It also assumes bargaining and binding agreements. For this reason it makes no sense where voting is secret, because that excludes meaningful bargaining and binding agreement. In order to know to what share of the prize you are entitled, if any, we have

[1]This is analogous to two populations of the same size, with the same distribution of relative wealth; yet one of these populations can be much richer than the other. Absolute wealth is the primary concept, whereas relative wealth is derived from it by normalization.

[2]The case of Barry is particularly ironic, because he vehemently rejects the Banzhaf index, which he mistakenly regards as a game-theoretic measure of P-power; and being in any case hostile to the very idea of P-power he dismissed this index as a "gimmick". But he fails to notice that his own measure of "decisiveness" is . . . essentially the Banzhaf measure. [See Barry (1980, pp. 191, 338).] To be precise, it is exactly the same as Penrose's original proposal, which equals 1/2 of the Banzhaf measure.

to know how you and others have voted.

Also, P-power is an *essentially relative* notion. Absolute P-power makes no coherent sense. For example, there is no reasonable sense in which a voter of $\mathcal{M}_7$ can be regarded as having more (or less) P-power than a voter of $\mathcal{B}_7$. This is because the absolute size of the total prize is determined exogenously; so the games $\mathcal{M}_7$ and $\mathcal{B}_7$ themselves tell us nothing about that.

The Shapley–Shubik index and the Deegan–Packel index are clearly attempts to explicate and formalize this pre-formal notion of P-power. The Johnston index in my view is a hybrid, based on a confused and misconceived attempt to graft a P-power "correction" on the Banzhaf index. In this way it transformed a good index of I-power into a bad index of P-power.[3]

I think that the Public Good Index (PGI) too makes sense only as basically an attempt to explicate and formalize P-power, but incorporating some elements from the sphere of I-power. It looks like an adaptation in the direction of I-power of the Deegan–Packel index, which was justified by its authors explicitly in terms of P-power—the very title of Deegan and Packel (1982) is a slogan of P-power!

All these four indices share one feature, which disqualifies them from measuring a priori relative I-power. Unlike the Banzhaf index, they are not derived by normalization from a meaningful measure of absolute power. For an index of a priori P-power this is no disqualification because, as I have just pointed out, P-power is in any case an essentially relative notion.

The reason why there is no obvious single compelling index of P-power is that such an index should be based on a totally convincing bargaining model for n-person cooperative simple games. But in fact there is no such compelling model for simple games (let alone general cooperative games) with $n > 2$.[4]

I will now proceed to my specific comments on VCP.

3. Critical Comments on VCP

• Ad p. 4, paragraph 1. It is true that the Banzhaf measure does not forecast the outcome of a division. But I think it is misleading to say simply that it is "an expected value", because this might suggest that it is the expected value *of some payoff to the voter*, which is not the case for the Banzhaf measure (but is the case for indices of P-power). In the literature you can indeed find comments by voting-power theorists saying or implying that the Banzhaf

[3]For further details on these matters see MVP, Chapter 6.

[4]For an elaboration of this see MVP, Com. 2.2.2, Section 3.1 and Section 6.1.

measure (or the Banzhaf index) is game-theoretic, and provides an estimate of expected payoffs to the voters. But in my opinion this is just a mistake. It is part of the syndrome to which I have alluded in the Introduction: failure to make the distinction between I-power and P-power. Conflation of the two notions results in illegitimate importation to the sphere of I-power of modes of thought from the sphere of P-power.[5]

• Ad p. 7, last paragraph. The Shapley–Shubik index is not the pioneer. The true pioneer, as far as I know, was Lionel Penrose (1946).

• Ad p. 8, paragraph 1. Why the restriction to *proper* games? Mathematically, the Shapley value and the Shapley–Shubik index are perfectly well defined for improper games. The problem is only that of interpretation of an improper game itself. This is a real problem from the viewpoint of cooperative game theory, and hence of P-power, but it is not at all a problem for I-power. Excluding improper games makes sense only if you are thinking of P-power.[6]

The situation is the same as for other indices of P-power: they are mathematically well-defined for all voting games, and normally the definition is stated in full generality. But they cannot be *used* for improper games, not because the index is undefined, but because from the viewpoint of P-power an improper game has no clear meaning.

So the reluctance of the authors to consider improper games in VCP tends to confirm my assessment that PGI is fundamentally an index of P-power, because from the viewpoint of I-power improper games are not at all problematic.

• Ad p. 10. I agree with the authors of VCP (and many others) that the queue model (in which the n voters queue up to vote for the bill in question, and all $n!$ queues are equiprobable) does not provide a convincing justification for the Shapley–Shubik index, and (as Shapley and Shubik themselves admit) it is "just a convenient conceptual device". But much more can be said about this.

First, there is another, apparently more sophisticated, version of the same queue model (that is, a version with an apparently less simplistic scenario). This version has been used, for example by Laver (1978), to defend the Shapley–Shubik index. But on close examination this version also turns out

[5]However, note that in another sense the Banzhaf measure can be regarded as expected value—not of the payoff to the *voter*, but to an *external agent* who buys that voter's vote. On this see MVP, Thm. 3.2.14 and Com. 3.2.15.

[6]See MVP, Rem. 2.1.2(iv), Com. 6.1.7 and Rem. 6.2.2(iii).

to be untenable as a justification of the Shapley–Shubik index *qua* measure of a priori P-power.[7]

In MVP we also discuss another model (or family of models), presented earlier in Felsenthal and Machover (1996), which we call the *roll-call model*, and which also yields the Shapley–Shubik index. But this model too is not really convincing as a justification of the Shapley–Shubik index *qua* measure of a priori P-power.

As for the Straffin model mentioned in VCP, the situation is quite different. This model, as pointed out in VCP, is probabilistic. And I would add: it is not really game-theoretic. In fact, it is concerned not with P-power but with I-power. You can see this from the facts that, firstly, it does not involve payoffs; and, secondly, no bargaining of any kind takes place. In fact, this model makes as much sense for secret voting as for open voting. We analyse it in detail in MVP.[8] We show there that Straffin's model indeed provides a justification of the Shapley–Shubik index: not as a measure of a priori P-power, but of *a posteriori* I-power. Rather than positing total prior ignorance about the voters' propensities and interactions (which is the standpoint of a priori measures) it assumes that the voters behave as clones, in a highly correlated way.

In the end, the Shapley–Shubik index depends for whatever justification it has *as a measure of a priori P-power* not on this or that representation by some bargaining model (for no convincing model of this kind is known for simple games with more than two voters) but on the fact that it is a special case of the Shapley value. And the Shapley value in turn depends for its justification not on some bargaining model (for, *a fortiori*, no convincing model of this kind is known for general cooperative games with more than two players) but on the conviction provided by Shapley's postulates. This justification, however, is rather vulnerable.[9]

4. Critical Comments on PGI

Finally, I come to the Public Good Index (PGI). I must admit that I find it untenable; firstly, because the scenario or "story" (in effect, bargaining model) offered for justifying it is in my view untenable; and, secondly, because its mathematical behaviour is intuitively wrong. It is the *conjunction* of

[7] See MVP, Com. 6.3.9.

[8] End of Section 6.3, from Def. 6.3.11 to the end of that section.

[9] On these matters see MVP, Coms. 6.2.8, 6.2.26, 6.2.27, 6.3.10, 7.9.19, 7.10.1 and 7.10.2.

these two reasons that, for me, is decisive. If the scenario were compellingly convincing, then one would have to accept (however reluctantly) that our intuition is faulty. But then it would be our scientific duty to explain *why* it is faulty![10] And if the mathematical behaviour of PGI were intuitively acceptable, then we could go on searching for a better story.

Let me take the story first. It only makes sense as a story about P-power, where the voting behaviour is not motivated by policy seeking (the voter's attitude to the proposed bill per se) but by office seeking (the voter's wish to enjoy the fruit of winning per se, irrespective of the content of the proposed bill). It is also clear that it presupposes open voting; otherwise it is not known who is entitled to enjoy the fruit of winning. But then we are told that

(1) the outcome of passing a bill is a public good;

(2) its value (in effect, payoff) to all users is the same;

(3) no entry costs or transaction costs are involved in coalition formation;

(4) there is no rivalry in consumption of the public good;

(5) only members of a minimal winning coalition (MWC) are allowed to enjoy the public good, so as to exclude free-riding.

I find (2) somewhat suspect, because normally the worth of a public good to a given user depends on the user's circumstances, preferences, etc. But perhaps (2) can be justified on the grounds of a priori ignorance.

However, I find the conjunction of (1), (3), (4) and (5) incoherent. If there are no costs and consumption involves no rivalry, then the notion of free-riding becomes meaningless.

Bargaining and a process of coalition-formation must be presupposed in this story: otherwise, how can one justify the special role assigned here to *minimal* winning coalitions? The usual—and in my view the only possible—justification of giving MWCs a privileged role in determining voting power is an appeal to a process of bargaining and conscious coalition formation, in which "redundant" voters are thrown out of a coalition.

But what is there for the voters to bargain *about*? Why should enjoyment of the fruits be confined to the members of an MWC? Why should I deny you enjoyment of a public good, if it costs me nothing and your enjoyment doesn't

[10] This is the kind of job we do in Chapter 7 of MVP, where we analyse in detail the reasons for certain phenomena of voting power that may seem paradoxical.

compete with mine in any way? I know that game theory assumes extreme selfishness, but even it does not prescribe wanton gratuitous vindictiveness.

Compare this with the situation of the Shapley–Shubik and Deegan–Packel indices, which assume that a *constant* sum is divided among the members of a winning coalition that has voted for a given bill. Here voting will only take place after bargaining, and the way each voter votes on the bill will be completely determined by the outcome of the bargaining. Bargaining is necessary in order to lead to a binding agreement as to how the constant sum (the prize of victory) will be divided among the members of the winning coalition. Here it is quite reasonable to suppose that if a coalition S can win even without some member a, then the other members will conspire against a and promise her nothing, so as to leave more to be divided among themselves. But then a will not vote for the bill, because there is no reason for her to do so. In this situation, the Deegan–Packel "story" that only MWCs will form has some prima-facie plausibility.[11] But if the total payoff is not constant, and each member of a winning coalition can enjoy it fully (irrespective of how many other members there are) then what is there to bargain *about*? Why should MWCs have any special status compared to non-minimal winning coalitions?

In VCP (top of p. 14 [????]) it is claimed that a coalition S which is not a MWC may nevertheless be formed by "luck"—that is, by chance—and Barry (1980) is invoked in support on this. But Barry can provide no support here. As far as I can see, he rejects the notion of office-seeking voting behaviour and hence (implicitly) of P-power. The power he is concerned with is I-power. And, like others who have formalized I-power, he assumes a priori ("behind the veil of ignorance") random voting—which can well be secret!—whereby a voter can end up on the successful side by luck. And he (necessarily!) ends up, like the others, with the Penrose–Banzhaf–Rae–Coleman measure, albeit without realizing it.

If the payoff of the passage (or failure) of a given bill to any given voter depends only on the bill itself (as Barry assumes) and not on who voted for or against it, then there can be no role for bargaining; and each voter will vote for or against the bill (independently of the others) according to whether s/he prefers the bill to pass or fail. This is what the notion of I-power presupposes. In this case you can say that from the viewpoint of a given voter, how *another* voter will vote is a matter of luck.

[11] It is *only* prima facie, because a more detailed analysis shows that this assumption is not so reasonable, especially in conjunction with their other two assumptions: that all MWCs are equally probable, and that each MWC will divide the prize equally among its members. See MVP, Comment 6.4.4.

However, if voting depends on prior bargaining, leading to a binding agreement, then "luck" (that is, probability) cannot have any role *after* the bargaining has ended and the agreement concluded. A voter to whom the binding agreement promises payoff 0 will certainly not vote for the bill, because s/he has no incentive to do so. In this scenario, how voters vote depends not on their preferences regarding the bill, but on their ability to enjoy the prize (or part of it). So, *either* we accept that members of a winning coalition may agree in advance to give positive payoff to a non-critical member of the coalition; *or* we say that this never happens. In the former case, there is no justification for basing the measure of voting power exclusively on MWCs. In the latter case, we cannot admit that a non-minimal coalition will ever be formed, even by "luck".

VCP also seems to imply that if a non-MWC S happens to form and win, then the outcome is the same as if an MWC T included in S had formed and won. But there may be more than one such T. Consider for example the weighted voting game $[4; 3, 2, 1, 1]$, with voters a, b, c and d (with quota 4 and weights listed in alphabetic order). If coalition $\{a, b, c\}$ is formed, who will be allowed to enjoy the "public" good, a and b or a and c?

One could consider modifying the story by stipulating that formation of a coalition does involve costs, and (as Deegan and Packel assume) only MWCs get formed. But then, quite apart from having to justify these assumptions—and, as we show in MVP, Com. 6.4.4, justifying the latter assumption would be difficult—another big problem would arise. If formation of a coalition involves costs, then the tacit but crucial assumption that all MWCs have equal probability of forming would be far from natural, let alone compelling. The reason is that the MWCs in one and the same voting game can be very unlike one another. One may consist of a small number of big voters, another of very many small voters, and there are endless other combinations. Even in the absence of costs, the assumption of equi-probability is problematic; in the presence of costs it becomes very questionable indeed.

Now let me address the counter-intuitive mathematical behaviour of PGI. This concerns not only non-monotonicity (which is mentioned explicitly in VCP), but some other counter-intuitive phenomena such as the donation and bloc paradoxes, which afflict PGI. As I mentioned before, and as VCP seems to agree (p. 16 last para, cont. on p. 17 [????]), this is an acute problem if one is not convinced by the story behind the model.

However, the mention of the Fischer–Schotter "paradox" in this connection is, in my opinion, a red herring. The Fischer–Schotter phenomenon is counter-intuitive only to a very naïve and inexperienced observer. In our paper (Felsenthal and Machover 1995), as well as in MVP, Section 7.3, we analyse this phenomenon and show why it is perfectly reasonable for any decent measure of voting power. Things are very different with true non-monotonicity (that is, when a measure attributes "more power" to a lighter voter than to a heavier one *in the same weighted voting game*) and the donation and bloc paradoxes. In my view (and, as VCP admits, in the view of most political scientists and economists) such phenomena are unacceptable for a valid measure of absolute I-power or index of P-power. I find the counter-arguments of Brams and Fishburn (1995) quite unpersuasive.

Finally, I agree with the dictum that "if we could trust our intuition, then power indices in general would be rather unnecessary". Yes; but in my view the duty of science is not only to replace naïve intuition; it is also to "educate" intuition and to show, by analysis that delves beneath the surface, which naïve intuitions are misleading and why. In Chapter 7 of MVP we try to do so for the main known paradoxes (both serious and not so serious), some of which we invented ourselves.... When we are unable to find a satisfactory explanation or justification for an apparently counter-intuitive piece of behaviour by a given measure, we feel forced to conclude that the measure in question is suspect.[12]

References

Barry, B. (1980), "Is it better to be powerful or lucky?", *Political Studies*, 28, pp. 183–194, 338–352.

Banzhaf, J.F. (1965), "Weighted voting doesn't work: a mathematical analysis", *Rutgers Law Review*, 19, pp. 317–343.

Brams, S.J. and P.C. Fishburn (1995), "When is size a liability?", *Journal of Theoretical Politics*, 7, pp. 301–316.

Coleman, J.S. (1971), "Control of collectivities and the power of a collectivity to act", in: B. Lieberman (ed.), *Social Choice*, New York: Gordon and Breach.

Deegan, J. and E.W. Packel (1982), "To the (minimal winning) victors go the (equally divided) spoils: a new index of power for simple n-person games', in: S.J. Brams, W.F. Lucas and P.D. Straffin (eds.), *Political and Related Models* (Vol. 2 in series *Models in Applied Mathematics* edited by W.F. Lucas); New York: Springer.

[12]Cf. the case of the Shapley–Shubik index's flagrant violation of the Added Blocker postulate, discussed in MVP, Section 7.9.

Felsenthal, D.S. and M. Machover (1995), "Postulates and paradoxes of relative voting power: a critical re-appraisal", *Theory and Decision*, 38, pp. 195–229.

Felsenthal, D.S. and M. Machover (1996), "Alternative forms of the Shapley value and the Shapley–Shubik index", *Public Choice*, 87, pp. 315–318.

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

Felsenthal, D.S., M. Machover and W. Zwicker (1998), "The bicameral postulates and indices of a priori voting power", *Theory and Decision*, 44, pp. 83–116.

Laver, M. (1978), "The problem of measuring power in Europe", *Environment and Planning A*, 10, pp. 901–906.

Holler, M.J. and M. Widgrén (1999), "The value of a coalition is power", *Homo Oeconomicus*, [????????].

Penrose, L.S. (1946), "The elementary statistics of majority voting", *Journal of the Royal Statistical Society*, 109, pp. 53–57.

Rae, D.W. (1969), "Decision rules and individual values in constitutional choice", *American Political Science Review*, 63, pp. 40–56.