Theoretical vs. Empirical Power Indices: Do Preferences Matter?

Badinger, Harald; Mühlböck, Monika; Nindl, Elisabeth; Reuter, Wolf Heinrich

DOI:
10.57938/ca5b8571-da01-4887-a0bc-a18ea07f8edf

Published: 01/01/2013

Document Version:
Publisher's PDF, also known as Version of record

Department of Economics
Working Paper No. 153

Theoretical vs. Empirical Power Indices: Do Preferences Matter?

Harald Badinger
Monika Mühlböck
Elisabeth Nindl
Wolf Heinrich Reuter

June 2013

WIRTSCHAFTSUNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Theoretical vs. Empirical Power Indices: Do Preferences Matter?*

Harald Badinger[†‡] Monika Mühlböck[§] Elisabeth Nindl[†]
Wolf Heinrich Reuter[†]

June 2013

Abstract

This paper considers empirically whether preference-based (empirical) power indices differ significantly from their preference-free (theoretical) counterparts. Drawing on the to date most comprehensive sample of EU Council votes (1993-2011), we use item-response models to estimate the EU27 member states' preferences (ideal points) in a one-dimensional policy space. Their posterior distributions are then used for the calculation of empirical versions of the Banzhaf, the Shapley-Shubik, and other power indices, invoking the concepts of connected coalitions and bloc voting. Our ideal point estimates point to significant differences in member states' preferences, which often translate into significant differences of empirical (versus theoretical) power under individual voting. However, the formation of voting blocs appears to offset differences in countries' ideal points as the bloc size grows. Interestingly, this result does not hold up for the Shapley-Shubik index, whose empirical variant differs from the theoretical one both under individual and bloc voting.

JEL Codes: D72, C71, C72
Keywords: EU Council, Spatial Voting, Power Index

*Financial support by the Austrian Central Bank (OeNB, Anniversary Fund, project number: 14028) is gratefully acknowledged.

[†]Vienna University of Economics and Business, Department of Economics, Althanstrasse 39-45, A-1090 Vienna, Austria. E-mail: harald.badinger@wu.ac.at, elisabeth.nindl@wu.ac.at, wolf.reuter@wu.ac.at.

[‡]Austrian Institute of Economic Research (WIFO), Arsenal, Objekt 20, A-1030 Vienna, Austria.

[§]University of Salzburg, Department of Political Science, Rudolfskai 42, A-5020 Salzburg, Austria. E-mail: monika.muehlboeck@sbg.ac.at.

1 Introduction

While the theory of power indices in voting games is well researched, a key open question that is still subject to debate in the literature is whether power indices should account for actors' preferences (Braham and Holler, 2005; Napel and Widgrén, 2005). This is not only relevant from an academic perspective, but also of high policy relevance when it comes to negotiating revisions of voting rules, since 'empirical' (preference-based) power indices may differ significantly from their theoretical (preference-free) counterparts.

Against this background the striking lack of comprehensive empirical evidence on the relevance of preferences as determinants of actors' voting power is surprising. This paper does not aim at adding to the theoretical controversy about the benefits and drawbacks of preference-based power indices. Rather we analyze empirically whether the inclusion of policy preferences actually leads to a significant shift in actors' voting power.

Our study builds on the to date most comprehensive data set of EU Council votes, providing us with a comprehensive and high quality dataset and a large number of actors (with potentially large variations in policy preferences), while at the same time involving a large number of (27) actors that is (computationally challenging but) still manageable.

Standard measures of voting power such as the Shapley and Shubik (1954) or the Banzhaf (1965) index, defined as an actor's individual ability to influence the outcome of a vote on an unspecified issue, do not consider actors' preferences. Under a probabilistic interpretation, this is reflected in the assumption about random voting behavior, namely that actors' votes are independent of each other (Banzhaf) or that all actors are homogenous in the sense that all have the same probability of voting in favor of a proposal (Shapley-Shubik).

These assumptions, however, are not necessarily consistent with actual (preference-based) voting behavior. Preferences may matter for two reasons: First, actors are likely to form coalitions and vote according to similar preferences. Based on the assumptions of a spatial voting model, Garrett and Tsebelis (1999) presume that only *connected coalitions* will form, i.e., only actors that are aligned next to each other in the policy space will vote together. They argue that classical power indices will thus overstate the power of extremist and understate power of centrist actors.

Second, actors with 'similar' preferences will sometimes form *a priori coalitions* (*voting blocs*) before the actual voting takes place (Malawski, 2004). Since the power of the voting bloc is potentially larger than the sum of block members' voting power, forming voting blocs is a means for countries to increase their power (Widgrén, 1995). Taken toghether, one could argue that only connected coalitions among voting blocs will form in the voting procedure.[1]

These features of voting behaviour, the formation of *voting blocs* and *connected coalitions*, can also be observed in practical politics, e.g., in the EU Council of Ministers, where EU member states often try to form alliances with like-minded counterparts during the negotiations, such that preference-free voting behavior seems unlikely in practice. As a result, the omission of preferences in the calculation of power indices might lead to systematically biased estimates of actors' power (such as EU member states' power in the EU Council).[2]

This paper considers empirically, whether preferences do in fact lead to a shift in voting power. To this end we study empirically, using voting data from the EU Council of Ministers, whether the theoretical (preference-free) power indices of EU member states differ significantly from the empirical (preference-based) ones, which take the formation of voting blocs and connected coalitions into account.

Thereby, we build upon Pajala and Widgrén (2004), who evaluate voting power using empirical versions of the Banzhaf index incorporating actors' preferences, which are based on expert judgments of EU member states' positions in a one dimensional policy space (DEU dataset).[3] Their results suggest that empirical power indices converge to the theoretical power indices as the number of items increases. While these findings are suggestive, the study by Pajala and Widgrén (2004) relies on a rather small data set including information on 45 legislative proposals over the period 1995-2000, arising from the lack of more comprehensive data on preferences of EU member states over specific issues.

Notwithstanding the various benefits of the DEU data set, there are also some caveats: Apart from the unavoidable degree of subjectivity involved in inferring preferences from expert interviews, this approach delivers only point estimates of countries' preferences (positions), without giving information about their variability. However, countries' preferences and the implied preference-based power indices vary over issues and also over time, since they are influenced by domestic and international political developments. To overcome this difficulty, Hagemann (2007, 2008) and Hagemann and Høyland (2008) suggest using Bayesian Ideal Point Estimation (Clinton et al., 2004; Bafumi et al., 2005) based on the logistic item-response model (IRM), which allows

[1] The following terminology will be used throughout the paper: *Preference-free* and *theoretical power indices* are used interchangeably. *Empirical power indices* (or *preference-based power indices*) are used to refer to power indices, where preferences are taken into account through the formation of *a priori coalitions* (*voting blocs*) or through the assumption that only *connected coalitions* (among single actors or voting blocs) will form.

[2] Preference-based power indices have been criticized on theoretical grounds, the key argument being that power is a generic ability determined by the rules of a game and not by individual preferences over outcomes (Braham and Holler, 2005). Another point of discussion is the distinction between decisiveness and luck, where the first refers to an actor's impact on an outcome as a combination of preferences and capabilities, and the latter reflects simply a coincident match of preferences and the voting outcome (Braham and Holler, 2005; Selck and Steunenberg, 2004).

[3] See Thomson et al. (2006)

to obtain the posterior distribution of actors' preferences (ideal points) based on the observed voting behaviour.

The present paper adds to this strand of the literature by using preference (ideal point) estimates for the calculation of empirical power indices, building on a new and more comprehensive dataset of EU Council votes, using a wider range of power indices, and explicitly accounting for the stochastic nature of preference estimates by random sampling and simulation. In particular, it makes the following contributions.

First, we compile a new data set comprising Council voting decisions on 3353 proposals over the period 1993-2011, which is the most comprehensive dataset used in the literature so far. The time coverage enables us to make comparisons of preferences and the implied empirical voting power over time and policy areas.

Second, this dataset is used to estimate EU member states' preferences (and their posterior distribution), which we use as an alternative to the positional data based on expert interviews that are available only for small subsets of voting data. Apart from increasing the data coverage, this approach allows us to account for the uncertainty in estimates of member states' preferences by using Bayesian estimation of logistic item-response models, which yields a posterior distribution of EU member states' preferences rather than only single point estimates.

Third, these estimated ideal points (preferences) are then used to calculate alternative empirical variants of the Banzhaf index (and other measures of power such as the Shapley-Shubik index), invoking the concepts of voting blocs (a priori coalitions) and connected coalitions. Thereby, we explicitly take into account the uncertainty of the preference estimates and calculate empirical power indices based on 1000 draws from the empirical posterior distribution of the ideal point estimates. This yields a distribution of the empirical power indices, whose sample averages can be compared and tested against their theoretical (preference-free) counterparts. The analysis is carried out for both the whole sample, and for subsamples by policy area.

Our findings suggest that EU member states' preferences differ significantly from each other, both for the full sample and for many policy areas. In addition, preferences appear to vary over time, which is reflected in strong changes in the composition of voting blocs over years, i.e., no stable coalitions among EU member states appear to exist over time. Regarding the implications for power, preferenes appear to have two, potentially offsetting effects: First, by ruling out subsets of unconnected coaltions and permutations, they lead to significant power changes under individual voting. However, once we allow for the formation of voting blocs, the difference between the empirical Banzhaf index (and the Johnston, Deegan-Packel, and Holler-Packel index) and its theoretical, preference-free counterpart fades away as the bloc size increases. Interestingly, the Shapley-Shubik Index is the only index, where this result does not hold up and where we find a significant difference between the empirical and theoretical index even under bloc voting.

The remainder of the paper is organized as follows. Section 2 defines the empirical power indices used. Section 3 outlines the application of item-response models to estimate EU member states preferences from observed voting behaviour. Section 4 describes the voting data from the EU Council of Ministers over the period 1993-2011 and presents the estimates of EU member states' ideal points, both for the full sample and by policy area. Section 5 calculates the empirical power indices reflecting countries' preferences, accounting for the uncertainty in the ideal point estimates, and tests for equality with their theoretical (preference-free) counterparts. Section 6 summarizes the main findings and concludes.

2 Theoretical and Empirical Power Indices

In this section, we formally define the concepts of theoretical (preference-free) and empirical (preference-based) power indices that will be used in the quantitative analysis for the EU Council of Ministers. We illustrate these concepts for the Banzhaf (1965) index, following Pajala and Widgrén (2004). Going beyond previous studies, we will then also define empirical counterparts of other commonly used theoretical power indices, namely the Shapley-Shubik index (Shapley and Shubik, 1954), the Johnston index (Johnston, 1978), the Deegan-Packel index (Deegan and Packel, 1978), and the Holler-Packel index (Holler and Packel, 1983).

2.1 The Banzhaf Index

2.1.1 The Theoretical Banzhaf Index

The (normalized) Banzhaf index of country i (BFI_i) gives the share of country i's 'swings', i.e., winning coalitions where the removal of actor i makes them losing, in the swings of all EU member states and is defined as

$$BFI_i = \frac{\sum_{S \subseteq N} [\nu(S) - \nu(S \backslash \{i\})]}{\sum_{j \in N} \sum_{S \subseteq N} [\nu(S) - \nu(S \backslash \{j\})]}, \tag{1}$$

where N is the set of EU member states (n is the number of member states), S is a coalition (formed of s countries), and ν is a function such that $\nu(S) = 1$ if S is a winning coalition, and $\nu(S) = 0$ otherwise.

By construction, it holds that $\sum_{i=1}^{n} BFI_i = 1$. Notice that for our sample period from 1993-2011, the number of EU member states (n) ranges from 12 to 27 (EU12: 1993-1994, EU15: 1995-4/2004, EU25: 5/2004-2006, EU27: as of 2007). For each regime, there is one theoretical Banzhaf index, depending on the rules of the game (voting weights, thresholds) reflected in the indicator function ν. In our empirical analysis, the focus will be on the weighted qualified majority voting under the EU27 as laid down in the treaty of Nice (and prolonged by the treaty of Lisbon), a regime that has been in place since 2007 and will be replaced - after a transitional regime as of 11/2014 - by a system of double majority voting as of 03/2017 (see Section 5).

2.1.2 Empirical Banzhaf Indices

Following Pajala and Widgrén (2004), we define three empirical variants of the Banzhaf index which differ from the theoretical index in equation (1) by taking EU member states' preferences into account: the normal, middle, and boundary variation of the Banzhaf index.

The Normal Variation

The normal variation (BFI^n) differs from the standard Banzhaf index only by the use of voting blocs instead of individual countries. Hence, the definition in equation (1) applies, with the only modification that there are $\bar{i} = 1, \ldots, \bar{n}$ voting blocs rather than $i = 1, \ldots, n$ EU member states, and that coalitions among the set of voting blocs ($\bar{N}$) rather than among the set of EU member states (N) are considered:

$$BFI^n_{\bar{i}} = \frac{\sum_{\bar{S} \subseteq \bar{N}} [\nu(\bar{S}) - \nu(\bar{S} \backslash \{\bar{i}\})]}{\sum_{\bar{j} \in \bar{N}} \sum_{\bar{S} \subseteq \bar{N}} [\nu(\bar{S}) - \nu(\bar{S} \backslash \{\bar{j}\})]}. \quad (2)$$

Since we are ultimately interested in the power of EU member states, the blocs' power has to be distributed among its members. A natural choice is to allocate the blocs total power to its members according to their weights in the bloc, i.e., $BFI^n_i = BFI^n_{\bar{i}} \omega_i$, where ω_i is the weight of country i in the bloc it belongs to, which we calculate as share of country i's votes in the total votes of the bloc.[4]

A simple example modified from Pajala and Widgrén (2004) serves to illustrate the main points. Assume there are four voting blocs, which are aligned in the policy space as A-B-C-D with weights A: 18, B: 8, C: 22, and D: 39. The threshold to pass a proposal is 61. Columns (1) and (2) in Table 1 list all possible coalitions and winning coalitions. Since only bloc D is pivotal in coalition ABCD (turns the losing coalition ABC into a winning one), there are 10 (bloc) swings under the normal variation, and blocs A and B score 1/10, C scores 3/10 whereas bloc D scores 5/10.

The Middle Variation

For the middle variation (BFI^m), only connected coalitions are considered. With voting blocs aligned in the policy space as A-B-C-D, a coalition consisting only of A and C would not be a connected coalition, because the actor in the middle is missing (see Table 1). An example for a coalition that is connected and winning would be ABCD.

[4] For alternative definitions of the distribution of power among bloc members see Alonso-Meijide et al. (2009).

In this coalition, A and D are the 'boundary' actors, while the actors in the middle are the 'centrist' actors. While 'boundary' actors are only defined to be critical if they can in fact swing the vote by leaving the coalition, the 'centrist' actors are defined as being always critical.[5] The formal definition of the middle variation is given by

$$BFI_{\bar{i}}^{m} = \frac{\sum_{\bar{S}_c \subseteq \bar{N}} [\nu_c(\bar{S}_c) - \nu_c(\bar{S}_c \backslash \{\bar{i}\})]}{\sum_{\bar{j} \in \bar{N}} \sum_{\bar{S}_c \subseteq \bar{N}} [\nu_c(\bar{S}_c) - \nu_c(\bar{S}_c \backslash \{\bar{j}\})]}, \tag{3}$$

where $\bar{S}_c$ is a connected coalition, $\nu_c(\bar{S}) = 1$ if $\bar{S}$ is a winning and connected coalition, and $\nu_c(\bar{S}) = 0$ otherwise.

In Table 1, there are three winning and connected coalitions, ABCD, BCD, and CD. Of the boundary actors, D is critical in all of them. B is boundary actor in BCD but not critical. C is a boundary actor in CD and critical. The centrist actors (B and C in coalition ABCD, C in collation BCD) are critical by assumption such that there are 7 swings in total, and bloc B scores 1/7, while blocs C and D score 3/7 each.

– Table 1 –

The Boundary Variation

For the boundary variation (BFI^b), only connected winning coalitions are considered as in the concept of the middle variation. However, centrist actors are defined as not being able to swing a vote at all; the rationale behind this is that it is unthinkable that a centrist leaves the coalition when the actors on both sides of him stay in the coalition. Thus, only boundary actors can have a swing. The formal definition of BFI^b for bloc $\bar{i}$ is given by

$$BFI_{\bar{i}}^{b} = \frac{\sum_{\bar{S}_c \subseteq \bar{N}} [\nu_c(\bar{S}_c) - \nu_c(\bar{S}_c \backslash \{\bar{i}\})] b(\bar{i})}{\sum_{\bar{j} \in \bar{N}} \sum_{\bar{S}_c \subseteq \bar{N}} [\nu_c(\bar{S}_c) - \nu_c(\bar{S}_c \backslash \{\bar{j}\})] b(\bar{j})}, \tag{4}$$

where $b(\bar{i}) = 1$ if $\bar{i}$ is a boundary bloc, and $b(\bar{i}) = 0$ if $\bar{i}$ is a centrist bloc. In our example in Table 1, blocs D and C have 3 and 1 boundary swings respectively and score 3/4 and 1/4.

Notice the the boundary variation does not fit well the concept of the theoretical Banhaf index (that ignores the ordering of the actors) and is more closely related to the Shapley-Shubik index; as Pajala and Widgrén (2004) point out, the boundary variation resembles an empirical variant of the Shapley-Shubik index. Hence, we will consider the results for the boundary variation for completeness, but focus on an exact empirical analog to the Shapley-Shubik (and compare it with the theoretical Shapley-Shubik index) in the Section 5.5.

[5] The rationale behind this is that the coalition would not be connected anymore if a centrist actor leaves.

2.2 Extensions

2.2.1 The Modified Middle Variation

We suggest another variant of the middle variation of the Banzhaf index, which is based on the following consideration regarding the role of centrist actors. Assume that the proposal in the policy space is 'to the right' of all blocks A-B-C-D. Then, if centrist bloc B leaves the coalition, also bloc A will leave (since it is located farther away from the proposal than B), such that only C and D remain. However, in that case actor B is not critical by definition, but only if the coalition of the remaining blocs C and D is losing.

A difficulty in the implementation is that knowledge of the position of the proposal relative to the centrist blocs is required. If the proposals were located to the left of B, then if B leaves, C and D would leave as well, and only bloc A would remain. Hence, without further knowledge, both possibilities have to be considered.

We refer to this index as modified middle variation $BFI_{\bar{i}}^{\check{m}}$, which defines a centrist bloc $\bar{i}$ only as critical, if its withdrawal from the coalition, along with the blocs located to the right of $\bar{i}$, and also to the left of $\bar{i}$, turns the winning coalition into a losing coalition. Formally, we have

$$BFI_{\bar{i}}^{\check{m}} = \frac{\sum_{\bar{S}_c \subseteq \bar{N}} [\nu_c(\bar{S}_c) - \max[\nu_c(\bar{S}_c \backslash \{\bar{i}_{\leq}\}), (\nu_c(\bar{S}_c \backslash \{\bar{i}_{\geq}\})]]}{\sum_{\bar{j} \in \bar{N}} \sum_{\bar{S}_c \subseteq \bar{N}} [\nu_c(\bar{S}_c) - \max[\nu_c(\bar{S}_c \backslash \{\bar{j}_{\leq}\}), \nu_c(\bar{S}_c \backslash \{\bar{j}_{\geq}\})]]}, \tag{5}$$

where $\bar{S} \backslash \{\bar{i}_{\geq}\}$ ($\bar{S} \backslash \{\bar{i}_{\leq}\}$) denotes the coalition $\bar{S}$ excluding $\bar{i}$ and the blocs to the right (left) of $\bar{i}$.

In our example in Table 1, in contrast to the middle variation, for the modified middle variation, bloc B in coalition ABCD is not critical anymore, as the coalition of C and D is winning. Thus, bloc B (as well as A) has no power at all, while C and D each score 1/2. It should be noted that - without information on the location of the proposal and status-quo - none of the two middle variations is preferable on theoretical grounds; hence, it is worth exploring whether the outcome of the two approaches differs from each other.

2.2.2 Individual Voting versus Bloc Voting

Wheras Pajala and Widgrén (2004) apply the aforementioned definitions of the empirical power indices (the middle and the boundary variation) to $\bar{N}$ voting blocs, we will - as a straightforward extension - calculate the empirical power indices not only for voting blocs but also under the assumption of individual voting of the EU member states.

In Pajala and Widgrén (2004), the sole focus on voting blocs is natural, since the EU member states' preferences (positions) based on expert judgments from the DEU dataset (in discret numbers) are either different or exactly the same. This is not the case for our estimates of EU member states' ideal points based on actual voting behaviour, which are defined as continuous variables.

To these empirical power indices under individual voting of EU member states, the definitions in equations (4)-(6) apply, replacing the set of voting blocs $\bar{N}$ by the set of member states N and replacing the coalitions among voting blocs $\bar{S}$ by coalitions among member states S.

2.2.3 Other Power Indices

Other 'Banzhaf-like' Power Indices

Pajala and Widgrén (2004) consider the Banzhaf index and its empirical counterparts only. While the Banzhaf index is one of the most widely used power indices, several closely related ('Banzhaf-like') power indices have been suggested in the literature: the Johnston index (Johnston, 1978), which accounts for the number of critical actors (swingers) in a winning coalition; the Deegan-Packel index (Deegan and Packel, 1978), which considers only minimal winning coalitions (i.e., coalitions, where each actor is critical) and accounts for the number of critical actors, and the Holler-Packel index (Holler and Packel, 1983), which differs from the Deegan-Packel index by ignoring the number of critical actors in the minimal coalition.

The concept of voting blocs and connected coalitions as reflected in the variations of the BFI can be applied to define corresponding variations of these alternative power indices. Here, we do not consider all possible variations, but rather focus on the - in our view - most intuitive concept of the (modified) middle variation of these indices, both under individual voting and under bloc voting. The detailed definitions of all theoretical and empirical power indices considered in the present paper are given in Appendix A.2 of the paper.

The Shapley-Shubik Index

Another widely used measure of power is the Shapley-Shubik index (Shapley and Shubik, 1954). Unlike the aforementioned indices, it is not a straightforward extension of the Banzhaf index, but is conceptually different in taking the ordering of the actors into account; in an alternative interpretation, the Banzhaf and the Shapley-Shubik index are based on different probabilistic assumptions regarding actors' voting behaviour.[6]

The theoretical Shapley-Shubik (1954) index of country i gives the share of orderings (permutations) of the set of n countries, in which country i is pivotal and is defined as

[6] See Straffin (1977) for a discussion of the probabilistic assumptions about actors' voting behaviour underlying the BFI and SSI; Paterson (2005) provides corresponding results regarding the assumptions with respect to the voting polls.

$$SSI_i = \sum_{S \subseteq N} \frac{(s-1)!(n-s)!}{n!}[\nu(S) - \nu(S \backslash \{j\})]. \tag{6}$$

Based on Edelmann (1997) and Perlinger (2000), we derive an empirical counterpart to the Shapley-Shubik index (SSI_i^c), which restricts the set of allowable permutations, accounting for the location of the actors in the policy space. It is given by

$$SSI_i^c = \sum_{S_c \subseteq N} \frac{2^{(|S_c|-2)}\binom{n-|S_c|}{p-1}}{2^{n-1}}[\nu_c(S_c) - \nu_c(S_c \backslash \{i\})]b(i), \tag{7}$$

where, aligning the countries in a one dimensional policy space according to their preferences (ideal points), p is the lower bound of the coalition $S = [p, \dots q]$, i.e., the rank of the 'leftmost' country, and $b(i) = 1$ if i is a boundary actor, i.e., if actor i is either in position p or in position q, and 0 otherwise. A detailed definition of allowable permutations and the derviation of the empirical SSI are provided in Appendix A.3 of the paper.

Again the definition in equation (7) can be applied under individual and bloc voting, replacing the set of voting blocs $\bar{N}$ by the set of member states N and replacing the coalitions among voting blocs $\bar{S}$ by coalitions among member states S.

Summing up, we will calculate and compare the theoretical Banzhaf index with four empirical counterparts under individual voting and bloc voting. In addition, we will also consider the Shapley-Shubik, the Johnston, Deegan-Packel, and Holler-Packel index, and compare them with the empirical counterpart of the middle variation under individual and bloc voting.

3 Item-Response Models and Preference Estimation from Voting Data

The calculation of the empirical power indices defined in Section 2 requires information on the preferences of EU member states. In the following, we outline our approach to estimate EU member states' ideal points in the policy space, following Hagemann (2008), who in turn build on work by Clinton et al. (2004) and Bafumi et al. (2005).

The idea of the preference-based (spatial) voting approach is that actors are ordered according to their preferences in a low-dimensional Euclidian policy space (e.g., but not necessarily, on an ideological left-right scale). The locations of the actors' preferences are denoted as 'ideal points'. It is assumed that actors vote in favor or against a proposal, depending on whether their ideal point is closer to the proposal or the status quo (Steunenberg et al., 1999; Clinton et al., 2004; Napel and Widgrén, 2004).[7]

In the present paper, ideal points of EU member states will be estimated from voting data in the EU Council of Ministers. Hence, there are $i = 1, \ldots, n$ EU member states having voted on $j = 1, \ldots, m$ items (proposals), which results in a binary $n \times m$ matrix $\mathbf{Y} = (y_{ij})$ of member states' individual voting decisions, where y_{ij} indicates whether country i has voted in favor of proposal j ($y_{ij} = 1$) or against proposal j ($y_{ij} = 0$).

Each country's preference is defined as ideal point (or ideal point) θ_i, which is located in a (one-dimensional) policy space. The status quo and the proposal are located in the policy space with positions ψ_j and ζ_j, respectively. Countries are assumed to have a quadratic utility function, which assigns a higher utility to positions that are closer to the country's ideal point in the policy space. Hence, voting in favor of a proposal is associated with utility $U_i(\zeta_j) = -||\theta_i - \zeta_j||^2 + \eta_{ij}$, whereas voting against the proposal is associated with utility $U_i(\psi_j) = -||\theta_i - \psi_j||^2 + \nu_{ij}$, where η_{ij} and ν_{ij} are independent error terms, reflecting uncertainty in judging the relative position of the proposal and status quo, with $E(\eta_{ij}) = E(\nu_{ij}) = 0$ and $Var(\eta_{ij} - \nu_{ij}) = \sigma_j^2$.

Under utility maximizing behavior, the probability that country i votes in favor of proposal j is given by

$$\begin{aligned} P(y_{ij} = 1) &= P(U_i(\zeta_j) > U_i(\psi_j)) \\ &= P(\nu_{ij} - \eta_{ij} < (\theta_i - \psi_j)^2 - (\theta_i - \zeta_j)^2) \\ &= P(\nu_{ij} - \eta_{ij} < 2(\zeta_j - \psi_j)\theta_i + \psi_j^2 - \zeta_j^2) \\ &= P((\nu_{ij} - \eta_{ij})/\sigma_j < \beta_j\theta_i - \alpha_j), \end{aligned} \tag{8}$$

where $\alpha_j = (\zeta_j^2 - \psi_j^2)/\sigma_j$ and $\beta_j = 2(\zeta_j - \psi_j)/\sigma_j$. Hence, the probability of a 'yes'-vote (π_{ij}) depends on country i's ideal point, and the properties of item j relative to the status quo, characterized by the 'difficulty parameter' $\alpha_j = (\zeta_j^2 - \psi_j^2)/\sigma_j$, and the 'discrimination parameter' $\beta_j = 2(\zeta_j - \psi_j)/\sigma_j$, indicating the distance (and direction) between the location of the proposal and the status quo (the no vote) in the policy space.[8]

Following Bafumi et al. (2005), we assume a standard logistic distribution, such that

$$P(y_{ij} = 1) = \pi_{ij}(\theta_i, \alpha_j, \beta_j) = \frac{1}{1 + \exp(\alpha_j - \beta_j\theta_i)} \tag{9}$$

[7] The framework for ideal point estimation is based on item-response models (IRM) that calculate the probability of success of an individual in a test situation based on two factors, the subject's ability and the difficulty of the item (Rasch, 1980).

[8] The discrimination parameter indicates how well an issue differentiates between legislators, with high values corresponding to a strong correlation between the ideal point and the probability of voting as expected. If β_j is equal to zero, the probability of voting 'yes' is solely determined by the underlying distribution of the error terms. The larger β_j, the more an item discriminates among countries and the stronger is the relation between the individual ideal point and the probability of voting 'yes'. In the standard case of a non-zero discrimination parameter, if the ideal point and the difficulty parameter α_j are very close, the actor is indifferent on a certain proposal (Bafumi et al., 2005).

and the likelihood of the implied logit model with (unobserved) regressor θ_i, conditional on the observed voting behavior $\mathbf{Y}$, is given by

$$L(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\alpha}|\boldsymbol{Y}) = \prod_{i=1}^{n}\prod_{j=1}^{m}(\pi_{ij})^{y_{ij}} \times (1-\pi_{ij})^{1-y_{ij}}, \tag{10}$$

where $\boldsymbol{\theta} = (\theta_i)$ is an $n \times 1$ vector and $\boldsymbol{\beta} = (\beta_j)$ and $\boldsymbol{\alpha} = (\alpha_j)$ are $m \times 1$ vectors.

In practice, voting behavior can deviate from pure utility maximization due to outside pressures or strategic considerations. Following Bafumi et al. (2005), this 'non-sincere' voting behavior can be accounted for by introducing error rates δ_0 and δ_1 into equation (10), yielding the following generalized version of the logit model

$$\pi_{ij}(\theta_i, \alpha_j, \beta_j, \delta_0, \delta_1) = \delta_0 + \frac{(1-\delta_0-\delta_1)}{1+\exp(\alpha_j - \beta_j\theta_i)}. \tag{11}$$

The logit model underlying equation (11) cannot be estimated using standard maximum likelihood methods, since we observe only data on voting outcomes. As suggested in Clinton et al. (2004), we will use a simulation-based Bayesian Monte Carlo Markov Chain (MCMC) approach to obtain estimates of the (posterior) distribution of countries' ideal points, conditional on the observed voting data.

The basic idea of this procedure is as follows. If α and β were known, the ideal points could be estimated. If the ideal points were known, α and β could be estimated. The MCMC algorithm repeatedly performs imputations of the unknown parameters and regressions, alternating between the 'estimation' of the ideal points, difficulty, and discrimination parameters, thereby sampling utility differentials from their predictive density (given the current values of the other parameters and the voting data).[9] As a result, we obtain the posterior distribution of the unknown parameters α, β and $\boldsymbol{\theta}$, where our primary interest relates to the posterior distribution of the countries' ideal points $\boldsymbol{\theta}$. Before we turn to the estimation results for the EU Council of Ministers, a brief description of the data is given (Clinton et al., 2004).

4 Preferences of EU Member States: Data and Ideal Point Estimates

4.1 Data

We compiled a new dataset of EU Council votes over the period 1993-2011 by merging existing data sources and collecting new data for the period 2007-2011. In particular, we merged Council voting data from Mattila and Lane (2001) for 1993-1998, from Hayes-Renshaw and Wallace (2006) for 1999-04/2004, and from Mattila (2008) for 05/2004-

[9] See Martin et al. (2011) for a more detailed description of the algorithm.

12/2006, and collected new data for the period 2007-2011 by web scraping of Council documents, using the interface by Buhl and Rasmussen.[10] This adds up to a total of 69195 individual voting decisions of EU member states on 3353 proposals, which is the to date most comprehensive and largest existing database on Council voting. Since unanimous votes contain no systematic information about differences in EU member states' preferences, only contested votes, i.e., votes where at least one member state voted against the proposal or abstained, are considered such that we end up with a total of 16035 voting decisions on 899 proposals.[11]

Of course, merging datasets raises several issues that deserve discussion. First, while our new data for 2007-2011 incorporate only votes on legislative proposals, pre-2007 data also include non-legislative votes (such as Council decisions). Second, data for the pre-1999 period were not publicly available and provided by the Council Secretariat and are thus unlikely to fully cover all Council votes.

These shortcomings of the data have to be borne in mind. However, for the objective of the present paper, namely making use of the voting data to estimate EU member states' ideal points (preferences), we regard these issues of minor relevance. First, there is no strong reason to assume that countries' preferences vary systematically between legislative and non-legislative decisions, such that the potential overrepresentation of legislative votes in the pre-2007 period should not introduce any distortions.

Second, the incomplete data for the pre-1999 period would only pose a problem, if the missing observations were systematically related to countries' preferences. However, there is no evidence for such a systematic exclusion (Mattila and Lane, 2001; Hayes-Renshaw and Wallace, 2006). In sum, given the benefits from the comprehensive period coverage and the large number of observations, we regard the use of a merged dataset as justified and suited with respect to the objective of the present paper.

One general characteristic of Council voting data that might affect the estimation of preferences should be mentioned. Due to the so-called 'culture of consensus' in the Council, proposals typically only reach the voting stage if most of the initial conflicts between countries have been resolved (Heisenberg, 2005). Thus, the number of actual votes against proposals is very low. While such informal agreements and strategic voting behavior is captured to some extent by introducing error rates in equation (11), this is a limitation inherent to all studies using Council votes (Mühlböck, 2011). Collecting data and exploring the nature of the bargaining process that takes place before proposals enter the stage of actual voting remains a challenging and potentially fruitful avenue for future research.

A final issue in coding the voting data is the treatment of abstentions. While

[10]See the application programming interface (API) homepage by Buhl and Rassmussen: `http://api.epdb.eu`.

[11]Excluding non-contested votes also ensures consistency of the dataset, since the Hayes-Renshaw and Wallace (2006) data over the period 1999-04/2004 include contested votes only.

abstentions in voting bodies are often treated either as missing observations or as a third vote choice, within the Council's consensus-seeking culture, an abstention is used to signal that a country holds an opposing view; a diplomatic version of a 'No'-vote (Mühlböck, 2011). Thus, we count each abstention as support of the status quo and disapproval of the common position, i.e., abstentions are coded with 0, as are votes against a proposal.[12] If there are no observed votes of a country (e.g. because they were not eligible to vote, e.g., the UK on issues concerning the Schengen Area or the Euro), these observations are treated as missing data.

– Table 2 –

Table 2 provides an overview of the dataset of contested votes over the period 1993-2011, both for the full sample and for subsamples differentiated by time period and policy areas, showing the number of proposals voted on, the total number of individual voting decisions by member states, and the mean share of votes in favor of an item.

4.2 Ideal Points Estimates for EU27 Member States

In the following, we report the estimates of the item-response model given by equations (10) and (11) for the full sample over the period 1993-2011, yielding (the distribution of) the EU27 member states' ideal points, and then turn to the estimates by policy area.[13]

4.2.1 Results for the Full Sample

Results for all EU27 member states' mean preferences over the full sample period 1993-2011, along with the 95-% confidence intervals are illustrated in Figure 1. Obviously, information on the (simultaneous) voting behavior of all EU27 member states is available only for the most recent period from 2007-2011. Nevertheless, the voting decisions from the pre-accession periods (i.e., before 1995, 2004, and 2007) contain valuable information on the relative positions of the (12, 15, and 25) incumbent EU members states and are thus included in the estimation.

As can be seen from Figure 1, although there is considerable overlap of the confidence intervals of EU member states' ideal points, there are also several 'gaps' between the

[12] In this respect we depart from Hagemann (2008), who consider only abstentions in qualified majority voting (but not abstentions under unanimity) as votes against a proposal. Apart from the view that abstentions can be reasonably argued to reflect (passive) support of the status quo, our approach has the further advantage to generate slightly more variation in the voting data. Since the share of abstentions is small, this choice is not crucial for our results, however.

[13] The (outlier-robust) maximum likelihood estimation of the contaminated hierarchical logistic item-response model that allows for non-sincere voting behavior (equation (11)) was performed using the MCMCirtKdRob function in the MCMCpack library in *R*. A more detailed description of the implementation is given in Appendix A.1.

countries' positions.[14] Northern EU member states such as Sweden, Denmark, the Netherlands and Finland are located on one side of the policy space, whereas southern member states like Spain, France, Portugal, and Italy are located on the opposite side. Moreover, Central and Eastern European member states that joined the EU as of 2004 are located fairly close to each other in the policy space, in between northern and southern EU member states.

Given the multidimensional nature of ideal points and policy proposals and the variation in EU member states' relative positions over proposals and time, the estimates for the one-dimensional policy space should be interpreted with care. The alignment of member states in Figure 1 should not be understood to follow a meaningful policy dimension, e.g. in terms of ideological left-right scale. Rather, ideal point estimates should be taken and interpreted as what they are, namely countries' average preferences over numerous proposals in various policy areas and time periods, still allowing us to judge the proximity of countries reflected in their average voting behavior.

– Figure 1 –

Overall, our mean ideal point estimates are in line with previous studies suggesting a North-South pattern of coalitions in the Council, which has been complemented by an East-West pattern after the EU enlargement in 2004 (Mattila and Lane, 2001; Naurin and Lindahl, 2008; Mattila, 2008; Hagemann, 2008).

To judge whether the differences between EU member states' preferences are also significant in statistical terms, we perform pair-wise Wald tests on the equality of countries' ideal points.[15] In 130 cases (43.2% of all pair-wise tests) the pair-wise test of the hypothesis of equal ideal points is rejected at the 1% level, in 39 (13%) cases at the 5% and in 32 (10.6%) of all cases at the 10% level. The joint test that all EU member states share the same ideal point is rejected at the 1% level.

It should be noted that we move away from a full-fledged Bayesian approach and invoke classical testing methods in order to assess whether EU member states' ideal points (and, subsequently, the implied theoretical and empirical power indices) differ

[14]The fact that all ideal points and confidence intervals are negative follows from equation (11), the identification strategy, and the data set. actor i votes in favor of proposal j if β_j and θ_i have the same sign as this maximizes the value of the likelihood function. As neither the location of a proposal nor the yes- or no-positions are known, we restrict the country with the most/least yes-votes to be on the negative/positive side of the policy space (i.e. -1 and 1). Since we have data only on accepted proposals, it does not come as a surprise to observe only negative values of θ_i and thus β_j, as the yes-position will be closer to the proposal than the no-position.

[15]Refer to the estimate of the $n \times 1$ vector of countries' ideal points as $\hat{\boldsymbol{\theta}} = (\hat{\theta}_i)$ and its (estimated) variance-covariance matrix as $\hat{\boldsymbol{\Omega}} = (\hat{\sigma}_{ij})$, which are calculated as means, variances and covariances of the 10000 posterior 'observations'. The null $H_0 : \boldsymbol{R\theta} = 0$ with restriction matrix $\boldsymbol{R}$ can then be tested using a Wald test given by $\boldsymbol{m'(R\hat{\Omega}R')^{-1}m} \sim \chi^2_r$, where $\boldsymbol{m = R\hat{\theta}}$ is the discrepancy vector and r is the number of restrictions (See, e.g., Greene, 2003, p.95, 487). For a pair-wise test of identical preferences ($H_0 : \theta_i = \theta_j$) the matrix $\boldsymbol{R}$ is a $1 \times n$ vector with elements 1 in the i-th column, -1 in the j-th column, and zeros elsewhere. For a test that all ideal points are the same, the $\boldsymbol{R}$ matrix has $n - 1$ rows (restrictions).

significant from each other.[16] Alternatively, a full Bayesian approach could be pursued, comparing the posterior likelihood of the voting outcomes for restricted models with common ideal points for all (and subsets) of the countries. Such an approach, however, would be computationally very expensive, in particular for the calculation of the empirical power indices. Hence, taking a pragmatic perspective, we regard our approach, which combines Bayesian ideal point estimation to obtain standard errors of countries' ideal points with standard frequentist Wald-type testing procedures, as a reasonable compromise.

4.2.2 Results by Policy Area

To account for the potential multidimensional nature of the policy space and differences between policy areas, we estimate ideal points for different subsamples. The estimates of the subsamples show considerable variation over years and policy areas. Results for some policy areas of particular interest are displayed in Figure 2. The North-South divide obtained in previous studies does appear in some, but not all policy areas. While, e.g., the results for the policy area 'General Affairs' are pretty much in line with the findings in Figure 1 covering the whole dataset, there are no North-South or East-West coalition patterns to be found in the field of 'Health and Consumer Affairs'.

A possible explanation for these differences is that the policy area 'General affairs' consists of topics the Foreign ministers are concerned with, i.e. mostly Foreign Affairs, but also a range of issues not covered by other Council configurations. Due to this diversity of issues within that policy area, it is not surprising that the distribution of ideal point estimates is rather similar to the general distribution in Figure 1.

This is not the case for the policy area 'Health and Consumer Affairs'. Here, the extreme position of Germany might be due to the fact that Germany has already detailed legislation concerning health policy or consumer protection and thus more often supports the status quo (and rejects new EU legislation that potentially requires costly changes of national law) than other EU member states.

– Figure 2 –

Similar reasons might account for the ideal point distribution of the policy area 'Internal Market'. Here no clear distinction between Northern and Southern or Eastern and Western EU member states can be detected. Yet another pattern exists for the policy area 'Agriculture', where we find France and Ireland, two EU member states that are strongly supporting agricultural subsidies, on one side of the spectrum, and the UK and Sweden, who traditionally oppose such a policy on the other side.

– Table 3 –

[16]See Sala-i Martin et al. (2004) for a recent important methodological contribution that combines Bayesian and frequentist features.

Finally, it should be mentioned that not in all policy areas the null hypothesis of identical preferences can be rejected. Table 3 gives the joint tests on equality of the ideal point estimates for the different policy areas, showing that in 5 of 14 policy areas (Agriculture, Trade, Transparency, General Affairs, Internal Market) the hypothesis of equality is rejected at least at the 5% level. In the remaining policy areas the joint test is insignificant.

5 Estimating Empirical Power Indices

The estimates of EU member states' ideal points are used to calculate empirical power indices as defined in Section 2, invoking the concepts of voting blocs and connected coalitions. EU member states' power is calculated according to the present rules governing qualified majority voting laid down in the Nice Treaty (and the Treaty of Lisbon), where the necessary quorum for the adoption of a proposal is defined in terms of a certain share of the weighted votes and the population of the member states.[17]

As outlined in Section 2.2.2, the empirical power indices given by equations (3)-(6) will be calculated 'directly' for single EU member states assuming 'individual voting'; in that case no voting blocs are assumed to be formed a priori, but the preferences are accounted for by considering connected coalitions only. (The normal variation then corresponds to the theoretical power index.)

Moreover, the empirical indices given by equations (3)-(6) are calculated assuming 'bloc voting', where a priori unions (voting blocs) of EU member states with 'similar preferences' are formed before the voting takes place (and where the individual countries' power is calculated from the share of the country's votes in their voting bloc).

Before turning to the results for the empirical, individual- and bloc voting based power indices, we outline how we identify voting blocs from the estimates of the EU member states ideal points and how we account for the uncertainty involved in the ideal point estimates.

5.1 Definition of Voting Blocs

In our analysis, we define two EU member states to be part of the same bloc if their distance in the policy space is below a threshold value, which is determined endogenously

[17]See 'Consolidated versions of the Treaty on European Union and the Treaty on the Functioning of the European Union', Official Journal 2008/C 115/323 for the exact quota and voting weights of the EU27 member states.

by specifying a maximum number of countries that may belong to a bloc.[18]

Thus the threshold distance $\bar{d}_\nu$ and also the composition of the various voting blocs is endogenously determined by the predefined maximum number of countries per voting bloc. Limiting the maximum number of countries in a bloc is motivated by the fact that transaction and coordination costs among member states are increasing with bloc size, such that there is a decreasing marginal benefit of having another member added to a bloc.

As a baseline scenario, we consider a relatively large maximum bloc size of up to 5 countries. However, in order to check the implications of alternative bloc sizes for the estimation results, we will consider alternative values for the maximum number of countries per bloc ranging from 2 to 5.

5.2 Accounting for the Uncertainty in Ideal Point Estimates

To calculate the empirical power indices for the EU27 member states based on the ideal point estimates for voting data over the period 1993-2011, both for the full sample and by policy area, we start by using the mean ideal point estimates, yielding exactly one empirical power index for each country (for the total sample and for each policy area).

To take the uncertainty in the ideal point estimates (reflected in the posterior distribution of the estimates) into account, we do use 1000 repeated independent random draws from the EU member states' empirical posterior distribution of ideal points, and - for each draw - (build voting blocs and) calculate empirical power indices for each EU member state (derived from the power of the voting blocs). This yields an empirical distribution of empirical power indices, from which we calculate the sample average[19] along with its standard deviation. These estimates of empirical power indices will then be compared with and tested against their theoretical (preference-free) counterparts.

5.3 Empirical Banzhaf Indices: Results for Full Sample

Table 4 shows the average empirical Banzhaf power indices defined in equations (2)-(5) in Section 2 (the normal variation, middle variation, modified middle variation, and

[18] More formally, let $d_{ij} = |\hat{\theta}_i - \hat{\theta}_j|$ be the distance between the (mean) ideal point estimates for countries i and j ($i, j \in [1, n]$). $\bar{c}$ is the maximum number of countries per voting bloc. Let $\bar{d}_\nu$ be a threshold distance and $G_{\bar{d}_\nu}(t, p)$ be a set of $\bar{n}$ voting blocs $g_y (y \in [1, \bar{n}])$ corresponding to the threshold distance $\bar{d}_\nu$. Then $i \in g_y \Leftrightarrow \nexists j \in g_x, x \neq y : d_{ij} \leq \bar{d}_\nu$ and $|g_y| \leq \bar{c} \,\forall\, y \in [1, \bar{n}]$. The bloc assignment is implemented by starting with a high value for the threshold (all countries are part of the same bloc), and decreasing the threshold until none of the voting blocs comprises more than the exogenously fixed maximum number of countries per bloc.

[19] By the fundamental law of statistics, we expect the sample mean of the empirical power indices to converge to the one implied by the mean ideal point estimates; in fact, with 1000 draws they turn out virtually identical.

boundary variation) for the EU27 member states and their standard errors.[20]

The first four columns report the 'direct' estimates of countries' empirical power indices, ignoring the formation of voting blocs, whereas the last four columns report the countries' empirical power indices under bloc voting. Notice that the normal variation of the direct indices (first column) equals the theoretical Banzhaf index as defined in equation (1) and thus serves as benchmark.

5.3.1 Results Under Individual Voting

As can be observed from Table 4, we observe that preferences do matter for EU member states' power if we do not take into account bloc building. Most EU member states' empirical power indices, apart from the boundary variation, are statistically different from the theoretical, preference-free Banzhaf index reported in the first column. Almost all the t-tests reject the null hypothesis that the preference-based, empirical power indices are equal to the theoretical preference-free counterpart in the first column.

Hence, it is interesting to explore whether there is some systematic deviation of the point estimates of the empirical power indices from the theoretical ones. The large number of indices and possible comparisons makes general statements difficult. Nevertheless some regularities can be identified from Table 4.

Power under the middle variation differs sizeably from that under the normal variation for several countries. In particular, a number of small EU member states with a centrist position (such as Malta, Slovenia, Slovakia, Hungary) gain, whereas countries with an extremist position (UK, Spain) lose. Country size does not appear to play a crucial role here, since there are also large EU member states (such as Germany) that gain moderately compared with the theoretical, preference-free index in the first column.

Comparing the middle variation defined in equation (3) with its modified variant in equation (5), results in the second and third column turn out virtually identical. The quantitative figures are very close and there is also hardly a difference in the number of EU member states where the empirical index differs from the theoretical one. Apparently, in only a negliglible number of voting constellations does it appear to be the case that the 'exit' of a centrist actor leaves a winning 'left' or 'right' coalition, suggesting that the middle variation, despite its potential theoretical shortcomings compared with its modified counterpart, captures the concept of connected coalitions reasonably well.

Under the boundary variation, most small countries' power is reduced to values close to zero and is shifted to a few large EU member states such as France, Germany, Italy, Spain, and the UK. However, it should be borne in mind that most indices are not significantly different from the theoretical Banzhaf index in statistical terms.

[20]We also carried out the same analysis for the EU15 using ideal point estimates based on voting decisions over the period 1995-2004 and obtained qualitatively very similar results. Hence, the results for the EU15 are omitted for the sake of brevity.

Moreover, from a theoretical perspective, the Banzhaf index does not appear to fit well the concept of a boundary actor (see Section 2.1.2), which is more closely related to the Shapley-Shubik index, whose empirical counterpart will be considered in Section 5.5 below.

5.3.2 Results Under Bloc Voting

As can be see from the right part of Table 4, the picture changes when the individual countries' power derived from the power of the voting blocs is considered. With a maximum bloc size of 5 countries, the empirical power indices do not differ significantly from the theoretical Banzhaf index except for Sweden and the United Kingdom, a result which is most likely due to sampling variation; with a significance level of 5%, the result that 2 out of 27 countries (7%) show a significant deviation is not unexpected.

We expect the result to depend on the size of the voting blocs. Hence, we calculate individual countries' voting power derived from voting blocs with different numbers of maximum bloc size, ranging from 2 to 5. Consider the middle variation, which is the empirical index showing the largest number of significant deviations from the theoretical one. in the extreme case with at most 1 country per bloc (individual voting), the middle variation of the Banzhaf index turned out significantly different for 20 countries (or 74% of all 27 EU member states) at the 5% level. With a maximum group size of 2, the number of countries with significant deviations drops to 7 (26% of all 27 EU member states). Increasing the maximum bloc size further to 3 or more countries, the number of countries with significant devations is reduced to 2 (7% of all 27 EU member states), which is roughly equal to the Type I error with a significance level of 5%.

The question of which is the 'right' bloc size will vary over issues and cannot be answered in gerneral terms. However, a general finding is that prefercs appear to have two potentially offsetting results on countries' empirical power. On the one hand, they introduce significant differences between countries' empirical power indices and their theoretical counterparts under individual voting. On the other hand, by leading to the formation of voting blocs, they also reduce number and heterogeneity of actors by merging their prefercs - apparently by more than their standard deviation - such that the differences between the empirical power indices and their theoretical counterparts fade away in statistical terms under bloc voting as the maximum bloc size increases.

– Table 4 –

5.4 Empirical Banzhaf Indices: Results by Policy Area

Table 5 reports summary results for the empirical power indices, both for the full sample and also by policy area under bloc voting.[21] In particular, it gives - for each variant of the empirical power index - the root means squared error (RMSE) of the difference between the theoretical and empirical power indices over all 27 EU member states and the number of countries where the deviations of the empirical power index from the theoretical one turned out significant.

– Table 5 –

The RMSE is larger for the policy areas than on average for the full sample as far as the normal variation is concerned. Regarding the other variations of the index, the deviation is fairly small with values around 3% and tends to be slightly larger for some policy areas (compared with the full sample), in particular for those, where the ideal point estimates indicate significant differences in the EU member states' preferences (Agriculture, General Affairs, Internal Market, Trade, Transparency; see Table 3).

Finally, also for the policy areas, the *t*-tests reject the equality of theoretical and empircal power indices only in a very small number of cases, suggesting once more that the deviation of the empirical power indices (and the role of preferences) is moderate at best. This is a result we expect to hold if voting blocs change frequently and unsystematically from proposal to proposal, such that the effect of allowing for connected coalitions is averaged out over a large number of proposals under bloc voting.

To provide some informal evidence on this hypothesis, we estimate EU member states' ideal points for each year over the period 1993-2011, and consider the bloc formation for each year to see whether there are stable voting blocs (of a least two countries) over time. It turns out that no such stable coalitions can be identified, pointing to large variations in the relative ideal points from year to year (proposal to proposal) and hence a large variation in the size and composition of voting blocs. This result is also in line with Pajala and Widgrén (2004), who argue that no stable minimal coalitions exist over time.

5.5 Results for other 'Banzhaf-like' Indices and the Shapley-Shubik Index

Finally, we also estimate empirical variants under individual and bloc voting of other power indices (see Section 3.5.3). Table 6 displays the theoretical indices and their

[21]For computational reasons, we did not consider individual voting for all policy areas, but some explorative calculations suggest no qualitative differences compared with the full sample: Under individual voting, the empirical indices (in particular the middle and modified middle variation) differ significantly from the theoretical ones for many EU member states, while they hardly do under bloc voting with a maximum bloc size of 3 or more countries.

empirical counterparts defined in Appendix A.2, both under individual voting and under bloc voting with a maximum bloc size of 5 countries.

– Table 6 –

In line with the results for the Banzhaf index and its empirical variations, the picture for the alternative 'Banzhaf-like' power indices is very similar as evident from the left and right part of Table 6. Under individual voting, empirical power indices are different for a large share of the EU27 member states under individual voting, but this difference vanishes under bloc voting with a maximum bloc size of 5 countries. In contrast, the empirical SSI remains significantly different from its theoretical counterpart even under bloc voting.

5.6 Empirical Power and The Role of Bloc Size

Given the sensitivity of the results with respect to the maximum bloc size obtained for the BFI in Section 5.3, we repeat the calculation for all other 'Banzhaf-like' indices (JNI, DPI, HPI) and the empirical variant of the SSI for alternative bloc sizes. Table 7 summarizes the results for alternative maximum bloc sizes, ranging from 2 to 5, and provides the share of countries (out of the 27 EU member states) for which the empirical power indices differ from their theoretical counterparts.

– Table 7 –

Interestingly there is a apparent difference between the Banzhaf (and the 'Banzhaf-like' indices) compared with the SSI. For the BFI, JNI, DPI, and HPI, the differences between empirical and theoretical power indices fades away with increasing bloc size. Depending on the index, the threshold size, where the share of significant coefficients is roughly equal to the Type I error, ranges from 3 to 4. In stark contrast, the power implied by the empirical SSI, which differs from its theoretical counterpart for virtually all EU member states under individual voting, turns out to be rather insensitve against the formation of voting blocs. Even with a maximum bloc size of 5 countries, the share of significantly different indices still amounts to 37%.

This adds another interesting dimension where the results between the SSI and the BFI (and related indices) differ strongly from each other. While changes of single EU member states' power implied by treaty reform are often very similar in terms of both indices, the proportionality of the voting system and its efficiency is typically much larger for the SSI.[22] Our results suggest a further difference, namely regarding the sensitivity of the theoretical measures against the consideration of preferences in general, and the implications of the formation of voting blocs in particular. The central

[22]Paterson (2005) discusses the large differences between efficiency implied by the Banzhaf and the Shapley-Shubik approach in the EU Council of Ministers after the treaty of Nice.

role of pivotal boundary actors inherent in the logic of the SSI appears to be decisive for the results even under bloc voting.

6 Conclusions

This paper tests whether differences in actors' preferences translate into significant differences between theoretical power and empirical power indices, using voting data from the EU Council of Ministers over the period 1993-2011. To do so we consider empirical variants of the Banzhaf index and the Shapley-Shubik index, which take countries' preferences into account by restricting the set of allowable coalitions and permutations (depending on the alignment of the countries in the policy space) and by allowing for the formation of a prioiri unions and bloc voting.

EU member states' preferences in a one-dimensional policy space are calculated using logistic-item response models, which provide estimates of countries' average preferences over the sample period and their distribution. Using random draws of countries' preferences from this posterior distribution yields a posterior distribution of implied empirical power indices, both under individual voting and bloc voting, whose average is then tested for equality with the respective theoretical, preference-free index.

Results from the ideal point estimation show that EU member states' preferences do in fact differ significantly from each other, both for the full sample and for many policy areas. For a large share of EU member states, these differences in preferences translate into significant differences of individual countries' empirical power. However, the formation of voting blocs appears to offset to some extent differences in the countries' ideal points in the policy space. With growing size of the voting blocs, the difference between the empirical Banzhaf indices, derived from voting blocs with a maximum bloc size of more than 3 member states, and the theoretical, preference-free Banzhaf index fades away. Interestingly, this result does not hold up for the Shapley-Shubik index, whose empirical variant differs from the theoretical both under individual voting and bloc voting.

Overall, our estimates suggest that preferences matter, though with two apparently offsetting effects. One the one hand, they rule out certain coalitions (among countries remote from each other in the policy space), thereby leading to a change in empirical power under individual voting. The magnitude and direction of the change will depend not only on country size but also on the countries' position in the policy space, which may change over time. On the other hand, the formation of voting blocs reduces the number of and the heterogeneity among actors, and thereby the difference between empirical and theoretical power indices. Since preferences, the formation of voting blocs, as well as their size and composition, will vary over issues and time as can be observed in the EU Council of Ministers, general statements would be be misleading.

However, a general result is that preferences - besides the rules of the voting game - can be an important determinant of actors' power.

References

Alonso-Meijide, J., Bowles, C., Holler, M. and Napel, S. (2009), 'Monotonicity of Power in Games with a priori Unions', *Theory and Decision* **66**, 17–37.

Bafumi, J., Gelman, A., Park, D. and Kaplan, N. (2005), 'Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation', *Political Analysis* **13**(2), 171–87.

Banzhaf, J. I. (1965), 'Weighted Voting Doesn't Work: A Mathematical Analysis', *Rutgers Law Review* **19**(1965), 317–43.

Braham, M. and Holler, M. (2005), 'The Impossibility of a Preference-Based Power Index', *Journal of Theoretical Politics* **17**(1), 137–157.

Clinton, J., Jackman, S. and Rivers, D. (2004), 'The Statistical Analysis of Roll Call Data', *American Political Science Review* **98**(2), 355–70.

Deegan, J., J. and Packel, E. (1978), 'A New Index of Power for Simple N-Person Games', *International Journal of Game Theory* **7**, 113–123.

Edelmann, P. H. (1997), 'A Note on Voting', *Mathematical Social Sciences* **34**, 37–50.

Garrett, G. and Tsebelis, G. (1999), 'Why Resist the Temptation to Apply Power Indices to the European Union?', *Journal of Theoretical Politics* **11**(3), 291–308.

Greene, W. H. (2003), *Econometric Analysis*, 5th edn, Pearson Education.

Hagemann, S. (2007), 'Applying Ideal Point Estimation Methods to the Council of Ministers', *European Union Politics* **8**(2), 279–296.

Hagemann, S. (2008), Voting, Statements and Coalition-Building in the Council from 1999 to 2006, *in* D. Naurin and H. Wallace, eds, 'Unveiling the Council of the European Union: Games Governments Play in Brussels', Palgrave Macmillan, pp. 36–64.

Hagemann, S. and Høyland, B. (2008), 'Parties in the Council?', *Journal of European Public Policy* **15**(8), 1205–1221.

Hayes-Renshaw, F. and Wallace, H. (2006), *The Council of Ministers*, 2nd edn, Palgrave Macmillan, Houndmills, Basingstoke, Hampshire; New York.

Heisenberg, D. (2005), 'The Institution of 'Consensus' in the European Union: Formal versus Informal Decision-Making in the Council', *European Journal of Political Research* **44**(1), 65–90.

Holler, M. J. and Packel, E. W. (1983), ‘Power, Luck and the Right Index’, *Journal of Economics* **43**, 21–29.

Johnston, R. J. (1978), ‘On the Measurement of Power: Some Reactions to Laver’, *Environment and Planning A* **10**, 907–914.

Malawski, M. (2004), ‘’Counting’ Power Indices for Games with A Priori Unions’, *Theory and Decision* **56**, 125–140.

Martin, A. D., Quinn, K. M. and Park, J. H. (2011), ‘MCMCpack: Markov Chain Monte Carlo in R’, *Journal of Statistical Software* **42**(9), 1–21.

Mattila, M. (2008), Voting and Coalitions in the Council after the Enlargement, *in* D. Naurin and H. Wallace, eds, ‘Unveiling the Council of the European Union: Games Governments Play in Brussels’, Basingstoke: Palgrave Macmillan, pp. 35–68.

Mattila, M. and Lane, J.-E. (2001), ‘Why Unanimity in the Council?: A Roll Call Analysis of Council Voting’, *European Union Politics* **2**(1), 31–52.

Mühlböck, M. (2011), Biased Data? Deducing Policy Preferences from ‘Voting” in the Council of the European Union, Paper presented at the ECPR Joint Sessions, St. Gallen, 12-17 April 2011.

Napel, S. and Widgrén, M. (2004), ‘Power Measurement as Sensitivity Analysis: A Unified Approach’, *Journal of Theoretical Politics* **16**(4), 517–538.

Napel, S. and Widgrén, M. (2005), ‘The Possibility of a Preference-Based Power Index’, *Journal of Theoretical Politics* **17**(3), 377–387.

Naurin, D. and Lindahl, R. (2008), East-North-South. Coalition-Building in the Council Before and After Enlargement, *in* D. Naurin and H. Wallace, eds, ‘Unveiling the Council of the European Union: Games Governments Play in Brussels’, Basingstoke: Palgrave Macmillan, pp. 100–125.

Pajala, A. and Widgrén, M. (2004), ‘A Priori versus Empirical Voting Power in the EU Council of Ministers’, *European Union Politics* **5**(1), 73–97.

Paterson, I. (2005), A Lesser Known Probabilistic Approach to the Shapley-Shubik Index and Useful Related Voting Measures. Paper presented at the EPCS, Durham, March 31 - April 3, 2005.

Perlinger, T. (2000), ‘Voting Power in an Ideological Spectrum - The Markov-Polya Index’, *Mathematical Social Sciences* **40**(2), 215–226.

Rasch, G. (1980), *Probabilistic Models for Some Intelligence and Attainment Tests (Expanded edition with foreword and afterword by B.D. Wright)*, 2nd edn, Chicago: The University of Press.

Rivers, D. (2003), Identification of Multidimensional Spatial Voting models. Stanford University.

Sala-i Martin, X., Doppelhofer, G. and Miller, R. I. (2004), 'Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) approach', *American Economic Review* **94**(4), 813–835.

Selck, T. J. and Steunenberg, B. (2004), 'Between Power and Luck: The European Parliament in the EU Legislative Process', *European Union Politics* **5**(1), 25–46.

Shapley, L. and Shubik, M. (1954), 'A Method for Evaluating the Distribution of Power in a Committee System', *The American Political Science Review* **48**(3), 787–792.

Steunenberg, B., Schmidtchen, D. and 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. (1977), 'Homogeneity, Independence, and Power Indices', *Public Choice* **30(1)**, 107–18.

Thomson, R., Stokman, F. N., Achen, C. H. and König, T. (2006), *The European Union Decides*, Cambridge: Cambridge University Press.

Widgrén, M. (1995), 'Probabilistic Voting Power in the EU Council: The Cases of Trade Policy and Social Regulation', *Scandinavian Journal of Economics* **97**(2), 345–56.

Appendix

A.1 Estimation of EU Member States' Ideal Points

Following Hagemann (2008), we estimate the parameters of the one-dimensional contaminated hierarchical logistic item response model (IRM) in equation (10) with maximum likelihood, using the MCMCirtKdRob function in the MCMCpack library in *R*.[23] The Bayesian estimation of the likelihood function uses a Markov Chain Monte Carlo (MCMC) simulation method that draws from a prior distribution, estimates the parameters and updates the prior distribution. In particular, the MCMC algorithm draws 110000 times from the (updated) prior distribution. The first 10000 draws are discarded to find an arbitrary starting point in the country-item-space, then the maximum likelihood estimates of the mean and the variance of the parameters of each 10^{th} draw are recorded, yielding a posterior distribution made up of 10000 observations.

We assign non-informative priors to avoid influencing the posterior distribution with subjective prior beliefs.[24] Following Martin et al. (2011), country i's ideal point θ_i is assumed to follow a standard normal distribution, for the item parameters α_j and β_j the (normal) prior distribution $[\alpha_j, \beta_j]' \sim \mathcal{N}_{(K+1)}(b_{0,j}, B_{0,j})$ is assumed, where $B_{0,j}$ is a diagonal matrix for the prior precision of the independent normal prior on the item parameters. Like Hagemann (2008) we set the precision equal to $1/\sigma^2 = 0.25$ which implies a priori variance equal to 4. The error rates δ_0 and δ_1 are estimated from the data and are assumed to follow an independent uniform $(0, 0.1)$ prior distribution.

Finally, for identification constraints on θ_i have to be imposed; in a one-dimensional policy space two constraints are sufficient for identification (Clinton et al., 2004). In particular, we deduce the restrictions from the data and restrict the country with the largest share of yes-votes to be negative and the one with the largest share of no-votes to be positive. However, the posterior distributions turn out to be insensitive to the choice of identifying restrictions, given that our model is locally identified by two arbitrary constraints (Rivers, 2003).

A.2 Definition of Theoretical and Empirical Power Indices

Table A1 gives an overview of the power indices considered in the paper. The first column reports the standard definitions of the theoretical (preference-free) indices. The second column reports the empirical power indices, accounting for preferences by considering only connected coalitions (BFI, JNI, DPI, HPI) or permutations that constitute a maximal chain (SSI). We adopt the following notation:

[23] We would like to thank Bjørn Høyland for sharing his R-code (see `http://folk.uio.no/bjornkho/`).

[24] Clinton et al. (2004) find that the results for voting in the US senate and the US House of Representatives in general appear to be insensitive to the choice of the prior.

N is the set of EU member states ($n = |N|$ is the number of member states), and S is a coalition (formed of $s = |S|$ countries). The function ν is defined such that $\nu(S) = 1$ if S is a winning coalition, and $\nu(S) = 0$ otherwise. S_c is a connected coalition, i.e., a coalition comprising only actors that are aligned next to each in the policy space. The function ν_c is defined such that $\nu_c(S_c) = 1$ if S_c is a winning *and* connected coalition, and $\nu_c(S_c) = 0$ otherwise.

The term $\kappa(S)$ denotes the number of critical actors in a winning coalition S (required for the calculation of JNI); $\mathcal{M}(\nu)$ is the set of all minimal winning coalitions, $m(\nu) = |\mathcal{M}(\nu)|$ is the number of minimal winning coalitions, and $S \ni i$ indicates that the coalition S contains actor i (required for the calculation of the DPI and HPI). Finally, p is the lower bound and q is the upper bound of coalition $S = [p, \ldots, q]$, i.e., the 'leftmost' and 'rightmost' countries in a coalition (in terms of their ideal points in the one-dimensional policy space), $b(i) = 1$ if i is a boundary actor, i.e., actor i is either in position p or in position q, and 0 otherwise (required for the calculation of the SSI).

Under bloc voting the same definitions apply, with the only modification that there are $\bar{i} = 1, \ldots, \bar{n}$ voting blocs rather than $i = 1, \ldots, n$ EU member states, and (connected) coalitions $\bar{S}$ of voting blocs among the set of voting blocs $\bar{N}$ rather than coalitions S among the set of individual member states N are considered.

Table A1: *Definition of Power Indices*

	Theoretical Power Index	Empirical Power Index
BFI_i	$\frac{\sum_{S\subseteq N}[\nu(S)-\nu(S\backslash\{i\})]}{\sum_{j\subseteq N}\sum_{S\subseteq N}[\nu(S)-\nu(S\backslash\{j\})]}$	$\frac{\sum_{S_c\subseteq N}[\nu_c(S_c)-\nu_c(S_c\backslash\{i\})]}{\sum_{j\subseteq N}\sum_{S_c\subseteq N}[\nu_c(S_c)-\nu_c(S_c\backslash\{j\})]}$
SSI_i	$\sum_{S\subseteq N}\frac{(\|S\|-1)!(n-\|S\|)!}{n!}[\nu(S)-\nu(S-\{i\})]$	$\sum_{S_c\subseteq N}\frac{2^{(\|S_c\|-2)}\binom{n-\|S_c\|}{p-1}}{2^{n-1}}[\nu_c(S_c)-\nu_c(S_c\backslash\{i\})]b(i)$
JNI_i^*	$\sum_{\substack{S\subseteq N\\ S\ni i}}\frac{1}{\kappa(S)}[\nu(S)-\nu(S\backslash\{i\})]$	$\sum_{\substack{S_c\subseteq N\\ S_c\ni i}}\frac{1}{\kappa(S_c)}[\nu_c(S_c)-\nu_c(S_c\backslash\{i\})]$
DPI_i^*	$\frac{1}{m(\nu)}\sum_{\substack{S\in\mathcal{M}(\nu)\\ S\ni i}}\frac{1}{S}[\nu(S)-\nu(S\backslash\{i\})]$	$\frac{1}{m(\nu_c)}\sum_{\substack{S_c\in\mathcal{M}(\nu_c)\\ S_c\ni i}}\frac{1}{S_c}[\nu_c(S_c)-\nu_c(S_c\backslash\{i\})]$
HPI_i^*	$\frac{1}{m(\nu)}\sum_{\substack{S\in\mathcal{M}(\nu)\\ S\ni i}}[\nu(S)-\nu(S\backslash\{i\})]=\frac{m_i(\nu)}{m(\nu)}$	$\frac{1}{m(\nu_c)}\sum_{\substack{S_c\in\mathcal{M}(\nu_c)\\ S_c\ni i}}[\nu_c(S_c)-\nu_c(S_c\backslash\{i\})]=\frac{m_i(\nu_c)}{m(\nu_c)}$

* *Notes*: Definition of absolute indices; all indices are normalized such that they sum to 1. The variation of the indices considered is determined by the definition of the function ν.

A.3 A Preference-Based Shapley-Shubik Index

A.3.1 The Edelmann/Perlinger Index

Perlinger (2000) derives a preference-based Shapley-Shubik Index, based on a framework by Edelmann (1997), which restricts the number of allowable permutations to so called 'maximal chains' (M), defined as permutation of the n actors in the ('spectrum') game, where the actor in position i (of the permutation) is the ideological neighbour to an actor in position $1, 2, \ldots, i-1$ for all $i \in N$.[25] It is given by

$$\sum_{M}\frac{1}{m}[\nu(M^i)-\nu(M^i\backslash\{i\})], \tag{A.1}$$

where $m = 2^{n-1}$ denotes the total number of allowable permutations (maximal chains)

[25]Two member states are called ideological or preference-based neighbors if there are no other member states who, based on their preference estimates, are in a position between them.

in a spectrum game with n actors, and M^i denotes a coalition consisting of i and its predecessors in the maximal chain M. Finally, $\nu(M^i) = 1$ if M^i is a winning coalition and $\nu(M^i) = 0$ otherwise.

The reasoning behind equation (A.1) is that - as an equivalent to the empircal Banzhaf Index which allows only connected coalitions - a notion of allowable permutations of actors of the set N is required for an empirical Shapley-Shubik Index. Hence, the Edelmann/Perlinger-Index allows only those permutations of N that consist of connected coalitions at any point. For example, assume that $N = \{A, B, C\}$ and that actors are aligned as A-B-C in a one-dimensional policy space. Then, $\{B, C, A\}$ would be allowed, but $\{A, C, B\}$ would not, as $\{A, C\}$ is not a connected coalition. These allowable permutations are exactly the maximal chains.

A.3.2 An Alternative Definition of the Empirical Shapley-Shubik Index

For computational reasons - the number of maximal chains increases very quickly with the number of actors n - we derive an alternative formula (which is equivalent to Edelmann/Perlinger-Index), which does not sum over maximal chains but over connected coalitions (as we do for the empirical Banzhaf indices) and which also more directly reveals the relation between the theoretical and the empirical SSI.

The derivation builds on the definition of the standard Shapley-Shubik index, which gives the share of all permutations of actors of the set N, where actor i is pivotal:

$$SSI_i = \sum_{S \subseteq N} \frac{ab}{c} d = \sum_{S \subseteq N} \frac{(|S| - 1)!(n - |S|)!}{n!} [\nu(S) - \nu(S \backslash \{i\})] \qquad \text{(A.2)}$$

Note that in equation (A.2) the term $a \equiv n!$ gives the total number of orderings and $d \equiv [\nu(S) - \nu(S\backslash\{i\})] = 1$ if actor i is pivotal. Fixing the position of actor i, the number of orderings of actors preceeding actor i is given by $a \equiv (|S| - 1)! = (s - 1)!$ and the number of orderings of the actors following actor i is $b \equiv (|N| - s) = (n - s)!$.

In the following we provide corresponding defintions of the terms a, b, c, and d restricting the orderings to the number of allowable permutations, referred to as a_c, b_c, and c_c and sum over connected coalitions only. The number of maximal chains with actor i exactly in position $|S_c| = s$ is given by $a_c b_c$. Thereby, a_c equals the number of options to build a chain of the elements of S_c (a so-called *saturated chain* from $\emptyset$ to S_c in the notation of Edelmann (1997)), where actor i joins last. According to *Lemma 1* in Edelmann (1997),

$$a_c = 2^{(|S_c|-1)-1)} = 2^{|S_c|-2} \qquad \text{(A.3)}$$

The term b_c equals the number of options how to extend this saturated chain to a maximal chain, i.e. to allowably order the remaining actors of the set N (or, in other words, the number of saturated chains from S_c to $N = [1, \ldots, n]$). According to *Lemma*

2 in Edelmann (1997)

$$b_c = \binom{n - |S_c|}{p - 1}, \tag{A.4}$$

with p being the lower bound of $S_c = [p, \ldots, q]$.

As an equivalent to the number of permutations $a \equiv n!$ in the original Shapley-Shubik Index, the number of maximal chains in a spectrum game with n actors is given by $c_c = 2^{n-1}$ according to *Lemma 1* in Edelmann (1997).

Finally, only winning *and* connected coalitions, where actor i is pivotal (which is only the case if actor i is a boundary actor, i.e. either in position p or q of $S_c = [p, q]$) have to be considered, which is accomplished by redefining term d as

$$d_c = [\nu_c(S_c) - \nu_c(S_c \backslash \{i\})]b(i), \tag{A.5}$$

where $b(i) = 1$ if i is a boundary actor. Summing up, we arrive at the following equivalent defintion of the empirical Shapley-Shubik index:

$$SSI_i^c = \sum_{S_c \subseteq N} \frac{2^{(|S_c|-2)} \binom{n-|S_c|}{p-1}}{2^{n-1}} [\nu_c(S_c) - \nu_c(S_c \backslash \{i\})]b(i), \tag{A.6}$$

where S_c is a connected coalition, $\nu_c(S_c) = 1$ if S is a winning and connected coalition, p is the lower bound of $S_c = [p, q]$, and $b(i) = 1$ if i is a boundary actor, i.e. actor i is either in position p or in position q and 0 otherwise.

Table 1: *Illustration of Preference-based (Empirical) Power Indices*

(1) Coalition $\bar{S}$	(2) Winning	(3) Connected	(4) Normal	(5) Middle	(6) Mod. Middle	(7) Boundary
A		X				
AB		X				
AC						
AD						
ABC		X				
ABD	X		A, B, D			
ACD	X		C, D			
ABCD	X	X	D	B, C, D	C, D	D
B		X				
BC		X				
BD						
BCD	X	X	C, D	C, D	C, D	D
C		X				
CD	X	X	C, D	C, D	C, D	CD
D		X				
Total swings			10	7	6	4

Notes: There are four voting blocs ($\bar{N} = 4$), which are located in the policy space A-B-C-D with voting weights A: 18, B: 8, C: 22, and D: 39; the quorum is 61.

Table 2: *EU Council Voting Data: Overview and Some Descriptive Statistics*

	Proposals	Individual Voting Decisions[1)]	Share of 'Yes' Votes[2)]
Total	899	16035	88.33
By year			
1993	9	108	86.11
1994	55	660	86.66
1995	77	1155	90.82
1996	47	705	89.50
1997	51	765	89.80
1998	55	825	88.60
1999	47	705	89.07
2000	54	810	83.58
2001	56	840	86.90
2002	75	1125	83.28
2003	88	1320	85.90
2004	47	865	87.28
2005	56	1400	86.92
2006	70	1750	88.00
2007	16	429	94.17
2008	19	510	91.96
2009	27	723	93.64
2010	30	804	92.41
2011	20	536	91.6
By Policy Area[3)]			
Agriculture	239	3869	88.73
Cohesion	6	124	89.51
Ecofin	23	449	89.08
Energy	3	57	94.73
Environment	59	1122	89.21
Development	4	96	95.83
Fisheries	85	1432	90.57
General Affairs	53	1248	82.21
Health + Health & Consumer Affairs	41	736	88.85
Internal Market	124	2098	89.41
Institutions	10	256	91.01
Justice and Home Affairs	22	522	93.48
Research, Education & Culture	26	463	90.06
Social Policy	16	360	89.44
Statistical System	15	317	93.37
Telecommunication	19	367	92.91
Trade	15	225	75.11
Transparency	79	1179	79.81
Transport	58	1085	90.96

Notes: Data from 1993-2006 merged from Mattila and Lane (2001); Hayes-Renshaw and Wallace (2006), Mattila (2008); data as of 2007 collected from web scraping of Council documents using the interface by Buhl and Rasmussen (`http://api.epdb.eu`). [1)] Number of observations of individual voting decisions is equal to the number of proposals times the number of EU member states that have participated in the voting. [2)] Mean of y_{it} ($\times 100$ in %), corresponding to the share of 'Yes' votes in all voting decisions. [3)] Categorization into policy areas follows Hayes-Renshaw and Wallace (2006). Research, education, and culture were merged into one policy area due to the small number of observations. Two items could not be assigned to a paricular policy area; hence the proposals by policy area sum up to 897 items for the full period 1993-2011. Information on the policy areas Statistical System, Cohesion, Development, and Energy is only provided for completeness; due to the low number of contested votes, they are not considered in the following.

Table 3: *Joint Tests for Equal Ideal Points by Policy Area*

Policy Area	Test Statistic	Policy Area	Test Statistic	Policy Area	Test Statistic
Overall	300.28***	Health & Cons Affairs	34.60	Social Policy	0.01
Agriculture	135.19***	Internal Market	40.01**	Trade	51.08***
Ecofin	0.01	Institutions	0.00	Telecommunication	0.01
Environment	0.18	Institutions	0.00	Transparency	154.47***
General Affairs	77.15***	Justice & Home Affairs	0.00	Transport	0.04

Notes: Test statistic from joint Wald test on equality of countries' ideal points. ***, **, * indicate significance at the 1%, 5% and 10% level.

Table 4: *Empirical Banzhaf Indices, EU27 (2007-2011)*

	Individual Voting[1]				Bloc Voting[2]			
	$BFI_i^{n\,1}$	BFI_i^{m}	$BFI_i^{\check{m}}$	BFI_i^{b}	$BFI_{\bar{i}}^{n}$	$BFI_{\bar{i}}^{m}$	$BFI_{\bar{i}}^{\check{m}}$	$BFI_{\bar{i}}^{b}$
Austria	3.09	4.49*** (0.26)	4.85** (0.71)	0.72 (2.3)	2.97 (0.21)	4.74 (2.81)	5.21 (3.12)	1.04 (2.66)
Belgium	3.68	3.25 (1.62)	3.13 (2.05)	3.56 (5.33)	3.5 (0.29)	3.53 (2.5)	3.41 (2.79)	4.41 (4.96)
Bulgaria	3.09	4.43*** (0.44)	4.71* (0.9)	0.32* (1.59)	2.92 (0.24)	4.1 (2.62)	4.44 (2.89)	1.21 (2.36)
Cyprus	1.25	4.42*** (0.49)	4.74*** (0.95)	0.23 (1.32)	1.2 (0.1)	2.97 (2.83)	3.22 (3.14)	0.44 (1.27)
Czech Republic	3.68	4.53*** (0.16)	4.97*** (0.45)	0.32** (1.55)	3.53 (0.24)	4.98 (2.57)	5.52 (2.85)	0.48* (1.68)
Denmark	2.18	1.14** (0.5)	0.37*** (0.32)	4.15 (4.29)	2.13 (0.2)	1.99 (1.52)	0.50** (0.77)	2.44 (4.24)
Estonia	1.25	4.32*** (0.51)	4.38*** (1.2)	0.83 (2.51)	1.17 (0.11)	2.46 (2.75)	2.55 (2.91)	0.88 (1.59)
Finland	2.18	3.98*** (0.56)	3.51 (1.35)	3.47 (4.25)	2.02 (0.18)	2.89 (2.55)	2.65 (2.47)	3.73 (3.27)
France	7.78	2.73*** (1.6)	2.48*** (1.93)	13.44 (7.78)	8.3 (0.71)	4.3 (2.67)	4.02 (3.09)	12.99 (8.91)
Germany	7.79	4.43*** (0.31)	4.70*** (0.76)	7.77 (6.42)	8.09 (0.76)	6.17 (2.14)	6.65 (2.51)	10.42 (7.17)
Greece	3.68	3.36 (1.5)	3.14 (1.95)	4.16 (4.85)	3.52 (0.27)	3.71 (2.61)	3.56 (2.93)	4.39 (4.97)
Hungary	3.68	4.46 (0.49)	4.86 (0.8)	0.76 (2.5)	3.54 (0.26)	4.88 (2.58)	5.38 (2.93)	0.87 (2.61)
Ireland	2.18	3.98* (1.06)	4.01 (1.66)	1.81 (3.33)	2.07 (0.18)	3.29 (2.79)	3.36 (3.06)	2.03 (3.01)
Italy	7.78	3.43*** (1.45)	3.35** (1.83)	11.33 (9.35)	8.23 (0.75)	5.07 (2.74)	5.22 (3.25)	11.5 (9.39)
Latvia	1.25	4.51*** (0.21)	4.87*** (0.67)	0.00 (0.00)	1.19 (0.11)	2.85 (2.89)	3.13 (3.19)	0.26 (0.72)
Lithuania	2.18	4.51*** (0.16)	4.90*** (0.49)	0.48 (1.92)	2.07 (0.16)	3.7 (2.79)	4.05 (3.08)	0.55 (1.43)
Luxembourg	1.25	3.82*** (0.97)	3.48 (1.69)	1.62 (3.17)	1.19 (0.11)	2.59 (2.89)	2.46 (3.07)	1.36 (2.4)
Malta	0.94	4.53*** (0.23)	4.98*** (0.44)	0.08 (0.83)	0.9 (0.09)	2.71 (3.03)	3.00 (3.37)	0.1 (0.59)
Netherlands	3.97	3.18 (0.72)	2.13 (1.14)	5.29 (4.91)	3.75 (0.34)	3.42 (2.03)	2.31 (1.7)	5.32 (4.13)
Poland	7.43	4.43*** (0.42)	4.73*** (0.86)	4.06 (5.95)	7.59 (0.7)	6.24 (2.27)	6.83 (2.57)	4.7 (6.17)
Portugal	3.68	3.1 (1.46)	2.67 (1.91)	4.69 (5.28)	3.52 (0.27)	3.14 (2.53)	2.83 (2.78)	4.58 (5.00)
Romania	4.26	4.42 (0.57)	4.77 (0.9)	0.68* (2.16)	4.1 (0.31)	4.97 (2.53)	5.4 (2.8)	1.26 (2.86)
Slovakia	2.18	4.50*** (0.38)	4.95*** (0.58)	0.08*** (0.77)	2.09 (0.15)	4.00 (2.79)	4.45 (3.12)	0.21* (1.07)
Slovenia	1.25	4.51*** (0.32)	4.95*** (0.54)	0.07* (0.71)	1.2 (0.1)	3.19 (2.96)	3.55 (3.33)	0.11* (0.61)
Spain	7.43	2.41*** (1.81)	2.29** (2.05)	11.14 (6.39)	7.8 (0.66)	3.2 (2.76)	3.02 (3.01)	10.38 (8.01)
Sweden	3.09	0.29*** (0.43)	0.19*** (0.21)	5.18 (4.87)	3.06 (0.29)	0.29*** (0.61)	0.29*** (0.55)	2.99 (5.15)
United Kingdom	7.79	2.86*** (0.61)	1.90*** (0.8)	13.75 (5.7)	8.36 (0.67)	4.63** (1.41)	3.00*** (1.69)	11.35 (7.22)

Notes: Means and standard errors of EU27 member states' empirical power indices, based on 1000 draws from posterior distribution of member states' ideal points (period 1993-2011). Normal variation ($BFI_{\bar{i}}^{n}$), middle variation ($BFI_{\bar{i}}^{m}$), modified middle variation ($BFI_{i}^{\check{m}}$), boundary variation (($BFI_{\bar{i}}^{b}$)). All indices are normalized (displayed in %). ***, **, * indicate whether the null that the empirical index is equal to the theoretical one is rejected at the 1%, 5%, and 10% level by a t-test.

[1] Power indices of EU27 member states calculated 'directly' under individual voting. The normal variation under invididual voting is equal to the theoretical Banzhaf index.

[2] Power indices of EU27 member states under bloc voting, given by the power of the voting bloc times the share of the respective countries's votes in the bloc.

Table 5: *Empirical Banzhaf Indices by Policy Area (EU27, voting blocs)*

	β_i^n		β_i^m		$\check{\beta}_i^m$		β_i^b	
	RMSE	#	RMSE	#	RMSE	#	RMSE	#
Full Sample	2.12	0	2.38	2	2.43	3	2.96	0
Agriculture	3.64	0	3.41	2	3.08	2	3.32	0
Ecofin	3.19	0	3.69	0	3.25	0	2.37	0
Environment	3.65	0	3.42	0	3.27	0	2.84	0
General Affairs	3.37	0	2.76	0	2.89	1	3.02	0
Health	2.85	0	4.07	1	3.03	2	2.88	0
Institutions	3.17	0	2.61	0	3.09	0	2.67	0
Internal Market	3.87	0	2.28	2	3.01	2	2.51	0
Justice and Home	3.87	0	2.38	0	3.92	0	3.51	1
Social	3.31	0	2.80	0	3.51	0	2.73	1
Telecommunications	3.38	0	2.66	0	4.34	0	2.21	0
Trade	3.08	0	2.83	0	3.04	0	2.34	0
Transparency	4.68	0	2.64	0	3.29	1	2.18	0
Transport	3.96	0	4.04	0	2.87	0	2.69	0

Notes: RMSE is the root mean squared error of the difference between the theoretical and the respective empirical Banzhaf index over the EU27 member states; # denotes the number of countries' of which the empirical index is significantly different from the theoretical one (in terms of a t-test at the 5% level).

Table 6: *Empirical Shapley-Shubik Index and Middle Variation of Other 'Banzhaf-like' Power Indices*

	Theoretical Index				Individual Voting				Bloc Voting			
	SSI	*JNI*	*DPI*	*HPI*	*SSI*	*JNI*	*DPI*	*HPI*	*SSI*	*JNI*	*DPI*	*HPI*
Austria	2.82	3.52	3.53	3.54	0.04***	4.50***	4.55	4.54**	0.46	4.77	4.81	4.81
					(0.19)	(0.27)	(1.26)	(0.44)	(3.09)	(2.83)	(7.24)	(3.00)
Belgium	3.41	4.01	3.69	3.69	4.24	3.24	3.33	3.31	3.61	3.53	3.49	3.63
					(6.07)	(1.63)	(1.89)	(1.62)	(9.87)	(2.51)	(7.32)	(2.77)
Bulgaria	2.82	3.52	3.53	3.54	0.18	4.44**	4.45	4.45	0.16**	4.12	4.16	4.11
					(1.81)	(0.45)	(1.28)	(0.62)	(1.14)	(2.64)	(6.46)	(2.81)
Cyprus	1.10	1.65	3.03	3.07	0.07***	4.43***	4.41	4.46**	0.28	2.98	2.55	3.14
					(0.38)	(0.50)	(1.15)	(0.63)	(1.80)	(2.85)	(4.62)	(3.29)
Czech Republic	3.41	4.01	3.69	3.69	0.00***	4.54***	4.63	4.62***	0.02***	5.02	5.21	5.05
					(0.01)	(0.17)	(1.21)	(0.17)	(0.24)	(2.59)	(7.88)	(2.78)
Denmark	1.96	2.68	3.29	3.32	7.35	1.10***	1.04***	1.01***	5.96	1.85	1.30	1.21
					(5.52)	(0.49)	(0.75)	(0.66)	(10.56)	(1.42)	(3.21)	(1.74)
Estonia	1.10	1.65	3.03	3.07	0.02***	4.32***	4.27	4.25	0.05***	2.47	2.13	2.53
					(0.11)	(0.54)	(1.46)	(0.78)	(0.16)	(2.75)	(4.85)	(3.03)
Finland	1.96	2.68	3.29	3.2	0.04***	3.97**	3.79	3.75	0.25***	2.87	2.30	2.76
					(0.09)	(0.59)	(1.46)	(0.86)	(0.58)	(2.53)	(3.45)	(2.72)
France	8.68	6.39	4.90	4.83	18.00	2.74**	2.76	2.77	17.08	4.30	4.88	4.47
					(16.35)	(1.60)	(1.53)	(1.43)	(41.50)	(2.69)	(7.56)	(2.56)
Germany	8.67	6.39	4.90	4.83	0.09***	4.44***	4.38	4.40	0.59***	6.21	6.69	6.07
					(0.48)	(0.32)	(1.13)	(0.55)	(1.51)	(2.15)	(9.60)	(2.31)
Greece	3.41	4.01	3.69	3.69	3.85	3.35	3.37	3.41	4.01	3.71	3.70	3.72
					(6.05)	(1.52)	(1.53)	(1.48)	(12.05)	(2.61)	(6.78)	(2.86)
Hungary	3.41	4.01	3.69	3.69	0.45	4.47	4.57	4.55*	0.55	4.91	4.88	4.98
					(3.02)	(0.50)	(1.30)	(0.50)	(6.13)	(2.60)	(6.83)	(2.89)
Ireland	1.96	2.68	3.29	3.32	1.18	3.98	4.00	4.02	1.45	3.30	3.06	3.37
					(2.63)	(1.08)	(1.55)	(1.14)	(7.16)	(2.80)	(5.38)	(3.03)
Italy	8.68	6.39	4.90	4.83	11.54	3.44**	3.44	3.44	11.82	5.10	5.81	5.13
					(14.57)	(1.45)	(1.65)	(1.37)	(31.89)	(2.76)	(8.83)	(2.65)
Latvia	1.10	1.65	3.03	3.07	0.00***	4.52***	4.58	4.56***	0.01***	2.87	2.70	2.97
					(0.01)	(0.23)	(1.28)	(0.41)	(0.03)	(2.91)	(6.84)	(3.20)
Lithuania	1.96	2.68	3.29	3.32	0.00***	4.52***	4.57	4.57***	0.02***	3.72	3.38	3.80
					(0.00)	(0.17)	(1.15)	(0.29)	(0.14)	(2.80)	(4.91)	(3.10)
Luxembourg	1.10	1.65	3.03	3.07	0.73	3.81**	3.69	3.72	0.89	2.58	2.36	2.64
					(1.57)	(0.99)	(1.44)	(1.15)	(3.62)	(2.89)	(6.39)	(3.14)
Malta	0.82	1.25	2.59	2.64	0.01***	4.54***	4.63*	4.63***	0.02***	2.73	2.44	2.89
					(0.13)	(0.24)	(1.21)	(0.21)	(0.20)	(3.04)	(5.53)	(3.43)
Netherlands	3.68	4.19	3.77	3.76	0.90**	3.15	2.71	2.71	1.63	3.34	2.74	2.73
					(1.34)	(0.74)	(1.10)	(0.88)	(3.45)	(1.97)	(4.59)	(2.05)
Poland	8.00	6.19	4.78	4.72	0.19***	4.44***	4.45	4.43	0.44**	6.28	6.90	6.16
					(0.94)	(0.43)	(1.36)	(0.61)	(3.35)	(2.29)	(9.80)	(2.31)
Portugal	3.41	4.01	3.69	3.69	5.04	3.09	3.01	3.03	5.67	3.12	3.00	3.21
					(6.44)	(1.47)	(1.61)	(1.43)	(13.51)	(2.52)	(4.97)	(2.78)
Romania	3.99	4.36	3.85	3.83	0.26**	4.43	4.48	4.48	0.22	5.00	5.14	4.99
					(1.65)	(0.58)	(1.32)	(0.62)	(2.76)	(2.54)	(7.97)	(2.71)
Slovakia	1.96	2.68	3.29	3.32	0.09**	4.51***	4.56	4.58***	0.12	4.03	3.83	4.14
					(0.90)	(0.39)	(1.12)	(0.43)	(2.36)	(2.81)	(5.93)	(3.10)
Slovenia	1.10	1.65	3.03	3.07	0.00***	4.52***	4.56	4.60***	0.02***	3.21	2.73	3.39
					(0.01)	(0.32)	(1.07)	(0.36)	(0.24)	(2.98)	(4.21)	(3.41)
Spain	8.00	6.19	4.78	4.72	16.25	2.42**	2.68	2.67	17.56	3.20	4.40	3.61
					(13.47)	(1.81)	(1.83)	(1.57)	(35.33)	(2.78)	(7.62)	2.54)
Sweden	2.82	3.52	3.53	3.54	25.29**	0.28***	0.65***	0.63***	20.41	0.28***	1.09	0.83*
					(9.34)	(0.42)	(0.60)	(0.56)	(39.81)	(0.57)	(3.37)	(1.47)
United Kingdom	8.67	6.39	4.90	4.83	4.18	2.82***	2.45**	2.41***	6.70	4.52	4.33	3.68
					(3.47)	(0.62)	(1.07)	(0.73)	(12.00)	(1.40)	(6.74)	(1.72)

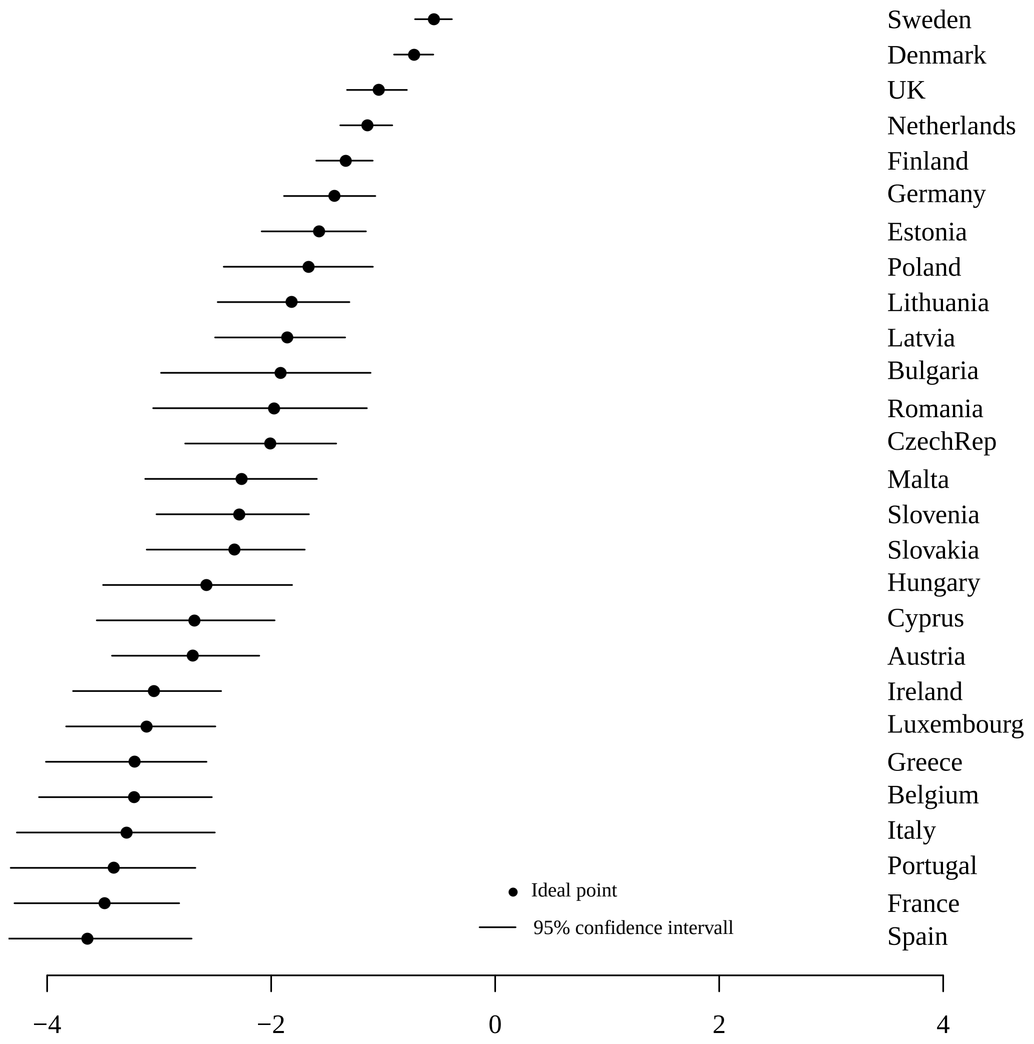
Notes: Means and standard errors of EU27 member states' empirical power indices, based on 1000 draws from posterior distribution of countries ideal points (1993-2011). *, **, *** indicate significance at the 10%, 5%, and 1%level (*t*-test). *SSI*... Shapley-Shubik index, *JNI*... Johnston index, *DPI*... Deegan-Packel index, *HPI*... Holler Packel index. All indices are normalized (displayed in %). See Table 5.

Table 7: *Bloc Voting and the Role of Maximum Bloc Size*

	Maximum Bloc Size				
	1	2	3	4	5
BFI	74.1	25.9	7.4	7.4	7.4
JNI	74.1	11.1	7.4	3.7	3.7
DPI	11.1	0.0	0.0	0.0	0.0
HPI	40.7	11.1	3.7	3.7	0.0
SSI	55.6	44.4	44.4	40.7	37.0

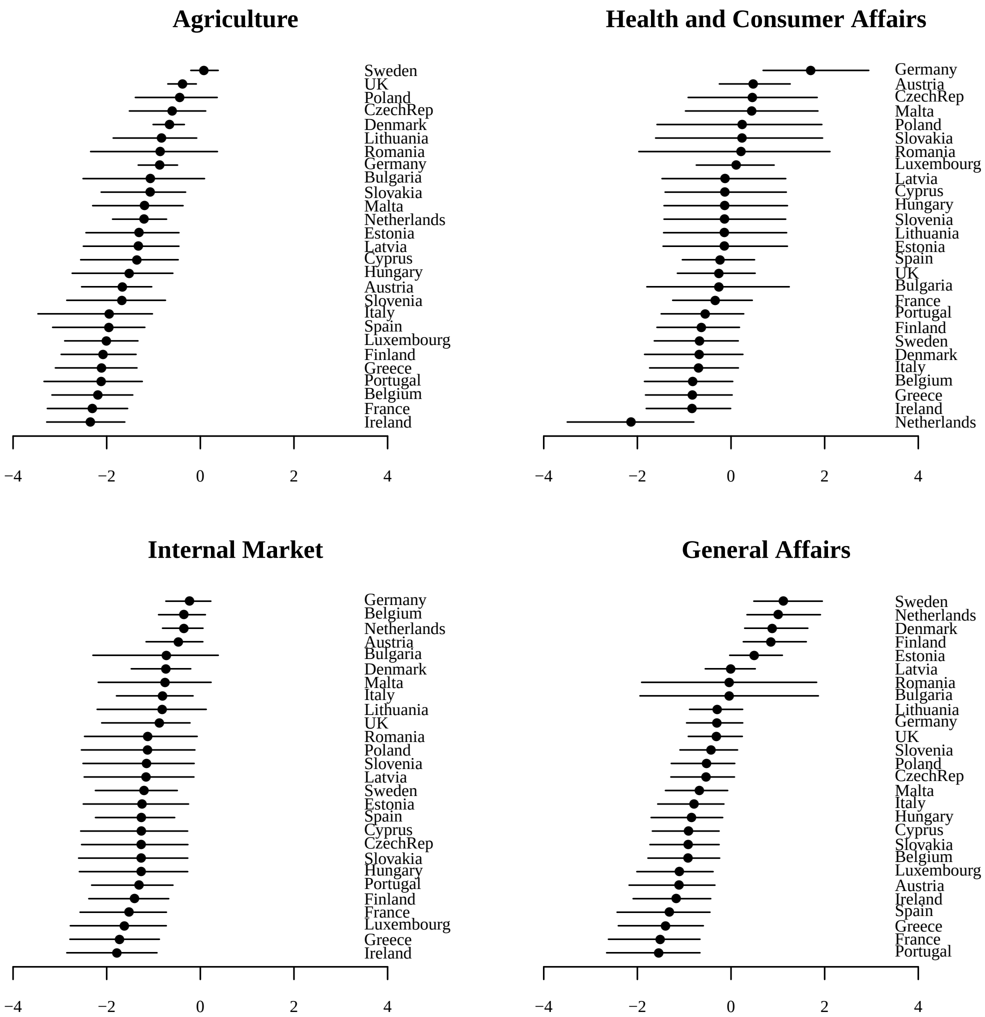
Notes: Share of the EU27 member states (in %), for wich the (middle variation) of the respective empirical index is significantly different from the theoretical one at the 5% level.

Figure 1: *Estimates of EU Member States' Ideal Points,* 1993-2011



Notes: Dots indicate location of mean ideal point estimates, lines show 95% confidence intervals.

Figure 2: *Estimates of EU Member States' Ideal Points by Policy Area,* 1993-2011



Notes: Dots indicate location of mean ideal points, lines show 95% confidence intervals.