The Story of the *Poor* Public Good Index

Manfred J. Holler(✉)

University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany
Manfred.Holler@uni-hamburg.de

Abstract. The paper starts from the hypothesis that the public good index (PGI) could be much more successful if it were introduced by a more prominent game theorist. It argues that the violation of local monotonicity, inherent to this measure of a priori voting power, can be an asset – especially if the public good interpretation is taken into consideration and the PGI is (re-)assigned to I-power, instead of P-power.

Keywords: Public good index · PGI · Banzhaf index · Shapley-Shubik index · Local monotonicity · I-Power · P-Power · NESS concept

JEL Classification: D71 · C78

1 Why the Public Good Index Is Poor

On the occasion of a workshop on "Institutions, Games and Experiments" at the Max Planck Institute of Economics at Jena, held in honor of Werner Güth's 70th birthday on January 31–February 3, 2014, I had the honor to sit next to Reinhard Selten during two dinners.[1] He told me that he was rather happy to work on a farm in Austria when he was a teenager and his family had to leave Breslau because of the approaching Russian troops. He still liked doing farmwork after his family had to leave Austria because they were German. But as we all know, this was not his final dedication. During the last forty years, Selten and I have met at various conferences and seminars. I remember several intensive discussions during these rare occasions. Yet, I was surprised that he still remembered some of the topics when we were sitting next to each other enjoying our dinners. One topic he remembered was my research on mixed strategy equilibrium. He had warned me that my results might be redundant – and he was right.[2] Another topic was the Public Good Index (PGI). His conclusion was that if a (more) prominent game theorist had introduced this measure of a priori voting power, PGI would be more popular than the Banzhaf index [4] and perhaps as popular as the Shapley-Shubik index [29].

[1] Reinhard Selten was born October 5, 1930 in Breslau (today Wrocław), and died August 23, 2016 in Poznań. In 1994, he received a Nobel Prize in Economics together with John Nash and John Harsanyi.

[2] My results were already published in [3], however, this paper was not quoted in the discussion of the mixed strategy equilibrium and its application during 1980–2010.

N. T. Nguyen et al. (Eds.): TCCI XXXIV, LNCS 11890, pp. 171–179, 2019.
https://doi.org/10.1007/978-3-662-60555-4_12

Obviously, this was a compliment for the measure, but not necessarily for me. In fact, I have hesitated to talk about Selten's evaluation for more than two years pondering over his resolution. Then I concluded that it should be of general interest to see how much the career of a theoretical concept can depend on the person's status promoting the concept. We should expect that this does not only hold for a theoretical concept, but for scientific work per se. I am sure we can find many examples supporting this hypothesis. For many scientists this seems obvious and needs no further discussion. However, perhaps it does need further discussion because this bias is so obvious – and, in many cases, a serious problem to scientific work. I do not want to give a general analysis of this problem, but outline an example of which I have rather intimate knowledge: the story of the PGI. Section 2 introduces the index, presents some facts of the history of the PGI and outlines the monotonicity problem inherent to this measure. In Sect. 3, the public good interpretation of the PGI will be discussed with respect to the concepts of I-power and P-power. I will argue that the wrong assignment of the PGI, i.e., P-power, leads to confusion and a debasement of this measure. Section 4 concludes this paper with the observation that postulates and stories discriminate between the power measures, however, more generally, the discussion within the community decides about the interpretation and success of an analytical instrument or a theoretical concept.

2 The PGI, Its Prehistory and Its Monotonicity Problem

In their critical review of power measures, Felsenthal and Machover [12: 486] write:

> "As far as we know, the first person to be concerned with the measurement of voting power was Luther Martin, a Maryland delegate to the 1787 Constitutional Convention held in Philadelphia. Martin was worried that the voting power of the large states in the US House of Representatives would be disproportionately too large compared to that of the small states, assuming that the representatives of each of the 13 states would always vote as a bloc. In a pamphlet published the following year he not only exposed the fallacy of equating voting power with weight (in this case size of a voting bloc), but made an attempt – albeit unsystematic and somewhat crude – to measure voting power."

Felsenthal and Machover [12: 486] conclude that "Martin's approach is broadly based on the notion of what we have called 'I-power' …. But the measure implicit in his argument apparently relates the voting power of a voter *a* just to the number of minimal winning coalitions to which *a* belongs; and therefore seems to us closest to Holler's index…," which is the PGI. In more recent times – almost 200 years after Luther Martin – it seems that I was the first to apply this measure. In [15],[3] I made use of it to evaluate the distribution of voting power in the Finnish Parliament (Eduskunta) underlying the formation of governments for the period 1948–1978. I tried to compare the theoretical power values with the observed participation of the various parties in the governing of

[3] This was in the year when Deegan and Packel published their power index referring to minimum winning coalitions, only. See [8]. I am absolutely sure that we did not know of each other, but Riker's "size principle" [28: 32ff] was "in the air." In 1980, I met Ed Packel at his home in Lake Forest, Illinois; I cannot recall that we discussed the "priority issue."

the country.[4] Finland had up to ten parties in its parliament and all parties were in some coalition governments during this period – but the very right party was never in coalition with the very left one, still both were in some government coalitions. During this period, the President of the Republic controlled most of the policymaking[5] – for example, he defined foreign policy and represented the government abroad. He was the commander-in-chief of the armed forces and appointed the members of the government, i.e., the Council of the State. The legislative power was vested with the Parliament. In the exercise of his authority, the President was bound to cooperate with the Parliament through the Council of State, "which must enjoy the confidence of parliament." While there was almost every year a new government, and members of seven different political parties have held position as ministers, Urho Kekkonen was Finnish president from 1956 to 1981.

I had doubts about the PGI because, in this study, this measure produced power values that showed a violation of local monotonicity: party blocs with a larger seat share were assigned smaller power values than party blocs with smaller seat shares. "This property of the MWC index causes doubt concerning its validity" [15: 33]. Note that, instead of PGI, I used the label MWC, i.e., minimum winning coalition – the public good idea did as yet not come into my mind. As an alternative to the PGI, I applied the normalized Banzhaf index which seemed "more adequate in the context of this analysis" [15: 33]. It satisfies local monotonicity.

In order to illustrate the monotonicity problem of the PGI we choose the "notorious example" of a weighted voting game v= (*51*; 35, 20, 15, 15, 15),[6] i.e., a decision rule $d = 51$ and a vote distribution w = (35, 20, 15, 15, 15) assigned to a set of voters $N = \{1, 2, 3, 4, 5\}$ such that, e.g., $w_2 = 20$. (Voters represent a bloc of votes – the blocs of voters are the players in the voting game v.) The set of minimum winning coalitions of this game (i.e., coalitions that do not contain any surplus players) is

$$\mathrm{M}(v) = \{\{1,2\},\{1,3,4\},\{1,3,5\},\{1,4,5\},\{2,3,4,5\}\}$$

so that the corresponding PGI is

$$\mathrm{h}(v) = \left(\frac{4}{15},\frac{2}{15},\frac{3}{15},\frac{3}{15},\frac{3}{15}\right).$$

This result indicates that voter 1 is in four minimum winning coalitions of game v, and so on. In its raw version the index counts the number of minimum winning coalitions, i.e., of the "decisive sets" of game v, which have a particular voter i as a member. The numbers 4, 2, 3, 3, and 3 represent the "decisiveness" of the players in game v.

[4] See [16] for the empirical results and the analysis.

[5] There were amendments to the constitution and today the power of the President is mostly symbolic. However, officially, the President still leads the nation's foreign politics together with the Council of State and is the Commander-in-Chief of the armed forces.

[6] See [16] for an early discussion of this example.

This counting implies:

(a) Each member of a minimum winning coalition gets the same value, i.e., there is no sharing.
(b) These values are identical for all minimum winning coalitions.
(c) A minimum winning coalition's contribution to the calculation of the PGI is equal to its cardinality.

Independent of the chosen power measure, the set of minimum winning coalitions M(v) fully captures the characteristics of the game v with respect to coalition formation and therefore with respect to power. It is an alternative representation of game v. Therefore, the PGI uses all we know of the game v.

Note that the denominator 15 is the total number of voters who are in some minimum winning coalition of this game. It is chosen for standardizing the PGI so that

$$\sum_i h_i(v) = 1$$

The denominator is not meant to imply a sharing rule. However, this leads us to the interpretation of the index formula. But let us first further discuss the non-monotonicity result illustrated by this simple example above.

Felsenthal and Machover [10: 211] write that "it seems intuitively obvious that if $w_i \leq w_j$ then every voter j has at least as much voting power as voter i, because any contribution that i can make to the passage of a resolution can be equalled or bettered by j." But is j as welcome as i in the forming of coalitions? More generally, do we really *know* that power is monotonic in vote shares? One of the starting points of power index research is the assumption that the vote distribution is a poor proxy for the distribution of voting power. If we could *prove* that the vote-power relation is monotonic, then vote distribution is perhaps a reasonable proxy. However, there is no such proof. Can we trust our intuition?[7]

Felsenthal and Machover [11: 221ff] conclude that "any reasonable power index" should be required to satisfy local monotonicity. Any a priori measure of power that violates local monotonicity is "pathological" and should be disqualified as a valid yardstick for measuring power [11: 221ff]. As a response to this statement, Holler and Napel [20, 21] argue that a violation of local monotonicity could be a characteristic feature of a game, perhaps indicating an instable power relation between the voting blocs with problems to coalition formation. Is the non-monotonicity property of the PGI a pathology or can it serve as an indicator, revealing certain peculiarities of a game?[8] The more popular power measures, i.e., the Shapley-Shubik index and the Banzhaf index, satisfy local monotonicity irrespective of the property of the game.

[7] "…if we could trust our intuition, then power indices in general would be rather useless. The number of paradoxes related to the application of these measures, which are the result of a deviation from intuition, indicates that our intuition most likely needs help when it comes to evaluating power – or forming 'reasonable expectations' with respect to power" [18: 607], also see [17].

[8] For further discussion of this argument, see [22, 23].

However, they show non-monotonicities if the vote shares are redistributed. Then there is no guarantee that an increase of a share results in an increase in the (relative) power value. It could well happen that the corresponding power value decreases.[9]

> *Power is a social phenomenon. It does not only depend on the resources we have but also on the resources of the other decision-makers and the distribution of these resources.*

There are many other paradoxes with the Shapley-Shubik index and the Banzhaf index (see, e.g., [7]). Alonso-Meijide and Bowles [1] demonstrate that the Shapley-Shubik index violates local monotonicity if there are a priori unions and, as a consequence, coalitions are no longer equally likely.[10] Note the PGI gives a weight of zero to winning coalitions that contain surplus voters, i.e., voters that are not critical to the winning of a coalition. Thus, coalitions are not equally likely in this case.

Alternatively, we can hope to learn more about weighted voting games if we look for the property of those games which satisfy local monotonicity even when power is measured by the PGI. This is the research program outlined and illustrated in [24]. An interesting variation of this research program has been offered by Freixas and Kurz [14]. They analyze the question how much solidarity we can afford, as measured by the PGI, to guarantee local monotonicity if the "rest of power" is captured by the Banzhaf index. To answer this question, they look for convex combinations of the PGI and the Banzhaf index of a game and their potential of monotonicity. On the one hand, the resulting indices, satisfying local monotonicity, are closer to the Banzhaf than to the PGI, on the other, they are all the more "solidary" than the Banzhaf index.

Earlier, Widgrén [32] proved the following equation that relates the normalized Banzhaf index (β_i) and the PGI (h_i): $\beta_i = (1 - \pi)\, h_i + \pi\varepsilon_i$. Here π represents the share of winning coalitions that contain surplus players and ε_i represents the share of such coalitions for a particular player i. Widgrén probabilistic model allows to discuss the relationship between the PGI and the Banzhaf index and therefore points to elements which are responsible for the PGI's violation of local monotonicity if observed.

Loosely speaking, the difference between the PGI and the normalized Banzhaf index boils down to those winning coalitions that are not minimal. Holler [16, 19] argues that these coalitions should not be considered because they imply a potential for freeriding if the decisions are on public goods – as is often the case in policy making.[11] This does not mean that surplus coalitions do not form, but they should not be considered when measuring power. The focus on minimum winning coalitions excludes freeriding. If winning coalitions form which contain surplus players, "then it is by luck, similarity of preferences, tradition, etc. – *but not because of power*" [18, 607].

[9] This well-known "paradox of redistribution" was introduced in [13].

[10] See [2] for further discussion.

[11] "The *basic principles* underlying the public good index are (a) the public good property, i.e. non-rivalry in consumption and non-excludability of access, and (b) the non-freeriding property. It is obvious from these principles that (strict) minimum winning coalitions should be considered when it comes to measuring power. All other coalitions are either non-winning or contain at least one member that does not contribute to winning. If coalitions of the second type form, then it is by luck or because of similarity of preferences, tradition, etc.—but not because of power, as there is a potential for freeriding" [22: 9].

3 I-Power, P-Power, Solidarity and the Public Good Interpretation

It is not obvious how and when the concept of solidarity emerged to characterize the PGI. Moreover, it is not clear what this concept contributes in addition to the public good interpretation of the PGI introduced in [16].[12] This interpretation was inspired by Brian Barry's [5: 189] observation that the coalition value is a collective good and a concept of dividing the value of a coalition "violates the first principle of political analysis, which is that public policy is a public good (or bad)." For illustration, he continues: "If the death penalty is reintroduced, that pleases those who favour it and displeases those who do not. Similarly, a tax break is a good or bad for people according to their situation. The gains are not confined to those who voted on the winning side nor are the losses confined to those who are on the losing side" [5, 189].

From Paul Samuelson we learned that there is nothing to share in the case of (pure) public goods as both non-rivalry in consumption and non-excludability of access apply. The argument is that those critical to the winning of a coalition exert some power to achieve the public goods which they prefer to alternative outcomes. It is assumed that different minimum winning coalitions bring forward different public goods as collective outcome.

In order to characterize the properties of various indices Felsenthal and Machover (1998) introduced the concepts of I-power and P-power where I-power is meant to capture the influence of (a bloc of) voters on the outcome while P-power summarizes the shares of the spoils voters derive from "winning an election" and holding an office. P-power is about "sharing a cake" while I-power is about having a say in determining the outcome. For instance, [11, 12] suggest that the Banzhaf index describes I-power, an agent's potential influence over the outcome, whereas the Shapley-Shubik index represents P-power, an agent's expected share in a fixed prize. However, as demonstrated by Turnovec [30, 31], with respect to the formal structure of the two measures the distinction does not hold: both measures can be interpreted expressing I-power or, alternatively, P-power. Turnovec convincingly argues that a "power index speaks about the properties of a model, not about the properties of the power as such" [30: 613]. Whether a specific power index is appropriate or not, depends on the properties of the model of collective decision-making which we want to analyse. Therefore, it is not surprising "that there is not just one power index. However, the category in which a specific index falls is not always obvious" [25: 290]. If the model considers the result of the collective decision making regarding a public good, then the sharing approach is inappropriate. Therefore, it is somewhat surprising that Felsenthal and Machover [11] classify the PGI among the P-power measures. It is difficult to see why the PGI does not qualify as an I-power measure like, according to Felsenthal and Machover, the Banzhaf index does. (See the formal relationship of the two measures outlined above.) Moreover, as quoted above, Felsenthal and Machover conclude that Luther Martin's approach is (a) broadly based on the notion of I-power, and (b) it seems closest to

[12] In the sequel, Holler and Packel [26] axiomatized the PGI. Napel [27] completed this axiomatization.

"Holler's index" - which is the PGI. It is therefore inconsistent assigning the PGI to P-power.

This assignment is of course not without consequences. In the case of P-power and sharing a cake, the violation of local monotonicity could be a problem and perhaps earn the label pathology. But does this still hold in case that we want to measure the influence of blocs of voters on the outcome? There is a lot of empirical evidence that this influence might be non-monotonic – like in the case of the Finnish Parliament (see above).[13]

It seems natural that the PGI does not achieve good results if tested with respect to postulates stated for P-power. Felsenthal [9: 368] gives a substantial list of six postulates that a "reasonable P-power index should satisfy." We read

> "…the various indices proposed to date for estimating the expected share in the fixed prize of the members in an *n*-person cooperative game – which, following Felsenthal and Machover (1998, ch. 6), we shall call *P-power indices* – must be assessed by examining which postulates – i.e., intuitively compelling conditions – they satisfy. Failure to satisfy these conditions is a suspect counter-intuitive behavior, which can be regarded as paradoxical or, in extreme cases, pathological, and may indicate that a P-power index guilty of it must be discarded" [9: 368].

In the list of P-power indices analyzed by Felsenthal we find the PGI which violates four of the six postulates that Felsenthal proposed for P-power indices – and therefore should be discarded. Indeed, the PGI should be discarded from the list of P-power indices as it is not meant to be a P-power index. One could have argued that this poor result should have signaled that the PGI does not represent P-power. The underlying public good model does not support sharing. It seems that Felsenthal gets rather close to this insight:

> "But the prize of victory according to the PGI index is regarded not as a unit of TU, a private good, to be divided among the members of a victorious MWC, but as a public good enjoyed in its entirety by all members of this coalition (but only by them!)" [9: 377f].

One should add that some people, not in the coalition, will be lucky and also enjoy this public good X, but they have no influence on whether X is produced or not. For criminals this luck has, however, a "negative sign" in the case that additional police officers are hired for increasing the security of the citizens.

Public goods are a close relative to externalities and therefore invite freeriding. Consequently, if we relate power to causality, we should exclusively consider decision-makers which are decisive for their production, i.e., "necessary elements of a sufficient set" – which justifies to focus on minimum winning coalitions only. The NESS-concept ("necessary elements of a sufficient set") has been introduced in philosophy to discuss causality in collective actions.[14] The set of the "necessary elements of a sufficient set" is a minimum winning coalition. This suggests the PGI as measure of causality. However, irritated by non-monotonicity results, Braham and van Hees propose a *weak*

[13] In this case, one might argue that the situation allows for an interpretation that considers P-power. Under the umbrella of the power of the President, the numerous coalition governments were sharing the benefits of holding an office. The contribution to the political result was perhaps secondary to many participants in this game.

[14] See [6] with reference to [33].

NESS-test ending up with the Banzhaf index instead of the PGI (which corresponds to the *strong* NESS-test).

4 Concluding Remarks

The discrimination against non-monotonicity, implicit in the weak NESS-test, ignores that possible causal relations are not only determined by the resources of an agent but also depend on the resources of the other agents and how they are employed. In the case of power, the weak NESS-test does not consider that power is a social phenomenon and therefore, in general, not only depending on the resources of a single decision maker.

We have seen that the objections to the PGI because of its non-monotonicity results are closely related to its P-power assignment. There is nothing in the mathematical structure of the various indices which says whether the index expresses sharing influence (I-power) or sharing a cake (P-power). There are the underlying stories and the scientific community that put the measures in the various characterizing boxes. Still, a concept can get misplaced – and doomed. It is the discussion within the community that decides on its performance and its status. I hope that this paper helps that the PGI is no longer discarded for producing pathologies, but applied clarifying the notion of power in cases of collective decision-making when the outcome can be considered a public good.

References

1. Alonso-Meijide, J.M., Bowles, C.: Power indices restricted by a priori unions can be easily computed and are useful: a generating function-based application to the IMF. Ann. Oper. Res. **137**, 21–44 (2005)
2. Alonso-Meijide, J.M., Bowles, C., Holler, M.J., Napel, S.: Monotonicity of power in games with a priori unions. Theor. Decis. **66**, 17–37 (2009)
3. Aumann, R.J., Maschler, M.: Some thoughts on the minimax principle. Manag. Sci. **18**, 54–63 (1972)
4. Banzhaf III, J.F.: Weighted voting doesn't work: a mathematical analysis. Rutgers Law Rev. **19**, 317–343 (1965)
5. Barry, B.: Is it better to be powerful or lucky? Parts 1 and 2. Polit. Stud. **28**, 183–194 and 338–352 (1980)
6. Braham, M., van Hees, M.: Degrees of causation. Erkenntnis **71**, 323–344 (2009)
7. Brams, S.J.: Game Theory and Politics. Free Press, New York (1975)
8. Deegan, J., Packel, E.W.: A new index of power for simple n-person-games. Int. J. Game Theory **7**, 113–123 (1978)
9. Felsenthal, D.S.: A well-behaved index of a priori p-power for simple n-person games. Homo Oeconomicus **33**, 367–381 (2016)
10. Felsenthal, D.S., Machover, M.: Postulates and paradoxes of relative voting power - a critical appraisal. Theor. Decis. **38**, 195–229 (1995)
11. Felsenthal, D.S., Machover, M.: The Measurement of Voting Power. Theory and Practice, Problems and Paradoxes. Edward Elgar, Cheltenham (1998)

12. Felsenthal, D.S., Machover, M.: Voting power measurement: a story of misreinvention. Soc. Choice Welfare **25**, 485–506 (2005)
13. Fischer, D., Schotter, A.: The inevitability of the paradox of redistribution in the allocation of voting weights. Public Choice **33**, 46–67 (1980)
14. Freixas, J., Kurz, S.: The cost of getting local monotonicity. Eur. J. Oper. Res. **251**, 600–612 (2016)
15. Holler, M.J.: A priori party power and government formation. Munich Soc. Sci. Rev. **1**, 25–41 (1978). (republished in: Holler, M.J. (ed.) Power, Voting, and Voting Power. Physica-Verlag, Würzburg and Wien (1982); and Munich Social Science Review, New Series, vol. 2 (2018))
16. Holler, M.J.: Forming coalitions and measuring voting power. Polit. Stud. **30**, 262–271 (1982)
17. Holler, M.J.: The public good index. In: Holler, M.J. (ed.) Coalitions and Collective Action. Physica-Verlag, Würzburg and Vienna (1984)
18. Holler, M.J.: Power, monotonicity and expectations. Control Cybern. **26**, 605–607 (1997)
19. Holler, M.J.: Two stories, one power index. J. Theor. Polit. **10**, 179–190 (1998)
20. Holler, M.J., Napel, S.: Local monotonicity of power: axiom or just a property. Qual. Quant. **38**, 637–647 (2004)
21. Holler, M.J., Napel, S.: Monotonicity of power and power measures. Theor. Decis. **56**, 93–111 (2004)
22. Holler, M.J., Nurmi, H.: Reflections on power, voting, and voting power. In: Holler, M.J., Nurmi, H. (eds.) Power, Voting, and Voting Power: 30 Years After, pp. 1–24. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-35929-3_1
23. Holler, M.J., Nurmi, H.: Pathology or revelation? The public good index. In: Fara, R., Leech, D., Salles, M. (eds.) Voting Power and Procedures. SCW, pp. 247–257. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-05158-1_14
24. Holler, M.J., Ono, R., Steffen, F.: Constrained monotonicity and the measurement of power. Theor. Decis. **50**, 385–397 (2001)
25. Holler, M.J., Owen, G.: On the present and future of power measures. Homo Oeconomicus **19**, 281–295 (2002); republished as Present and future of power measures. In: Kacprzyk, J., Wagner, D. (eds.) Group Decisions and Voting, pp. 31–46. Academicka Oficyna Wydawnicza, Warsaw (2003)
26. Holler, M.J., Packel, E.W.: Power, luck and the right index. Zeitschrift für Nationalökonomie (J. Econ.) **43**, 21–29 (1983)
27. Napel, S.: The holler-packel axiomatization of the public good index completed. Homo Oeconomicus **15**, 513–521 (1999)
28. Riker, W.H.: The Theory of Political Coalitions. Yale University Press, New Haven (1962)
29. Shapley, L.S., Shubik, M.: A method of evaluating the distribution of power in a committee system. Am. Polit. Sci. Rev. **48**, 787–792 (1954)
30. Turnovec, F.: Power, power indices and intuition. Control Cybern. **26**, 613–615 (1997)
31. Turnovec, F.: Power indices: swings or pivots? In: Wiberg, M. (ed.) Reasoned choices. Essays in Honour of Academy Professor Hannu Nurmi on the Occasion of his 60th birthday. Digipaino, Turku (2004)
32. Widgrén, M.: On the probabilistic relationship between the Public Good Index and the Normalized Banzhaf index. In: Holler, M.J., Owen, G. (eds.) Power Measures, vol. 2. Homo Oeconomicus **19**, 373–386 (2002)
33. Wright, R.: Causation, responsibility, risk, probability, naked statistics, and proof: Pruning the bramble bush by clarifying the concepts. Iowa Law Rev. **73**, 1001–1077 (1988)