# On the Possibility of a Preference-Based Power Index: The Strategic Power Index Revisited

**Dieter Schmidtchen and Bernard Steunenberg**

## 1 Introduction

The *strategic power index* (hereafter SPI) is an alternative method for evaluating the distribution of power in policy games, which has been introduced several years ago. Whereas traditional power indices are based on the notion that players need to form some kind of majority or winning coalition, the SPI employs the analytical tools of non-cooperative game theory. Key features of decision-making situations, such as actor preferences, the outcome or policy space, as well as the rules of the decision-making process, are part of a game-theoretical analysis, which forms the basis of the calculation of this index. Since the analysis allows players to act strategically, this index is labeled *strategic*. It reflects the power-related features of the position of political actors in a context as modelled in a non-cooperative game.[1] Since our first proposal of this index, several comments have been raised against the SPI. In this

[1]This index was presented, in non-normalized form, for the first time in 1996 (see Steunenberg et al. 1996). There is another attempt to develop a strategic power index labelled *strict power index* (Napel and Widgren 2002). As with our index, spatial preferences and strategic agenda setting are its main building blocks. However, in the framework of the strict power index, power is defined as the ability of a player "to change the current state of affair" (Napel and Widgren 2002: 4). Following the reasoning of traditional power indices power relates to the ability of being decisive or pivotal.

D. Schmidtchen (✉)
Center for the Study of Law and Economics, University of the Saarland, Building 31, PO-Box 15 11 50, 66041 Saarbrücken, Germany
e-mail: csle@rz.uni-sb.de

B. Steunenberg
Department of Public Administration, Leiden University, Campus Den Haag (SBS 2), Schouwburgstraat 2, 2511 VA Den Haag, Leiden, Netherlands
e-mail: b.steunenberg@cdh.leidenuniv.ed

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_15,

paper we will discuss these comments and clarify our position on voting indices and the SPI.

The SPI is very different from conventional power indices, such as the Banzhaf, Shapley–Shubik and others. These indices take the set of players and the bare decision-making rules as their domain and measure voting power by the extent to which a player in any collective body that makes yes-or-no decisions by vote may turn a losing coalition into a 'winning' coalition for all mathematically possible permutations of players (Shapley–Shubik index) or the relative number of times a player is decisive in a vote (Banzhaf index).[2] As mentioned by Felsenthal and Machover (1998) conventional voting power analyses are either based on cooperative games[3] or are entirely probabilistic measures.[4] What is measured is a priori power. By doing this, conventional power indices do not take account of positive or negative correlations of players preferences.[5] Comments also point to the limited capability of traditional power indices to model players' strategic interaction and a complicated institutional structure typical for real world decision-making (see Garrett and Tsebelis 1997, 1999a, b; Steunenberg et al. 1996, 1997, 1999; Schmidtchen and Steunenberg 2002).

The SPI rests on a notion of power as the ability 'to get what you want'. Important is to distinguish the modelling of a decision-making situation—a game—from the calculation of the index. While the modelling of decision-making situations may be done in different ways—for instance, by introducing new players, changing the sequence of play, information sets, or action sets of players—the calculation

[2]These 'classical' indices have been supplemented with more recent power measures, such as the Johnston index, the Deegan–Packel index and the Holler index. The main differences between these indices are the ways in which coalition members share the benefits of their cooperation, and the kind of coalition players chose to form (see Colomer 1999). For a comparative investigation of traditional power indices see Felsenthal and Machover (1998), Holler and Owen (2001) and Laruelle and Valenciano (2008).

[3]For example, the Shapley–Shubik index measuring what Felsenthal and Machover call P-power, which posits an office-seeking motivation of voting behavior (see Felsenthal and Machover 1998: 171).

[4]See Penrose (1946), Banzhaf (1965), Coleman (1971, 1986), which take a policy-seeking viewpoint focusing on the degree to which a member's vote is able to influence the outcome of a vote. These indices reflect I-power in the sense of Felsenthal and Machover (1998: 36).

[5]That is not to say that traditional power indices are unable to take account of voters' preferences or spatial voting (see Straffin 1994). In probabilistic characterizations of voting power indices each voter $i$'s probability $p_i$ of voting "yes" on a proposal is a random variable. Taking the $p_i$ as an indicator of the acceptability of a proposal to voter $i$ (see Straffin 1994: 1137), homogeneous as well as heterogeneous preferences can be modelled. If each $p_i$ is chosen independently from the uniform distribution on [0,1] we have the Banzhaf index. The independence assumption means that the acceptability of a proposal to voter $i$ is independent of its acceptability to any other voter $j$ (see Straffin 1994: 1137). Note, that $p_i = 1/2$, which means that voter $j$ voting "yes" is similar to flipping a coin. Note further, that the probability characterization of the Banzhaf index is restricted to its non-normalized version. If random variable $p$ is chosen from the uniform distribution on [0,1], and $p_i = p$ for all $i$ (homogeneity assumption), we have the Shapley–Shubik index. Here the acceptability of a proposal is the same to all voters.

of the index remains the same. The index uses, among others, the distances between outcomes and ideal points, which follows a distribution of preference configurations, including status quo points. The resulting expected or average distance for a player is compared with a 'neutral' or dummy player, which helps to differentiate between a player's success and luck. The smaller the expected distance, the more power is attributed to a player compared to the 'neutral' player who does not have any decision-making rights in the game. The index uses a distribution of states of the world, that is, various combinations of preferences and initial policies (status quo points). It levels out the effect of luck, that is, of being close to the equilibrium outcome in a specific game by using numerous different preference configurations and taking averages. The intuition is that the power of a player resides only in the game form or the rules of a game and not in the way a specific game is played (Steunenberg et al. 1996, 1997, 1999; Schmidtchen and Steunenberg 2002).

The SPI gave rise to several comments in the literature. Garrett and Tsebelis (1999b) argue that the SPI—although an improvement compared to conventional indices—nevertheless suffers from a drawback generated by the statistics used in it. Felsenthal and Machover (2001) proved a theorem stating that the SPI is a modified Banzhaf index. Napel and Widgren (2004: 519) give credit to it being the first unified approach to the measurement of decision-making power in that it combines an ex post analysis of well defined games with the ex ante prospect of being successful in the game form underlying these games. However, this first attempt to provide such a framework is considered to be 'problematic'. Napel and Widgren (2004: 524) point to a potential for confounding power and success "that may, but need not, result from it". In an earlier paper, they speak of a confusion of cause and effect (Napel and Widgren 2002: 2). They claim that "(o)nly for particular distribution assumptions ... luck (is) 'leveled out' by taking averages" (Napel and Widgren 2004: 524). The SPI is judged to be "a good measure of expected success but in general, it fails to capture power" and it may even become negative (Napel and Widgren 2004: 524). A fundamental critique has come from Braham and Holler (2005) who deny the possibility of a preference-based power index on the ground that it is incompatible with "a fixed core of meaning of power", i.e., the basic notion of power as a generic ability.

This paper contains our responses and is organized as follows. Section 2 describes the logic of the SPI. In Sect. 3 we, first, present the arguments leading Braham and Holler to deny the possibility of a preference-based power index. We then demonstrate why these arguments are not convincing. Section 4 deals with the argument put forward by Napel and Widgren that the SPI is not a true power index since it confuses power and luck. Section 5 addresses the question of whether the SPI can become negative. Section 6 is concerned with the Felsenthal and Machover proposition that the SPI is nothing but a modified Banzhaf index. Section 7 concludes the paper and presents our outlook.

## 2 The Strategic Power Index

As discussed by Steunenberg et al. (1999), the strategic power index approaches power as a player's ability to affect the equilibrium outcome in a game. The basic intuition is that the stronger a player's influence on the outcome under a specific game form, the more powerful this player is. The index is based on several elements: the modeling of a decision-making process using the tools of noncooperative game theory, the definition of a state space (outcome space), a distribution of state variables, and the use of an index to measure power.

The first element of the approach is the development of a game-theoretical model of a decision-making process, which includes various structural elements such as the set of players, their action sets, the possible sequence of moves, the distribution of information, the set of outcomes, and an outcome function mapping the space of strategy profiles into a set of outcomes. These elements define what is called a game form.

The second element is to define the preferences of the players. Let $n \in N$ be the number of players in a game form π and $X \subseteq \Re^m$ an $m$-dimensional and finite outcome space. For this space players are assumed to have Euclidean preferences which can be characterized by player $i$'s ideal point $x_i = (x^1{}_\mathrm{i}, x^2{}_\mathrm{i}, \ldots x^\mathrm{m}{}_\mathrm{i})$.[6] Let $q \in X$ denote the status quo, that is, the hypothetical state of affairs before the start of the decision-making process. This can be the current policy, or the situation without such a policy. We call a combination of a particular ideal point for each player and the status quo a 'state of the world', which will be denoted as $\xi = (x_1, x_2, \ldots x_n, q)$.

The third step is to feed the 'state' variables into the game form, π, based on some distribution. Combining a specific state, ξ, with the game form, leads to a specific game with some (unique) equilibrium outcome $x^\pi(\xi)$.[7] In this context, each particular state of the world is assumed to be the instance of a random variable $\overline{\xi} = (\overline{x}_1, \overline{x}_2, \ldots \overline{x}_n, \overline{q})$. In order to assess how well a player perform in a game form, we determine the expected distance between the equilibrium outcome and the player's ideal point for all possible configurations of preferences and the status quo, or states of the world. The expected or mean distance between the equilibrium

[6]For a more general version of the SPI, it is only necessary that $X$ is some metric space, i.e. a space on which a metric (distance function) is defined, which, for every two points in $X$, gives the distance between them as a nonnegative real number. Such a metric space must satisfy the axioms of symmetry, positive definiteness and triangle inequality. The most familiar metric space is the (one- or multidimensional) Euclidean space which we assume in this paper. The Euclidean space is translation and rotation invariant and stretching, shrinking or mirroring at the origin does not alter the SPI.

[7]At this point we focus on a unique equilibrium outcome only for expositional convenience. The strategic power index can also be applied to games for which multiple equilibria exist. If the game does not have a unique equilibrium, but multiple equilibria, the simple Euclidean distance can be replaced by the average Euclidean distance, i.e. the sum of the Euclidean distances between each equilibrium outcome and the player's ideal point for all equilibria in a particular state of the world, divided by the number of equilibria.

outcomes for some game form, π, and player $i$'s ideal point is given by

$$\Delta_i^\pi = \int \delta_i^\pi f\left(\overline{\xi}\right) d\left(\overline{\xi}\right)$$

where

$$\delta_i^\pi = \sqrt{\sum_{k=1}^{m}\left(x^\pi\left(\overline{\xi}\right)^k - \overline{x}_i^k\right)^2}$$

is the Euclidean distance between the equilibrium outcome of the game and the ideal point of player $i$ in any particular state of the world, and $f\left(\overline{\xi}\right)$ is the density function if $\overline{\xi}$ is a continuous random variable.

The mean distance, as expressed by $\Delta_i^\pi$, provides information on how well a player is doing in the context of a game form. This distance allows us to assess a player's power vis-à-vis the other players: all other things being equal, a player is more powerful than another player if the expected distance between the equilibrium outcome and its ideal point is smaller than the expected distance for the other player. This measurement forms the basis of the proposed power index.

The last step of the approach concerns the development of an index. In order to distinguish 'power' from 'luck', which are both contained in the measurement of expected distances, we need some standardization. This is done by comparing the 'performance' of a player with that of a dummy. A dummy is defined as a player with preferences for the same space as actual players, but who does not have any decision-making rights in the game form. Moreover, dummy player's preferences are independent of the other players. As a consequence this player or his/her preferences do not matter for the outcome of the game. The dummy only experiences some equilibrium outcome that is set by the other players. Sometimes the dummy is 'lucky' in having an ideal point that is close to the equilibrium outcome. However, in other 'states of the world' the dummy may be less fortunate and encounter a policy outcome that is quite different from its preferred option. Consequently, the mean distance found for this player represents a minimum value that can be associated with a 'powerless' player and provides a baseline above which we can speak of power.

For a dummy player, $d$, the expected distance between his ideal point and the equilibrium outcome of a particular game based on game form π can be defined as $\Delta_d^\pi$. The (absolute) strategic power index for a player $i$ can now be defined as[8]:

[8]The relative power of player $i$ can be defined as $\widetilde{\Psi}_i^\pi = \frac{\Delta_d^\pi - \Delta_i^\pi}{\sum_{j=1}^{n}\left(\Delta_d^\pi - \Delta_i^\pi\right)}$. The relative power scores of all players add up to 1.

$$\Psi_i^\pi = \frac{\Delta_d^\pi - \Delta_i^\pi}{\Delta_d^\pi} = 1 - \frac{\Delta_i^\pi}{\Delta_d^\pi}.$$

This index lies in the interval [0,1] and increases with the power of player $i$. The expected distance for a player that is ‘powerful’ enough to dictate the outcome of a game under any preference configuration would be zero, leading to a corresponding value for the index of one. By contrast, if a player has an effect on the outcome of a game, similar to that of the dummy player (which, by definition, is ‘powerless’), the expected distance for this player is the same as for the dummy player, leading to a corresponding index value of zero.[9]

Based on this index, there is a natural way to approach the status quo bias of a game form, that is, the extent to which players are unable to act and to pull a new policy away from the current state of affairs. For a specific game form, that status quo bias can be measured by the expected distance between the equilibrium outcome and the status quo, which is defined as $\Delta_q^\pi$. Substituting this value for the expected distance found for a player in the strategic power index, we get

$$\Psi_q^\pi = \frac{\Delta_d^\pi - \Delta_q^\pi}{\Delta_d^\pi} = 1 - \frac{\Delta_q^\pi}{\Delta_d^\pi},$$

which is called the *inertia index*. A value of one for this index means that under some game form the status quo always prevails. The smaller the value for the index, the more players are able to move the equilibrium policy away from the status quo.

## 3 Impossibility of the SPI?

In this section, we deal with the critique put forward by Braham and Holler (2005) who argue that a preference-based power index is impossible. We, first, present the main argument leading Braham and Holler to deny the possibility of a SPI. Next, we discuss its main shortcomings.

### *3.1 A “Core Theorem of the Measurement of Power”*

Braham and Holler want to bring the “semantics of power into the centre of the debate about how to measure power” (Braham and Holler 2005: 139). By referring to the philosophical semantic analysis of power they take the notion of a “generic

[9] Since the ideal points for each player are independent random variables, the equilibrium outcomes can never be systematically biased *against* the interest of a particular player, and, therefore, no player can fare worse than the dummy player. Thus, the proposed index can never become negative.

ability to effect outcomes" to be "the natural 'fixed core of meaning' of power" (Braham and Holler 2005: 145). In their words:

> *If* player *i* wanted a particular outcome or set of outcomes and that *i* has an action (or sequence of actions) such that the performance of these actions under stated or implied conditions will result in that outcome or set of outcomes and would not result if *i* would not perform this action (or sequence of actions), *then* player *i* would perform this action (or sequence of actions) and the specified outcome or set of outcomes would obtain. That is, *i* is essential or non-redundant for an outcome or set of outcomes (Braham and Holler 2005: 145).

In the absence of that player's intervention the state of the world would be different (Braham and Holler 2005: 145). Regarding simple games, one would speak of a swing (Braham and Holler 2005: 145, n. 8) or a player being decisive or pivotal.

From the definition of power as a capacity or potential to affect outcomes Braham and Holler conclude that a measure of power cannot accommodate any reference to the preferences of the players with respect to affecting outcomes. A power ascription is, first, categorical, second, "leaves the matter of what *i* wants undefined" and, third, "does not say how much power *i* has, only that there exist circumstances in which *i* is non-redundant for the outcome; a measure of power—power index—aggregates these ascriptions of non-redundancy in some way" (Braham and Holler 2005: 145–146).

The central claim of the article is formulated as 'Core Theorem of the Measurement of Power', which is, as the authors concede, not a theorem in the formal sense of the term, but rather "a kind of conceptual impossibility result that is germane to the theory of power generally" (Braham and Holler 2005: 138). The 'theorem' is stated in the following way:

> Core Theorem of the Measurement of Power: If power is the ability of *i* to affect an outcome, then a measure of *i*'s power must exclude any reference to *i*'s preference (behavioural content) with respect to affecting that outcome (Braham and Holler 2005: 146).

Three reasons are given for this statement:

> (1) being disinclined to do something does not imply the inability to do it;
> (2) psychological states such as desires and wants are not normally applied to the concept of ability; and
> (3) the *exercise* of an ability is not to be conflated with its *possession* (Braham and Holler 2005: 146).

Braham and Holler are of the opinion that a preference-based power index such as the SPI violates these three conditions: it conflates disinclination with inability (Braham and Holler 2005: 146–148); redefines the game form, since a 'phobiafied' strategy, i.e. a strategy which is not rational being chosen, cannot be considered a strategy at all (Braham and Holler 2005: 148–150); commits the so called *exercise fallacy* by conflating the "*possession* of a disposition (*having* power) with its *exercise*" (Braham and Holler 2005: 151).

With regard to a game theoretical setting, Braham and Holler 'derive' a corollary of their theorem, which

> states that a player's power resides in, and only in, the strategies available to her given by the *game form* and not in the way that she plays the *game*. This implies that power is a value-independent concept. The upshot is that the Core Theorem renders unintelligible any attempt to formulate a measure of power in terms of the equilibrium of a non-cooperative game – the very idea of strategic power indices. Put bluntly, assessing how a player may play a game does not help us answer such questions as 'Is Smith more powerful than Jones?' or 'What is the extent of Smith's power?' because power concerns what players may be *able to do*, not the actions they *may* or *do* take (Braham and Holler 2005: 139).

Interestingly also to Braham and Holler power is linked to the game form and the strategies available to players. The SPI gathers information on the success of the various, available strategies, which are embedded in the game form, by comparing the outcomes for a distribution of states of the world. Following a broad and general distribution basically levels out the effect of specific values related to preferences and policies, which seem to be the main reason of Braham and Holler's objection.

## *3.2 Is There a Fixed Core of Meaning of Power? (Pitfalls of Essential Definitions)*

The essay of Braham and Holler is an exercise in semantics. They concede that they are "making liberal use of the philosophical semantic analysis of power conducted" (Braham and Holler 2005: 139), but they add: "It must not, therefore, be thought that we are refreshing old philosophical debates. Rather, we are bringing the *semantics* (italics added) of power into the centre of the debate about how to measure power" (Braham and Holler 2005: 138). In fact, they claim having formulated the 'right' ('true') definition of power, with 'general ability to affect outcomes' constituting its essence or intrinsic fundamental nature. In the philosophy of science those definitions are called 'essential definitions'. The problems associated with essential definitions are well known (Popper 1960, Chapter I.10): Is there one, and only one, notion of power? How do we know that 'general ability' is the essential property of power? How can we evaluate the definition in terms of the truth or falsity of the description given by it? Referring to "what we customarily mean by ability" (Braham and Holler 2005: 144) is a doubtful criterion, raising more questions than solving ones.

We should try to avoid converting substantial problems in purely semantic (verbal) ones, since this paves the path for endless discourses. We should reject the view that we should aim at and can obtain ultimate explanations by looking for essences. Following the path of methodological nominalism, definitions such as 'power is a generic ability' should be read from right to left, as an answer to 'What *shall* we call a generic ability in a game form?', and not from left to right as an answer to 'What *is* power in a game form?' Accepting this rule, one would be rather reluctant in conducting an 'analysis of power per se', as done by Braham and Holler (2005: 154). Since we do not have a criterion for figuring out what 'power per se' actually is, it seems reasonable to take an instrumental stance to the definition and to ask, 'Which definition is helpful in answering scientific questions?', and: 'Why are we interested in a definition of power?' The answer clearly depends on where we want to use the term. Several possibilities come to mind: If power is part of a theory, the explanatory power of the theory might depend on the definition. For normative statements, the workability or the empirical relevance of a concept of power in the sense of 'What people are really interested in' might be decisive. From this perspective one might ask: Of what interest is it to know what a player is able to do, if it is not rational to do it? Why should we be interested in the potential or capacity of an action to alter outcomes if this does not is in accordance with equilibrium behavior?[10]

Our position is that power can and should be defined in several ways depending on the research question. In some contexts it might be useful to define power the way Braham and Holler did, i.e. applying the criterion of decisiveness, in others it is better to follow the SPI approach, which relies on the criterion of success. In Sect. 3.3.3, we will show that in take-it-or-leave-it committees—these are the voting bodies to which the SPI is applied—the criterion of success is the better measure of power. In the next sections we illustrate the relevance of the above arguments by analyzing some well-defined games.

## 3.3 *Thinking Strategically vs. 'Analysis of Power Per-Se' or: Why the Inclusion of Preferences Is Necessary*

Strategic interactions arise in two forms. The first is sequential. Here, players make alternating moves whereby earlier moves are observable to those choosing later. In a simultaneous game, players act at the same time in ignorance of the other player's current actions (game of imperfect information).

### 3.3.1 Constitutional Choice[11]

Consider three legislators, A, B, and C, who must vote in alphabetic order under a majority rule, on whether to increase their own salaries (see Ordeshook 1992: 41f.). Each legislator prefers to receive the pay raise, but each realizes that the constituents will not be pleased with a legislator voting to increase his own salary. There are four possible outcomes (see Ordeshook 1992: 41):

$o_1$: The raise passes, but the legislator votes against it.
$o_2$: The raise passes, and the legislator votes for it.
$o_3$: The raise fails, and the legislator votes against it.
$o_4$: The raise fails, but the legislator votes for it.

[10] See Barry's critique of the Shapley–Shubik index (Barry 1980).

[11] This part is from Schmidtchen and Steunenberg (2002: 208–210).

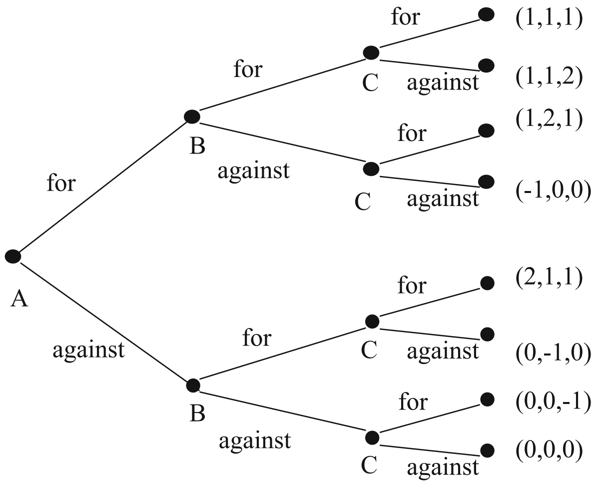


**Fig. 1** The pay-raise game in extensive form

Let $u$ denote utility, then the preferences of the legislators are summarized by the following numbers

$$u_{\mathrm{i}}(\mathrm{o}_1) = 2, u_{\mathrm{i}}(\mathrm{o}_2) = 1, u_{\mathrm{i}}(\mathrm{o}_3) = \mathrm{o}, u_{\mathrm{i}}(\mathrm{o}_4) = -1, \quad \text{with} \quad \mathrm{i} = \mathrm{A}, \ \mathrm{B}, \ \mathrm{C}.$$

Figure 1 represents this voting situation in extensive form, where the terminal nodes are associated with the payoff 3-tuple ($u_A$, $u_B$, $u_C$).

A game form analysis of this voting game, i.e. neglecting the payoffs of the players, which applies 'general ability to affect outcomes' as indicator of 'power per-se' would reveal that each player has power: If two players were to vote differently the third one is decisive, he decides which state of the world obtains—pay-raise or status quo. Since this result holds for each player, traditional power indices would assign equal power values to each of the three players. However, the game is a sequential one, which matters a great deal.

Assume society consists of our three players who had to choose, say, unanimously the game form underlying the game that is to be played afterwards.[12] Which game forms are candidates for getting unanimous support?

From the point of view of power as a 'general ability to affect outcomes' the game form of Fig. 1 is a candidate. Power seems equally distributed—the sequential order of play does not matter. We doubt that players are so stupid not to see that the order of play is highly relevant. The legislators must vote alphabetically—player A is moving first, player B moving second and player C moves last. Player A

[12]This is a traditional constitutional choice problem (Buchanan 1990). On the constitutional level, society must choose the rules (choice of rules) that govern decision-making on the post-constitutional level. On the post-constitutional level, choices have to be taken within the rules decided upon on the constitutional level.

| | A for | | A against | |
|---|---|---|---|---|
| | B for | B against | B for | B against |
| C for | 1, 1, 1 | 1, 2, 1 | 2, 1, 1 | 0, 0, -1 |
| C against | 1, 1, 2 | -1, 0, 0 | 0, -1,0 | 0, 0, 0 |

**Fig. 2** Pay-raise game in strategic form

enjoys a first mover advantage, which can be seen if we derive the subgame perfect equilibrium. The unique subgame perfect equilibrium is: (against, for, for). Player A receives the highest payoff (2), whereas the other players must content themselves with their second best outcomes. Note that if B and C were to change position the outcome of the game would not be affected. Clearly, each player would prefer to occupy the first mover position. A sequential game form such as in Fig. 1 would only have a chance to be chosen on the constitutional level, if uncertainty exists with respect to the first mover position. In such a case, each player must form beliefs about his position. With a perfect veil of ignorance these beliefs would be identical, leading to identical expected utilities for the players, given the majority rule $m = 2$. Similar calculations are required for the $m = 3$ and $m = 1$ rules.

Having done all these computations, the players can choose, on the constitutional level, which rule is best for them given the sequential order of play. But why should a sequential game be chosen at all? On the constitutional level players are free to choose a simultaneous game, which would change the structure of information of the pay-raise game dramatically.

Figure 2 portrays this game in strategic form (see Ordeshook 1992: 45); ordering of the payoffs (player C, player B, player A).

In contrast to the game portrayed in Fig. 1, where players A, B, C have 2, 4, 16 strategies, respectively, in the simultaneous game the strategy sets B and C are identical to A's—to vote for or against. Here, player C cannot condition on the choice of player A or B and B cannot condition on player A's choice. This game has four Nash equilibriums: (A for, B for, C against), (A for, C for, B against), (A against, B for, C for), (A against, B against, C against). Thus, in this section we reached a conclusion similar to that in the previous Section. A constitutional analysis, which restricts itself solely to the analysis of game forms, would be incomplete.

### 3.3.2 Inferior Players

Next, consider a 3-player simple game where the only winning coalitions are the grand coalition ABC and the two coalitions AB and AC (this example is from Napel and Widgren 2001: 213; Widgren and Napel 2001: 1–2). Looking at this game as a

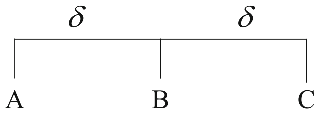


**Fig. 3** Preference constellation

coalitional form game the Banzhaf and Shapley–Shubik power vectors are $(\frac{3}{5}, \frac{1}{5}, \frac{1}{5})$ and $(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$, respectively.

From the point of view of non-cooperative game theory—following Napel and Widgren (2001: 213)—the game can be looked at as a sequential game, in which A makes, after flipping a coin, an ultimatum offer to B, asking for approval in return for an only marginal (and in the limit non-extent) concession to B's interest. A rational player B would have to accept the proposal, if a blocking coalition BC cannot be formed. B knows that if he rejects the proposal, A would move to C, who serves as a perfect substitute in forming a winning coalition. A similar reasoning holds in the case in which A makes the ultimatum offer to C. Thus, we would conclude, contrary to what power measures based on coalitional form games indicate, B and C are powerless in this game. Napel and Widgren (2001: 213–14) call players that are robbed of their power commonly associated with their swing inferior players.

Indeed, application of the machinery of the strategic power index shows that B and C are powerless. Consider a policy space with three possible outcomes and identical distance, denoted $\delta$, between two neighboring outcomes; player set {A, B, C} and D as dummy player. This player is not a true player but rather an outside observer. The ideal points are uniformly distributed on the policy space. Figure 3 shows one of the feasible preference constellations.

Translating the notion of an ultimatum game to our setting means that, whatever the distribution of ideal points (preference profile) of players A, B, C, the policy outcome always corresponds to A's ideal point. Thus, A's power score $\Psi_A = 1$. Since we assumed that the probability distribution of D's ideal points is the same as those of B and C, D's expected distance equals those of B and C: $\Delta_B = \Delta_C = \Delta_D$. Thus, $\Psi_B = \Psi_C = 0$.[13] The conclusion is that traditional power indices assign power scores to players without taking into account their position as inferior players. Thus, they neglect a factor that may be highly relevant from a player's point of view.

### 3.3.3 Agenda Setting Versus Veto Power[14]

Related to the first mover issue analyzed above is the problem of agenda setting power and veto power. Agenda setters are first movers, but not every first mover

[13]Note the difference to the strict power index approach favored by Widgren and Napel, where A is not treated as a 'pure' ultimatum player. Whereas A's SPI score is 1, the strict power index is 5/7. However, according to both indices B and C are powerless.

[14]This part is based on Schmidtchen and Steunenberg (2002: 212–214).

is an agenda setter. Classical power indices are not able to analyze and evaluate these distinctive types of power since all players are simply veto players. There are only two subsets of players, a 'winning' and a 'losing' one. Braham and Holler acknowledge that those indices are insensitive to the strategic aspects of power relations (Braham and Holler 2005: 141). They even illustrate this feature by two elementary examples. In one of their examples they consider a committee of seven players with each player having one vote and a 5/7-majority rule (Braham and Holler 2005: 142–143). They further assume a preference configuration ranking the players in a one-dimensional policy space and a proposal falling from heaven. They show that not all coalitions are rationally feasible, and that not every swing will be exercised by a rational agent (Braham and Holler 2005: 142). However, despite acknowledging the intuitive appeal of the critique encapsulated in the examples, they are still having the opinion that "it is fundamentally mistaken. The reason hinges on a conceptual issue: what we mean by a power ascription" (Braham and Holler 2005: 143). A substantial problem is converted into a verbal one! In most of the committees there are agenda setters. Furthermore, it is well known that a specific type of power, different from the power of a veto player, is associated with the position of an agenda setter as a first mover. A power concept that systematically neglects the sequential structure of collective decision-making is unable to measure this type of power and to address the problems associated with it.

To show why traditional voting power indices do not represent the distribution of power between an agenda setter and several veto players in a satisfactory and meaningful way, we choose, as a simple example, a decision-making procedure used in the European Union. With regard to legislative decision-making, the EC Treaty initially provided only for the unanimity version of the consultation procedure. This procedure allowed the Commission to propose new regulations or directives, which are subjected to unanimous consent by the Council. The latter implies that, in fact, each Council member has the right to veto the Commission's proposal. The European Parliament only needs to be consulted in this procedure. Since the Council can adopt a proposal regardless of the position Parliament takes, Parliament does not play a significant role and thus will not be discussed further.

Now assume that policies can be represented by a one-dimensional (left-right) outcome space and players have Euclidean preferences. In addition, assume that players have perfect and complete information. The Commission selects a proposal, which is then decided upon by the Council members. For our argument on the usefulness of voting power indices, we assume that Council members are not allowed to add new proposals to the agenda or to amend the Commission proposal. The interactions between the Commission and Council members now resemble the well-known agenda-setter model of Romer and Rosenthal (1978, 1979).

Figure 4 presents a preference configuration that may occur for the Commission, which is conceived as a unitary actor, and a five-member Council. In this figure $V_i$ and C denote the most preferred or ideal points of Council member *i* and the Commission, respectively, and $V_i(q)$ stands for member *i*'s point of indifference to the status quo q. The Commission, C, has a more progressive preference than most Council members, $V_i$. Nevertheless, the leftmost Council member, $V_1$, holds

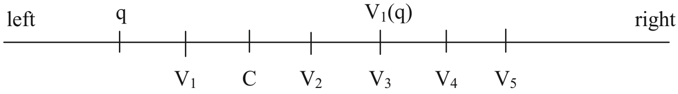


**Fig. 4** Preferences of the Commission and the Council Members

an even more extreme position. Given a status quo to the left of these players, the Commission will propose a measure that is equivalent to its own most preferred point. Since all Council members prefer this point to the status quo, the proposal will not be vetoed. So, in equilibrium, the outcome of this game is a legislative policy $x = C$.

In this context, all players have to approve a measure, and no measure can be taken without the support of each one of them. Each (last) player has the same probability of being pivotal, and each player is necessary to form the (minimum) winning coalition of all players. The Shapley–Shubik, Banzhaf, Johnston and Holler indices therefore allocate power values of 1/6 to each player. These individual scores would suggest that the Commission is as 'powerful' as the Council members. The aggregated score of the Council would even be 5/6, which implies that the Council would be *more* powerful than the Commission.

Adding together the scores of individual Council members to calculate the power of the Council leads to what we call an *aggregation bias*. This bias is the result of the fact that, in interbody analyses of voting power, the members of separate decision-making bodies are treated as if they were the members of a single committee. However, in the game as discussed, a proposal must be approved by both the Commission and the Council, regardless of the voting rule the Council uses to reach a collective decision. If the Commission does not belong to a coalition, then this coalition is not a winning coalition. Both players can be regarded as necessary players. Therefore, one would expect that both actors have power values of 1/2, and not 1/6 for the Commission and 5/6 for the Council. The bias, as revealed by these numbers, leads to an exaggeration of the Council's abilities and an understatement of the power of the Commission. The reason is that the abilities of these players to affect the equilibrium outcome differ: the Commission can take the initiative and draft a proposal, while Council members can only approve or reject this proposal. Council members may restrict the Commission's policy choice, but they cannot set the final proposal. The Commission enjoys discretion in choosing a new policy, which makes it more 'powerful' than the traditional indices indicate.

In addition, the power value of the Council, in a game with the Commission, is independent of the number of Council members. The individual values are only relevant to assess each member's power in shaping a Council decision and not a decision that has to be taken by several 'institutional' actors, including composite decision-making bodies.

### *3.4 Strategic Power: Ability of Being Successful*

The SPI measures a player's ability/capacity/potential (whether generic or not) on average to influence (affect) as a member of a voting body the equilibrium outcome of a voting game or, in other words, the ability/capacity/potential 'to get what you want' by incentivizing as a member among other members of a voting body an agenda-setter to present proposals which approach as close as possible the preferences of the respective player. It is an indicator of average success of affecting *equilibrium* outcomes.

This potential to affect equilibrium outcomes is determined by the game form, the state space and state variables (which are random variables). Taking into account the preferences of the players serves the purpose of determining rational behavior and to derive the equilibrium in a specific game.[15] Since the sole sources of power are the game form, the state space and the state variables we can fully subscribe to Braham and Holler's statement:

> Ordinarily speaking, a 'power' ascription refers to a person's ability: what a person is able to do. In the game theoretic context that we are discussing, the ability in question is to effect outcomes (i.e. 'force' or 'determine' outcomes) of the game. That is, a player has a strategy that, if chosen, will make a decisive difference to the outcome. *This basic definition is the same for a power index based upon a simple game and one that is ostensibly based upon a non-cooperative game* (italics added). The difference lies in the specification of the ability. In a simple game, the ability is turning a winning coalition into a losing coalition or vice versa, thereby being decisive for the acceptance or rejection of a bill, while, in a non-cooperative game, the ability is specified in terms of shifting the equilibrium in one's own favour (Braham and Holler 2005: 143).

Note that in both models of a decision-making procedure the veto-players have identical action sets: they can either reject or accept a proposal. But only in the non-cooperative game setting players are assumed to act rational, i.e. choosing that action which leads to the better individual payoff.

It depends on the decision-making rule whether or not a player is *decisive* as for the equilibrium outcome. With a unanimity rule each veto player is decisive in the sense that the rejection or acceptance of a proposal always, i.e. whatever the preference configuration, depends on the action chosen. With a rule of simple majority there are sometimes preference configurations in which the equilibrium outcome of the game, i.e. either the status quo or, if there is a proposal, its content, crucially depends on the action of a player; but sometimes the equilibrium outcome is determined irrespective of the action chosen by a player. Nevertheless, in the latter case still distances between the ideal points and the equilibrium outcomes can be calculated and they are included in our power measure.

From the discussion above it should be obvious that, contrary to what Braham and Holler (2005: 147–148) believe, taking into account the state space and state variables in measuring a player's power does not mean conflating disinclination

---

[15]Of course, our approach can also be applied to games of incomplete information, which would require making assumption regarding the possible types of players.

with inability. What the SPI measures is simply the ability/capacity/potential of a *rational player* to affect an (equilibrium) outcome, which is a subset of all possible outcomes. Finally, contrary to what Braham and Holler (2005: 150–152) believe, we do not conflate the possession of power with its exercise thereby committing the so-called exercise fallacy. What the players, the agenda-setter and the veto-players, do is exercising rational behavior. Whether or not, for example, a veto-player affects the equilibrium outcome depends on the decision-making rule and the rational behavior of all other players.

Meanwhile even adherents of the traditional power index approach question that there is only one notion of voting power, namely decisiveness. They realize that the notion of 'satisfaction' or 'success', "that is, focusing on the likelihood of having the result one voted for irrespective of whether one's vote was crucial for it or not" (Laruelle et al. 2006: 186) is a meaningful notion of "voting power" and might be more relevant than decisiveness from the voters' point of view (Laruelle et al. 2006: 189; Laruelle and Valenciano 2008). Whereas in so-called bargaining committees decisiveness is the adequate notion of power, in so called take-it-or-leave-it committees—these are the committees in which the set of players "is entitled only to vote for or against proposals submitted to it by an external agency" (Laruelle and Valenciano 2008: 53)—success is the better one (Laruelle and Valenciano 2008). We agree but there remains still a difference to our measure of success: Laruelle et al. measure success by a probability, whereas we take the expected distance between a player's ideal points and the equilibrium outcomes.[16] But irrespective of this difference, what Laruelle et al. (2006: 201/203) conclude is worth to be quoted:

> Perhaps the fascination raised by the notion of 'power' has caused a distortion of focus in the field. It can be argued that decisiveness seems intuitively closer to the notion of 'power' than that of success, but this does not grant greater credit to recommendations based on this interpretation. In other words, the relevant question is not what notion is closer to the intuitive idea of 'power', but is a more adequate basis for normative recommendations. And as a base for normative recommendations (e.g., in connection with important issues, as that of the most adequate voting rule in a 'take-it-or-leave-it' committee of representatives) it seems more relevant the notion of success than that of decisiveness.

The upshot of these deliberations is that a decision-making process can be modeled in several ways: as a simple game, using a coalitional or purely probabilistic approach, or as a non-cooperative game. Whether the one or the other is superior depends upon the question to be addressed *and* whether the nature of the decision-making process—for example, the sequential moves of the players, the inter-body decision-making, or the possibility to vote strategically (Schmidtchen and Steunenberg 2002: 206–214)—is adequately captured. To paraphrase Braham and Holler: "Here lies the heart of the problem" (Braham and Holler 2005: 144).

[16]Another difference is worth to be mentioned: Whereas in the Laruelle et al. model proposals are submitted by an external agency, the agenda setter in our model is a player, thinking strategically.

## 4 The SPI: Confounding Power with Luck?

In a much-cited paper entitled "Is it Better to be Powerful or Lucky?" Brian Barry presented the following formula: success = luck + decisiveness (Barry 1980: 338). Although Barry had not been concerned with non-cooperative voting games and, moreover, defined the terms as probabilities, the logic of this formula applies in a modified way to the SPI as well. The modification consists in substituting, first, probabilities by distances between ideal points and equilibrium outcomes, and, second, decisiveness by strategic power. As for the latter substitution, recall that power has nothing to do with decisiveness but refers to the ability of getting desired (equilibrium) outcomes. In order to level out the effect of luck, we focus on the average or expected ability.

This procedure has been criticized by authors who are in favor of power indices quite similar to ours. For example, Napel and Widgrén propose—as we do—a unified framework for measuring power as determined by spatial preferences, strategic agenda setting and decision-making procedures (see Napel and Widgren 2004). Thus, they do not deny the possibility of a preference-based power index. However, they claim that the framework underpinning the SPI leads to a *strategic success index*, rather than a strategic power index. In their view, SPI measures "the ability of a player to make a difference in the outcome", i.e. power, only under very special circumstances (Napel and Widgren 2004: 524): "Only for particular distribution assumptions is luck 'leveled out' by taking averages" (Napel and Widgren 2004: 524), and: "Unless one regards average success as the defining characteristic of power (which neither Steunenberg et al. nor many others do), taking expectations will only by coincidence achieve what Steunenberg et al. aim at, namely' to level out the effect of 'luck' or a particular preference configuration on the outcome of a game'" (Napel and Widgren 2004: 524). Napel and Widgren concede that the SPI "is a good measure of average success but, in general, it fails to capture power" since the SPI confounds luck with power (Napel and Widgren 2004: 524).

We will discuss this critique in turn. Consider Fig. 4, which can be used to illustrate the importance of distinguishing 'power' from 'luck'. The equilibrium outcome of the game is x = C, that is, the most preferred position of the Commission. This outcome seems to be more favorable to Council member 2 than member 5, since the distance to $V_2$ is less than the distance to $V_5$. Is member 2 therefore also more powerful? Both players have the same abilities to affect the outcome, that is, to veto the Commission proposal. So, from this perspective, there is no difference in power. Nevertheless, the outcome is closer to member 2's preferences. This indicates that member 2 is more 'lucky' than member 5. Having a preference that lies close to the equilibrium outcome of a particular game does not necessarily mean that this player is also 'powerful'. Similarly, one may question whether Council member 1 is more 'powerful' than the other Council members, since this player defines the boundary, $V_1(q)$, where the Commission can no longer select its ideal point, should this player move to the right. If any other player can also occupy the position of this member, or the status quo can be located at any other point along the policy dimension, Council member 1 is just more 'lucky' than the others. Following Barry (1980), we regard (in this specific game) a player's *success*, which is defined as the extent to which the outcome of the decision-making process corresponds to its ideal point, as the composite effect of 'power' and 'luck'. Part of a player's success is therefore based on 'luck', the other part is due to the 'power' a player exerts.[17]

Whereas power can be associated with a player's ability to affect the final outcome [which is basically a matter of the rules of the game telling us *who* can do *what* and *when* and who gets *how much* when the game is over (see Binmore 1992: 25)], 'luck' is related to the preferences of the players and the location of the status quo, which are assumed to be exogenously determined. The latter can be illustrated by the role of the Commission in our example of the consultation procedure. The fact that the outcome of the game coincides with the Commission's most preferred point does not imply that the other players in the game are 'powerless'. This result depends on the preferences of the Council members and the location of q. A shift of $V_1$ to the left may, for instance, force the Commission to propose a policy $x = V_1(q)$. Thus, given the preference configuration, the Commission is 'lucky' that Council members have preferences that allow for the equilibrium outcome $x = C$. This clearly indicates that the *success* of a player in a given game is the combined result of *abilities* (defined by the rules of a game) and the specific *preference configuration*. To assess a player's power, a measure should be based on the former and not the latter.

To distinguish 'power' from 'luck', we propose a measure that is independent of the preferences of players in a specific game, which, together with the game form, determines the outcome of the game. This can be achieved by measuring a player's power under some game form with reference to the *mean* or *expected* distance between the equilibrium outcome and this player's ideal point for all possible combinations of players' preferences and all possible combinations of the status quo. In doing so, the power-luck confusion vanishes. The fact that our power

[17] Note the difference between our definition and Barry's definition, which has recently been given more precision by Laruelle and Valenciano (2008: 54–55, 58). In their view a player is successful *ex post*, i.e. once the players have voted on a given proposal, if he/she obtains an outcome—acceptance or rejection of a proposal—that he/she has been voted for. A voter has been decisive if he/she is successful and his/her vote was crucial (critical) to that outcome. Luck is simply success without decisiveness, i.e. a player's vote is irrelevant for the outcome. Thus, Laruelle and Valenciano interpret decisiveness, success and luck as binary variables.

Our definition of terms is more general than Laruelle and Valenciano's, first, in that it refers not only to veto-players but also includes the agenda-setter. Second, in our framework, a player is successful if his/her vote influences the content of the proposal such that the equilibrium outcome moves towards his/her ideal point (including the case in which the status quo remains). Contrary to Laruelle and Valenciano, in our framework a player can be more or less successful, since the distance between the equilibrium outcome and a player's ideal point can vary.

scores turned out to be sensitive to a change of the decision-making procedures (all other things being equal) gives further support to this conclusion.[18]

## 5 The SPI: Can It Become Negative?

Napel and Widgren claim that the SPI may become negative (Napel and Widgren 2004: 524; Napel and Widgren 2002: 9–11). We discuss two examples Napel and Widgren developed in support of their claim (Napel and Widgren 2002: 9–10; Napel and Widgren 2004: 524).

Consider a simple majority voting game with three players having equal voting weight and outcome space $X = \{-1, 0, 1\}$. Player 1's random ideal point, $\lambda_1$ is degenerate and always equal to 0, whereas the ideal points of players $i = \{2, 3\}$, $\lambda_i$, are uniformly distributed on $X$. The status quo is fixed on position 0. In only two out of nine states of the world $\xi = (q, \lambda_1, \lambda_2, \lambda_3)$ with $q = 0, \lambda_1 = 0$, and either $\lambda_2 = \lambda_3 = -1$ or $\lambda_2 = \lambda_3 = 1$ the status quo does not prevail. Since average distance is 2/9 for player 1 and 4/9 for both other players, player 1 appears to be the most successful and most powerful player. Napel and Widgren (2002: 10) conclude: "However, exactly the same equilibrium outcomes prevail when player 1's voting weight is reduced to zero, i.e. if he becomes a dummy player" (assuming that for an even number of players the status quo wins unless defeated by a majority). And they add: "According to Steunenberg et al.'s Strict (!) Power Index, he is still the most powerful player".

What is the reason for this seemingly strange conclusion? Given that the status quo is always q = 0 and the ideal point of player 1, $\lambda_1$, is supposed to be always at the status quo, the set up of the game implies a status quo bias. Therefore it is not surprising to see player 1 coming out as the most "powerful" player. This status quo bias still exists if player 1 has a voting weight of zero, since a majority, in fact unanimity, is needed to defeat it. Player 1 is clearly in both scenarios the most successful player, but not the most powerful. His/her superior performance is due to a restriction of the set of possible states of the world from 81, under the assumptions we would use to calculate the index, to 9, leading to luck for player 1 and bad luck on the side of players 2 and 3.

The SPI is normalized by the introduction of a dummy player. Contrary to Napel and Widgren's approach, this player is not a true player but rather an outside observer. In fact, by assuming player 1 being a dummy player Napel and Widgren transform the three-player game into a two-player game.

[18]Note again the difference between our approach and that proposed by Laruelle and Valenciano (2008: 58). They define the *ex ante version* of the three terms success, decisiveness and luck (irrelevance) using probabilities. The probability of a player being decisive is simply the difference between his/her probability of being successful minus the probability of being lucky.

We define a dummy player as "a player whose preferences vary over the same range as the preferences of the actual players, but that has no decision-making rights in the game" (Steunenberg et al. 1999: 348). Napel and Widgren claim that this definition of a dummy player is not always meaningful:

"What does it mean to 'vary over the same range' if the so-called actual players' ideal points (to stay in a spatial voting framework) have different supports; e.g. $\tilde{\lambda}_i$ is uniformly distributed on [0, 1] and $\tilde{\lambda}_j$ has triangular distribution on [1/2, 4]?" (Napel and Widgren 2002: 11).

The answer is that the expression 'same range' refers to the range in the policy space in which the ideal points of *all* players can be distributed. In the example given by Napel and Widgren it is the range [0, 4].

Next, consider the second example developed by Napel and Widgren with the purpose to illustrate that, contrary to our view (Steunenberg et al. 1999: 349, n. 7), "equilibrium outcomes can be systematically biased against the interest of a particular player" (Napel and Widgren 2002: 11). If so, the SPI can become negative. In a group of four boys, the oldest one is the agenda setter and makes proposals as for what to do in the afternoon. Proposals have to be accepted by a majority of the remaining three boys. All boys have independent preferences, which follow the same distribution. According to the SPI framework the oldest boy as agenda setter is the most powerful player, and the SPI value is the same for the remaining players. There is a little brother of the oldest boy who is allowed to participate in the afternoon activities of the group but does not have a say in selecting the program. Regarding its preferences Napel and Widgren (2002: 10) make the following crucial assumption:

> It is plausible to assume that he does not always agree with his elder brother's most desired outcome, but does so more often than with the others' ideal alternatives. Mathematically speaking, let the ideal points of the two brothers be positively correlated. Then, the mean distance between the group's equilibrium activity and its youngest member's most desired recreation will be smaller than that of those group members who actually have their vote on the outcome.

These examples seem to suggest that the SPI is a rather strange construct: negative power—what sense does that make? But we should be rather careful in drawing such a conclusion: First of all, Napel and Widgren concede that "(f)or a measure of average normalized success this (negative power, the authors) makes sense" (Napel and Widgren 2002: 11): "it simply indicates that (a) player ... is less successful on average than a neutral member of the decision body would be" (Napel and Widgren 2002: 11).[19] Second, note that a negative SPI is due and only due to the introduction of a dummy and not due to the internal logic of the SPI. The introduction of a dummy simply serves the purpose of transferring the players' expected distances into a range of 0 and 1 (normalization). Of course, the SPI makes perfectly sense without such a normalization. Third, even if the power of a player

[19]Napel and Widgren present an example in which a player $n$ always has a position "opposite" of his $n-1$ colleagues (Napel and Widgren 2002: 11).

after normalization turns out to be negative, the relative power of the players having voting rights is still correctly indicated.

However, we could also try an entirely different route of argument. Taking the idea of a veil of ignorance seriously—and this idea is at the heart of measuring a-priori power—one might well ask whether the examples presented by Napel and Widgren are actually to the point. Following Felsenthal and Machover (2001: 94), "to obtain an a priori strategic measure we must go behind a veil of ignorance: we must minimize the information built into the state space and the distribution of the state variables". The crucial question then is: what is the proper assumption regarding the distribution of the ideal points and the status quo? We feel that the veil of ignorance means that there is no information about the players' preferences and the status quo. Therefore, the principle of insufficient reason (or principle of indifference) requires assuming that the state variables are mutually independent and uniformly distributed on the state space.[20] From this point of view, a 'true' veil of ignorance (principle of insufficient reason or indifference) would imply two things, (1) independent distributed ideal points for all players having voting rights and the status quo and (2) the distribution of the dummy being identical to that of all other players. All the examples presented by Napel and Widgren violate this condition. Consider the little brother: despite the fact that little brother does not have a say in the game he is *not* a dummy player in the strict sense. He is simply lucky to have a preference closely related to that of the player that is most powerful. This is also the reason why the SPI would *not* assign power to Luxembourg, to take another example referred to by Napel and Widgren (2002: 11). Luxembourg is lucky having sometimes similar views with the other Benelux countries.

## 6 The SPI: A Banzhaf in Disguise?

In their comment on the 'Symposium Power Indices and the European Union' in the *Journal of Theoretical Politics* Felsenthal and Machover argue that strategic power is simply the Banzhaf power multiplied by a constant that depends on the shape of the state space (see Felsenthal and Machover 2001). If Felsenthal and Machover's conclusion were correct we would take this as support for our view that the SPI is a possible and reasonable measure of power. In the following we, first, present a sketch of the Theorem proven by Felsenthal and Machover, which is then followed by an evaluation of the results.

[20] We agree with Felsenthal and Machover (2001: 94) that this is not sufficient, "because the geometric structure of the state space itself also carries some information. In particular, any asymmetry of this space implies a bias in favour of some states and against others".

### 6.1 The Theorem

Consider a simple voting game ***W***. Let **S** denote a state space which is perfectly symmetric.[21] $X_i, \ldots X_n, Y, Z$ are independent random variables, all of which take their values in the state space. $X_i$, $Y$, $Z$ stand, respectively, for the ideal point of player *i*, the state if a proposed bill will be passed, and the status quo (i.e. the state that continues to prevail if the policy proposal is defeated).

Let $D_i$ denote the distance$|X_i - U|$ between *i*'s ideal point and the preferred state $U = Y$ or $U = Z$. This distance is a function of the random variables $X_i, \ldots X_n, Y, Z$ and can be regarded as a value of a random variable $D_i = f_i(X_i, \ldots X_n, Y, Z)$ which is completely determined by the simple voting game ***W*** and the joint distribution of the state variables (Felsenthal and Machover 2001: 93). As Felsenthal and Machover point out, "from the symmetry of **S** and the assumption that the $X_i$ are independent and uniformly distributed on **S** it follows that the preferred state of each voter is equally likely to be nearer to **Y** than to **Z** as the other way around. Therefore each voter will vote 'yes' or 'no' with probability ½ and they will do so independently of each other—just as in the Bernoulli model underlying the *Bz* measure" (Felsenthal and Machover 2001: 94).

Now, let R and r, respectively, denote the greater and smaller of the two distances $|X_i - Y|$ and $|X_i - Z|$. Then the distance $D_i$ can be defined as

$$D_i = (1 - p) \cdot R + p \cdot r,$$

with $p$ the probability that *i*'s voting decision agrees with the outcome of the vote.

Using Penrose's theorem, which state

$$p = \frac{1 + \beta'(W)}{2},$$

with $\beta'[W]$ the Banzhaf, one can define the mean value of $D_i$, denoted $\Delta_i[W]$,

$$\Delta_i[W] = \frac{1 - \beta'[W]}{2} \cdot R + \frac{1 - \beta'[W]}{2} \cdot r,$$

for player *i*, and

$$\Delta_d[W] = \frac{R + r}{r}$$

for the dummy player. This gives

$$\Psi i[W] = \frac{R - r}{R + r} \cdot \beta_i'[W].$$

[21] Examples are: in the discrete case the set of vertices of a regular polygon or regular polyhedron; in the continuous case a circle or the surface of a sphere of some higher dimension.

Felsenthal and Machover (2001: 95) conclude:

> Thus $\Psi_i[W]$ is simply the *Bz* power of *i* multiplied by a constant that depends on the shape of **S**. Note, in particular, that in the simplest possible case, where **S** consists of just two points, *r* is clearly 0, so in this case $\Psi_i[W] = \beta_i'[W]$ exactly. In our view, this result vindicates the *Bz* measure: not for the first time, a new approach to the measurement of a priori I-power has, yet again, led to $\beta'$. It also suggests that the strategic measure proposed by SS&K is a natural generalization of a priori I-power, which allows the incorporation of additional information, and thus the study of a posteriori voting power.

Felsenthal and Machover believe that our method of measuring power is a promising candidate for a unified approach (Felsenthal and Machover 2001: 96). Since the SPI not only depends on the set of voters and the decision-making rule but also on the choice of the state space and the joint distribution of the state variables, there exists

> an enormous latitude for building into the model all kinds of information concerning the actual state of the world, the kinds of bill to be put to the vote and affinities or disaffinities between voters (Felsenthal and Machover 2001: 93).

## 6.2 Evaluation[22]

We welcome the Felsenthal and Machover approach. It forms a very interesting foundation of the approach presented here, which would allow the SPI to be fully characterized by the set of the axioms the Banzhaf is founded on. This axiomatic characterization would facilitate comparisons with other power measures. Although the theorem proven by Felsenthal and Machover provides for important insights into the logic of the SPI, several comments seem in order.

First of all, we agree that Felsenthal and Machover succeeded in reformulating the algorithm of the SPI as far as simple voting games are concerned. Simple voting games take the proposals to be voted upon as exogenously given. Thus, they can be treated—as in Felsenthal and Machover—as a random variable. However, the most important feature of the SPI namely the strategic interaction and the procedural constraints are not taken into account (see also Napel and Widgren 2002: 12–13; Napel and Widgren 2004: 524): the bills proposed in the SPI framework are not randomly chosen but are the result of strategic thinking along the subgame perfect equilibrium path.

To illustrate, consider Fig. 4. We know already, in equilibrium, the outcome of this game is a legislative policy $x = C$. Here, the outcome of a specific sequential game is partly due to the value of the random variables and partly the result of strategic thinking on the side of all players. It is natural to think about how to introduce this factor in the Felsenthal and Machover set up. One could take account of strategic thinking by restricting the domain of proposed bills in the

[22] We are particularly indebted to Stefan Klößner for several illuminating suggestions.

state space. The question is whether we can find some reasonable equivalent to the equilibrium concept used in non-cooperative game theory. On the other hand, one might conjecture that since proposals depend on the state of the world, including the ideal points of the Commission, and since the state of the world is a probabilistic variable, also the proposals are. In fact, one might even be tempted to apply the terms winning and losing coalitions in the context of a non-cooperative model of a decision-making procedure.[23] If a majority of the players vote in favor of a proposal then one could say that they form a 'winning coalition'. However, one should speak of a 'quasi-coalition' since, as Felsenthal and Machover (2001: 84) rightly mention, "the very term 'coalition', as referring to an arbitrary set of voters, is perhaps somewhat misleading, as it seems to imply conscious coordination". Moreover, contrary to traditional power indices, the SPI takes account of the fact that the propensity to present a proposal and its content depends on the composition of potential winning coalitions. In other words: The agenda setter is looking for a winning coalition such that the distance between its ideal point and the proposal (generating a winning coalition) is smaller than the distance between its ideal point and the status quo. If there is no such a winning coalition the agenda setter remains silent.

Second, the assumption of perfectly symmetric state spaces is very restrictive and reduces the applicability of the theorem considerably.[24] Note, for example, that the only symmetric one-dimensional state space comprises two points. Moreover, those sets are non-convex, which rules out interpreting a convex combination $\lambda x + (1 - \lambda)y$, with $\lambda \in [0, 1]$, as a compromise between $x$ and $y$.

Third, perfect symmetry of a state space implies that yes/no decisions, interpreted as random variables, are stochastically independent. From this it follows that the distribution of ideal points, in nearly all cases, creates a 50 % a priori probability of a yes vote. Although the yes/no decisions of the players are stochastically independent, they are correlated in the following sense: They produce a distribution of the equilibrium outcomes for which the variance is considerably smaller than the variance of the status quo. The reason is that the proposals are less extreme than the status quo, and when player $i$ accepts the proposal it is more likely that player $j$ also does.

Fourth, probability $p$ and, for that, the definition of distance $D_i$, implies that the outcome of the division agrees with the way player $i$ voted (that is, the bill is passed (defeated) and $i$ votes 'yes' ('no') (see Felsenthal and Machover 2001: 94).

Fifth, Felsenthal and Machover are of the opinion that the strategic power measure "is a natural generalization of a priori I-power, which allows the incorporation of additional information, and thus the study of a posteriori voting power" (Felsenthal and Machover 2001: 95). The notion of I-power is that of "power as

---

[23] As done by Widgren and Napel (2001) and Napel and Widgren (2004). See also the discussion in Sect. 3.3.2.

[24] Simulations indicate that in the case of perfectly symmetric state spaces the values of the SPI match those of the Banzhaf index, but they differ considerably for asymmetric state spaces.

*influence*: a voter's ability to affect the outcome of a division of a voting body—whether the bill in question will be passed or defeated" (Felsenthal and Machover 2001: 84). They argue that the notion of I-power "has essentially nothing to do with cooperative game theory or, for that matter, with game theory generally, as it is normally understood. According to this notion, voting behaviour is motivated by *policy seeking*. The action of a given voter does not depend on what other voters may be expected to do, let alone on bargaining and concluding binding agreements with them.

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 (or public bad), which affects all voters, irrespective of how they have voted on that bill" (Felsenthal and Machover 2001: 84).

We agree with Felsenthal and Machover that the SPI has nothing to do with cooperative game theory, but we disagree with Felsenthal and Machover's characterization of the SPI as not being in essence a game theoretic concept. First of all, although in a simple voting game the action of a given voter does not depend on what other voters may be expected to do, it depends on what the agenda setter has done. Second, the action of the agenda setter clearly depends on what s/he expects the other players will do (backwards induction). Third, application of the SPI approach is not restricted to simple voting games but has been applied to interbody decision-making (Steunenberg et al. 1999; Schmidtchen and Steunenberg 2002). Furthermore, in models allowing for the possibility of negotiating, amending or modifying proposals, forming coalitions and linking decisions on different proposals there is even more room for strategic considerations. Finally, the reformulation of the strategic power index, as presented by Felsenthal and Machover, is based on payoffs, since one cannot calculate differences without knowing the ideal points for all players. In fact, the constant with which the Banzhaf index has to be multiplied is a payoff measure. It is implicitly assumed that voters care about distances and that decisions are (rationally) determined by the distance of the ideal point from both the proposed bill and the status quo. These distances are utility measures.

In a comment on our 2002 article Moshe Machover takes up the issue that the distances in the state space can be interpreted as some kind of payoff contradicting the proposition that the SPI is not in essence a game theoretic concept (Machover 2002: 226–227). He thinks "that the contradiction is only apparent, not real" (Machover 2002: 227). This belief follows from his characterization of the model underpinning the calculation of the SPI as consisting of two distinct parts: "The first part is a decision rule, a so-called 'simple' game or 'simple voting game'. The second part consists of a state space and state variables (which are random variables). The decision rule operates in the conventional way: it tells us how the outcome of a division is determined by the way each of the voters vote. The second part of the model serves to model the motivation that leads each of the voters to vote in a particular way" (Machover 2002: 225). Although Machover is right in identifying the two distinct parts, he neglects the crucial fact that the second part not only serves to model the incentives of the voters but also the strategic choice of the agenda setter. Moreover, although Machover explicitly concedes that the geometry of the state space and the distribution of the state variables *is* game theoretic, belongs to non-cooperative game theory and the voters' decision "*may well* be based on a calculation of expected payoff" (Machover 2002: 227), he nevertheless sticks to his position that the model is not game theoretic: "The point is that in the case of I-power ... these motivations and payoffs are exogenous to the decision rule. This is precisely the situation in S&S's model: the decision rule resides in one part of the model, while the motivations and payoffs reside in the other part" (Machover 2002: 227). True, but we cannot see why this ubiquitous feature of models, i.e., consisting of several parts which are conceptually different, deprives the SPI of its game theoretic nature. A good model integrates different parts such that new insights are generated.

Finally, we share Moshe Machover's position

> that a correct method of measuring actual voting power should be organically connected with the method of measuring a priori power. The reason for this is that actual power is the result of a superposition of real-life factors (such as preferences) on the 'bare' decision rule itself. S&S's two-part model does precisely that; and when the contribution of the second part is reduced to nothing, the result is the Penrose measure (Machover 2002: 225–226).

However, reducing the second part to nothing would mean eliminating any strategic element in a power measure. The constant in the formula derived by Felsenthal and Machover simply disappears.

## 7 Conclusion

In this paper we have discussed the critique raised against the SPI as a preference-based power index. Overall, we find that the critique is unfounded: First, the proposition that the SPI is impossible results from playing with semantics. Second, the SPI does not confound power with luck, since taking expectations eliminates luck. Third, the SPI can become negative due and only due to the procedure of normalization. The SPI makes perfectly sense without a normalization. Moreover, taking the veil of ignorance concept seriously, it cannot become negative. Fourth, the attempt to show that the SPI is nothing but a modified Banzhaf and, for this reason, is not game theoretic should be welcomed, since it takes preferences into account and supports our claim for a unified approach to the study of a priori and a posteriori (actual) voting power. However, it neglects any strategic interaction and important procedural features such as, for example, the sequential nature of the game. Voting does not take place in an institutional vacuum. Rules of order exist, which determine the type of proposals or amendments that can be made, and the agenda of the voting process that must be used. Moreover, the voting body may use a committee structure, in which committees—or subsets of voters—discuss and reformulate proposals before they are put to a final vote on the floor. Not only the vote as such, but also these structures determine the extent to which individual players are able to affect the outcome of a vote.

The SPI refers to the ability of a player to make a difference in the outcome of a policy game. This index has many desirable features. First, it can be based on a careful and detailed analysis of some decision-making process in which the preferences of all players and all relevant institutional complexities are taken into account. Second, like traditional voting power indices, the strategic power index measures a priori power. However, in contrast to these indices the strategic power index provides a unified method to study the composite edifice of a priori and a posteriori power as a whole.

**Acknowledgements** We thank Ronald Heiner, Stefan Klößner, Birgit Will and a referee for helpful comments and suggestions.

## References

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

Barry, B. (1980). Is it better to be powerful or lucky. Part I and Part II. *Political Studies, 18*, 183–194. 338–352.

Binmore, K. (1992). *Fun and games. A text on game theory*. D.C. Heath and Company: Lexington.

Braham, M., & Holler, M. (2005). The impossibility of a preference-based power index. *Journal of Theoretical Politics, 17*(1), 137–157.

Buchanan, J. (1990). The domain of constitutional economics. *Constitutional Political Economy, 1*(1), 1–18.

Felsenthal, D., & Machover, M. (1998). *The measurement of voting power*. Cheltenham: Edward Elgar.

Felsenthal, D., & Machover, M. (2001). Myths and meaning of voting power: Comments on a symposium. *Journal of Theoretical Politics, 13*, 81–97.

Garrett, G., & Tsebelis, G. (1997). Why power indices cannot explain decision-making in the European Union. In D. Schmidtchen & R. Cooter (Eds.), *Constitutional law and economics of the European Union* (pp. 11–31). Cheltenham: Edward Elgar.

Garrett, G., & Tsebelis, G. (1999a). Why resist the temptation to apply power indices to the European Union. *Journal of Theoretical Politics, 11*(3), 291–308.

Garrett, G., & Tsebelis, G. (1999b). More reasons to resist the temptation of power indices in the European Union. *Journal of Theoretical Politics, 11*(3), 331–338.

Holler, M. J., & Owen, G. (2001). *Power indices and coalition formation*. Dordrecht: Kluwer.

Laruelle, A., & Valenciano, F. (2008). *Voting and collective decision-making*. Cambridge: Cambridge University Press.

Laruelle, A., Martinez, R., & Valenciano, F. (2006). Success versus decisiveness: Conceptual discussion and case study. *Journal of Theoretical Politics, 18*, 185–205.

Machover, M. (2002). Comment on strategic power in policy games. In M. Holler, H. Kliemt, D. Schmidtchen, & M. E. Streit (Eds.), *Power and fairness* (pp. 225–227). Mohr Siebeck: Tübingen.

Napel, S., & Widgren, M. (2001). Inferior players in simple games. *International Journal of Game Theory, 30*(2), 209–220.

Napel, S., & Widgren, M. (2002). *Strategic power revisited*. CESifo Working Paper, No. 736 (2).

Napel, S., & Widgren, M. (2004). Power measurement as sensitivity analysis. A unified approach. *Journal of Theoretical Politics, 16*(4), 517–538.

Ordeshook, P. (1992). *A political theory primer*. New York: Routledge.

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

Popper, R. (1960). *The poverty of historicism*. London: Routledge/Kegan Paul.

Romer, T., & Rosenthal, H. (1978). Political resource allocation. *Controlled Agendas, and the Status Quo, Public Choice, 33*, 27–43.

Romer, T., & Rosenthal, H. (1979). Bureaucrats versus voters: On the political economy of resource allocation by direct democracy. *Quarterly Journal of Economics, 93*, 561–587.

Schmidtchen, D., & Steunenberg, B. (2002). Strategic power in policy games. In M. Holler, H. Kliemt, D. Schmidtchen, & M. E. Streit (Eds.), *Power and fairness* (pp. 203–224). Tübingen: Mohr Siebeck.

Steunenberg, B., Koboldt, C., & Schmidtchen, D. (1996). Policy making, comitology, and the balance of power in the European Union. *International Review of Law and Economics, 16*, 329–344.

Steunenberg, B., Schmidtchen, D., & Koboldt, C. (1997). Beyond comitology: European policy-making with parliamentary involvement. *Aussenwirtschaft, 52*, 87–112.

Steunenberg, B., Schmidtchen, D., & Koboldt, C. (1999). Strategic power in the European Union: Evaluating the distribution of power in policy games. *Journal of Theoretical Politics, 11*(3), 339–366.

Straffin, P. D. (1994). Power and stability in politics. In R. J. Aumann & S. Hart (Eds.), *Handbook of game theory with economic applications* (Vol. 2, pp. 1127–1151). Amsterdam: Elsevier.

Widgren, M., & Napel, S. (2001). *The power of a spatially inferior player*. The Research Institute of the Finnish Economy, Discussion Paper 771.