A utilitarian assessment of alternative decision rules in the Council of Ministers

Beisbart, Claus; Bovens, Luc; Hartmann, Stephan

European Union Politics
DOI: 10.1177/1465116505057814
Volume 6 (4): 395–418

SAGE Publications
London, Thousand Oaks CA,
New Delhi

A Utilitarian Assessment of Alternative Decision Rules in the Council of Ministers

Claus Beisbart
University of Konstanz, Germany

Luc Bovens
London School of Economics and Political Science, UK

Stephan Hartmann
London School of Economics and Political Science, UK

ABSTRACT

We develop a utilitarian framework to assess different decision rules for the European Council of Ministers. The proposals to be decided on are conceptualized as utility vectors and a probability distribution is assumed over the utilities. We first show what decision rules yield the highest expected utilities for different means of the probability distribution. For proposals with high mean utility, simple benchmark rules (such as majority voting with proportional weights) tend to outperform rules that have been proposed in the political arena. For proposals with low mean utility, it is the other way round. We then compare the expected utilities for smaller and larger countries and look for Pareto-dominance relations. Finally, we provide an extension of the model, discuss its restrictions, and compare our approach with assessments of decision rules that are based on the Penrose measure of voting power.

KEY WORDS

- Council of Ministers
- European Constitution
- probabilistic modelling
- utilitarianism
- voting theory

Introduction

A central problem in negotiating the EU Constitution was how to define qualified majority voting in the Council of Ministers (Council or CM for short). In more general terms, such a problem arises for any federation in which the member countries have to decide on collective actions. A decision rule is needed that determines for which combinations of the countries' votes (i.e. for which voting profiles) a proposal is accepted. The problem is: which decision rule should be chosen? How should qualified majority voting be defined for the CM?

The problem of finding a decision rule not only was the subject of political negotiations in the EU, but also has been attracting much academic interest. Many authors start by requiring a fair distribution of powers. The power or influence of political agents (such as persons or countries) is then expressed in terms of some quantitative measure. The most frequently used measures are the *Penrose measure* (Penrose, 1952) and its normalization in the *Banzhaf index* (Banzhaf, 1965). The Penrose measure equals the probability of making a difference to whether a proposal is or is not accepted (see Felsenthal and Machover, 1998, for a general discussion and Felsenthal and Machover, 2001, for an application to the CM). An alternative measurement of voting power is provided by the *Shapley–Shubik index*, which is based on cooperative game theory (Shapley and Shubik, 1954; for an application to the CM see, e.g., Widgrén, 1994). Both the Penrose measure and the Shapley–Shubik index are a priori measures because they do not take into account any empirical information about the political agents' preferences.

Despite their extensive usage, power measures have come under attack (e.g. Garrett and Tsebelis, 1999, and Albert, 2003; for a response to Albert, see Felsenthal et al., 2003). One of the objections is that power measures do not take into account the actual preferences of the political agents, whereas, on a realistic view, the influence of a country in the CM depends on how its preferences are related to the other countries' preferences. For this reason, alternative approaches that build on a posteriori information on preferences have been developed. One prominent example is the *strategic power index* introduced by Steunenberg et al. (1999), which builds upon non-cooperative game theory. A unifying perspective on power measures and strategic power has recently been developed by Napel and Widgrén (2004).

In this article we propose to pursue a completely different line and assess alternative decision rules in the Council within a strictly utilitarian framework. We use this approach because the implementation of a decision rule is likely to have positive or negative effects on the people in the various EU countries. The reason is that the decision rule governs which proposals will be accepted, where the proposals have positive and negative effects on the people themselves. For example, if a proposal imposes tariffs on the import of olives from Tunisia, then people in the Mediterranean EU countries will typically benefit – they can sell their olives at a higher price – whereas people from non-Mediterranean EU countries will on average be worse off than before – they will be paying a higher price for their olives. If we quantify these effects in terms of utilities and put them together, we get some net effect for Europe as a whole. Now people will be better off if the decision rule will block proposals with a negative net effect and let pass proposals with a positive net effect. Accordingly, the best decision rule will filter the proposals in such a way as to maximize the net utility. Since one cannot foresee the proposals that might be made, we will set up a probabilistic model for proposals and evaluate what voting rule yields the highest *expected* utility. This rule, we suggest, should be chosen.

Though their investigation takes a different direction, Barberà and Jackson (2004) share our interest in assessing decision rules for a federal assembly on the basis of expected utility. There is also some similarity in motivation with strategic power indices (Steunenberg et al., 1999), although the details differ greatly.

The Council of Ministers is part of a larger framework of closely interacting institutions. To make things simpler we will abstract from the other institutions in the following way. First, the work of the *Commission* (which makes the proposals) will be used only to constrain the probability distribution for the proposals. The idea is that the Commission will function as a pre-selector of proposals; thus some kinds of proposals will only rarely be brought forward because there is no Commission to draft them. Secondly, we will neglect the work of the *European Parliament* – whose acceptance is needed for some considerable proportion of the proposals. To be able to do so, we can restrict ourselves to proposals that do not need acceptance by the Parliament. Alternatively, the work of the Parliament might be thought of as an additional pre-selection of proposals that can be absorbed into the probability distribution of our model, too. Further, in assessing decision rules for the CM, it is generally reasonable to concentrate on the CM as a first step. This kind of approach is, after all, taken by many authors (see Kirsch et al., 2004).

Our model is still rudimentary at this stage. But full empirical adequacy is not our aim in this paper. Rather, we seek to apply the utilitarian criterion on the basis of a few reasonable assumptions. We do not claim that our account is superior to the well-established power measures. Instead, our analysis allows for a different perspective on decision-making in the EU. It focuses on benefits and costs and takes seriously the possibility that the benefits for a country that votes for a proposal are different from the costs for a country that votes against the proposal.

The plan of our paper is as follows. In the next section we introduce a number of decision rules. We then lay out our model. After presenting our results, we discuss what political lessons can be learned from our results, expound a set of concerns, and examine the relation between our approach and one power measure approach. Mathematical details are given in the Appendix.[1] We will restrict ourselves to the 25-member EU.

Alternative decision rules

We will compare seven decision rules. Rules SMP through P62 are conceptually simple and act as interesting *benchmarks*.

(SMP) **Simple majority with proportional weights.** A proposal is accepted if and only if the population of the member countries that support the proposal is at least 50% of the total population of the European Union.

(SME) **Simple majority with equal weights.** A proposal is accepted if and only if at least 50% of the member countries support the proposal.

(D) **The double-majority model.** This model combines the conditions of SMP and SME. A proposal is accepted if and only if the population of the member countries that support the proposal is at least 50% of the total population of the European Union and at least 50% of the member countries support the proposal. The double-majority model was proposed by the European Commission at the Intergovernmental Conference in 2000.

(P62) **The Penrose-62 Rule.** A proposal is accepted if the voting weights of the countries in favour of the proposal add up to at least 62% of the total voting weights, where the voting weight of each country is proportional to the square root of its population. The motivation for this is as follows. Let us suppose that each country elects a representative for the Council from two candidates on the basis of a simple majority vote. Penrose (1946) showed that the voting power of each citizen in the election of the representative is inversely proportional to the square root of the population in his/her country. Therefore, if we set the voting power of each country in the Council proportional to the square root of its population, each citizen in the federation will have the same voting power at the level of the CM (square root rule, Penrose, 1946), where voting power is always measured in terms of the Penrose measure. As Slomczynski and Zyczkowski (2004) show, square root power within the Council can be achieved by giving the countries square root weights and by requiring a quota of acceptance of 62%. This proposal is defended in Kirsch et al. (2004), and an open letter in defence of square root weights was sent to the EU governments and institutions (Bilbao et al., 2004; see also Laruelle and Widgrén, 1998).

The following rules, (Acc) through (Con), are *political rules*. They have all been serious contenders for decision rules in the CM in the political arena:

(Acc) **The Accession Treaty.** The Accession Treaty (signed in April 2003) contains the decision rule that is currently in force in the 25-member EU (this rule is a slight modification of the decision rule in the Nice Treaty). Each country receives multiple votes. Larger countries have more votes than smaller countries, but there is no analytical formula that determines how many votes each country should get. A proposal is accepted if and only if: (acc.a) at least 72% of the votes support the proposal; and (acc.b) the population of the member countries that support the proposal is at least 62% of the total population of the European Union and (acc.c) at least 50% of the member countries support the proposal.

(Draft) **The original Draft Constitution** (as proposed by the European Convention in July 2003). The Draft Constitution in effect drops condition (acc.a), modifies the quotas of conditions (acc.b) and copies condition (acc.c). A proposal is accepted if and only if (draft.a) the population of the member countries that support the proposal is at least 60% of the total population of the European Union and (draft.b) at least 50% of the member countries support the proposal.

(Con) **The current Draft Constitution for proposals submitted by the Commission.** The current Constitution adopts variations of (draft.a) and (draft.b), where in each case the quota has been raised. But it leaves open an alternative route for the acceptance of a proposal when there is massive support from smaller countries, even though the population quota has not been met. Thus, a proposal is accepted if and only if (con.a) the population of the member countries that support the proposal is at least 65% of the total population of the European Union and (con.b), in the current 25-member EU, at least 15 countries support the proposal; alternatively, instead of (con.a) and (con.b), (con.c) at least 22 member countries support the proposal.[2]

If (con.c) holds, (con.b) is necessarily fulfilled, too. In effect, (con.c) prevents groups of three big countries from blocking decisions. (Con) was approved by the European governments at the Brussels summit in June 2004. But it will become effective only if it is ratified by all member countries. If it is not, then the Accession Treaty will apply.

Model

We will now set up a general utilitarian framework to assess decision rules (for our model see Schweizer, 1990). We will assume a hypothetical federation consisting of N countries and return to the case of the EU in the next section. Let us first consider a particular proposal that is brought to vote. The desirability of its acceptance or rejection for each person will be expressed in terms of personal utilities. Because we need to average over utilities for a group of people, we have to assume full interpersonal comparability of scales. Either one can think of these personal utilities as primitive notions, or one can interpret them as von Neumann-Morgenstern utilities, in which case, however, a number of additional restrictions have to be met. Without loss of generality we can stipulate that every person gets zero utility if the proposal is rejected.

Let us now consider some country i, and let v_i be the average utility that a person from country i will receive if the proposal is accepted in the Council of Ministers.[3] Each proposal can then be represented by an N-tuple of real-valued numbers v_i quantifying the utilities. Let r_i be the population fraction of country i within the EU population. On average, people in the EU will feel the net effect

$$v = \sum_{i=1}^{N} r_i v_i, \tag{1}$$

if the proposal is accepted.

The Council decides about the proposal following a decision rule. The utility u_i that an average person in country i gets from the decision can be written as

$$u_i = v_i D, \tag{2}$$

where D represents the outcome of the decision. D equals 1 if the proposal is accepted, and 0 otherwise.

The outcome of the decision depends on the countries' votes. Let $\lambda_i = 1$ if country i votes in favour of the proposal, and $\lambda_i = 0$ otherwise. D is then a function from the set of possible voting profiles $\{\langle\lambda_1, \ldots, \lambda_N\rangle \mid \lambda_i = 0, 1$ for $i = 1, \ldots, N\}$ into the set $\{0,1\}$. Specifying this function is a way of spelling out the decision rule adopted. To mark the dependence on the decision rule, the u_is might be given an additional label indicating the rule. We drop this label, unless it is not clear from the context which rule applies.

We assume that the countries vote on the basis of their own utilities only: $\lambda_i = \lambda_i(v_i)$. Moreover, we assume that country i votes in favour of a proposal just in case the utility v_i exceeds zero. This does not involve any substantial assumption – it means only that the country will vote in favour of the proposal if and only if acceptance would be better for its people on average.

We can now construct a probabilistic model for the proposals and introduce N random variables V_i. They take real values v_i that represent the utilities as defined above. The motivation for adopting a probabilistic model is this: when we are setting up decision rules (i.e. when we are deciding how to decide), we are operating at a meta-level. From this perspective we cannot foresee all the proposals that will be put forward in the future. But we are not completely ignorant about the proposals that will come to the vote, because, as we said earlier, proposals are pre-selected by the Commission. It is thus justified to account for the variety of proposals in terms of a probability distribution. We adopt the following assumptions for our model:

(a) Vi is independent of $V_i, \ldots, V_{i-1}, V_{i+1}, \ldots, V_N$ for all countries i, i.e. the probability density for a proposal factorizes into the product of the probability densities p_i for the individual countries i.
(b) The V_is are identically distributed for all countries i.
(c) The V_is are normally distributed, i.e. for all i: $p_i = N[\mu, \sigma]$, where μ and σ denote the mean and the standard deviation of the normal distribution.

(These assumptions will be further discussed in Sections 4 and 5.)

Utilitarianism commends the decision rule that maximizes expected utility. Insofar as the population sizes of the EU countries are not affected by the proposals,[4] we can equivalently maximize the expected utility per average person in the EU. To do this we introduce another set of random variables U_i that represent the utilities an average person in country i gets out of the decision process. The utility U of an average person in the EU is the weighted sum

$$U = \sum_{i=1}^{N} r_i \, U_i, \tag{3}$$

where the r_is again represent the countries' population fractions. The challenge here is to find the decision rule for which the expected utility, $E[U]$, is maximal.

To calculate $E[U]$, we first consider the $E[U_i]$s. Recalling eqn (2), we express U_i ($i = 1, \ldots, N$) in terms of the V_js ($j = 1, \ldots, N$),

$$U_i = V_i\, D\, (\langle \lambda_1(V_1), \ldots, \lambda_N\, (V_N)\rangle), \tag{4}$$

and take the mean with respect to the probability distributions over the V_js. We get

$$E[U_i] = \int \mathrm{d}v_1 \ldots \int \mathrm{d}v_N\, p(v_1) \ldots p(v_N)\, v_i\, D\, (\langle \lambda_1(v_1), \ldots, \lambda_N\, (v_N)\rangle). \tag{5}$$

Combining eqns (3) and (5) yields

$$E[U] = \sum_{i=1}^{N} r_i \int \mathrm{d}v_1 \ldots \int \mathrm{d}v_N\, p(v_1) \ldots p(v_N)\, v_i\, D\, (\langle \lambda_1(v_1), \ldots, \lambda_N\, (v_N)\rangle). \tag{6}$$

In the Appendix we show how to calculate the expressions (5) and (6). The results in the next section are based on these calculations.

Results

Expected utility

In our calculations for the EU we first assume that the standard deviation is fixed at $\sigma = 1$. We plot $E[U]$ as a function of μ in Figure 1. For large negative values of μ, every decision rule will block almost every proposal such that $E[U]$ stays at 0. For large positive values of μ, every decision rule will let pass almost every proposal and hence $E[U]$ goes to the value of μ. For values of μ around and slightly greater than 0, the expected utility increases as μ increases and the different decision rules yield noticeably different expected utilities. We specifically examine this region in Figure 2.

If we look at the left panel of Figure 2, there are many crossing points around $\mu = -0.35$ where the ranking over decision rules changes. For negative values of μ, SMP and SME yield manifestly negative expected utilities $E[U]$. But, having reached their minima, the curves for SMP and SME rise steeply and cross many other curves. In the right panel, where positive values of μ are considered, the benchmark rules SMP and D outperform all the political rules. More generally, the political rules do quite poorly in this range of μ values.

Let us look at the current decision rule, Acc. In our utilitarian framework it performs poorly in the range of positive μ values shown in Figure 1. Moreover, we see from Figure 1 that in general the other political rules provide an improvement over Acc for $\mu > 0$. Now consider the benchmark rule of SMP. This rule performs very well for positive values of μ. However, a small country might justifiably be nervous that larger countries will consistently outvote it. One remedy would be to combine the SMP with the SME, yielding the rule D: for a proposal to be accepted, both the majority of the

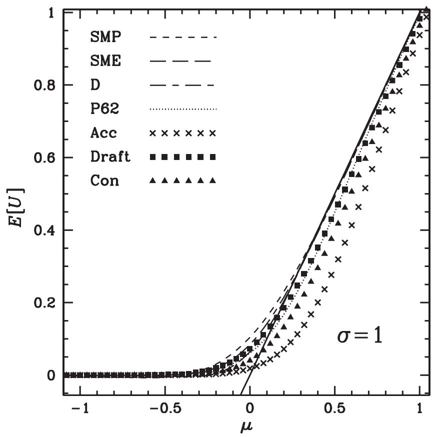

Figure 1 The expected utility, $E[U]$, as a function of the model parameter μ. The standard deviation σ is set to one. The thin solid line denotes the line $E[U] = \mu$.

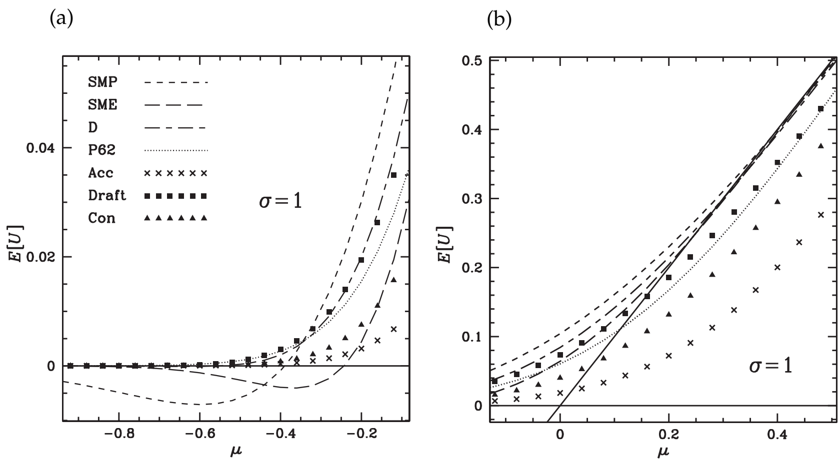

Figure 2 The expected utility, $E[U]$, as a function of μ for $\mu < 0$ (panel a) and for positive μ around 0 (panel b). The thin solid lines mark $E[U] = 0$ (both panels) and $E[U] = \mu/\sigma$ (right panel).

population *and* the majority of the countries must support it. Another solution is to give the smaller countries larger weights and to raise the quota of acceptance. But notice that the former remedy is much less costly in terms of expected utility than the latter. The D rule yields substantially higher $E[U]$ than the various political rules and P62 in the range of positive values of μ

shown in the right panel of Figure 2. At the same time it has no significant dip for negative values of μ.

To explain our results, notice that $E[U]$ is the product of the probability that a proposal is accepted, $P(A)$, and the conditional expected utility, $E[U \mid A]$:[5]

$$E[U] = E[U \mid A]P(A). \tag{7}$$

In the left panel of Figure 3, we plot $P(A)$ as a function of μ. Notice that most of the benchmark rules are more *permissive* than the political rules, that is, their curves start rising for smaller values of μ. The reason is that the political rules tend to be more complex, with a conjunction of criteria for acceptance and higher quotas. In the right panel, we plot $E[U \mid A]$ as a function of μ. Notice here that less permissive rules tend to be more *selective*, that is, they yield higher $E[U \mid A]$. When we take the product of both factors according to eqn (7), there are two forces at work. On the one hand, for lower values of μ, the SMP and SME are too permissive – they sometimes let proposals pass that have negative utilities for many countries, which results in $E[U] < 0$. On the other hand, for higher values of μ, these same rules are beneficial, because they make it possible that more proposals are being accepted and thus $E[U]$ becomes comparatively large for those rules. As a result of this interplay between permissiveness and selectivity, no rule globally outperforms any other – the choice of an optimal rule depends on the value of μ. If the proposals tend to yield negative utility for an average person in the EU, selectivity is more of a concern and less permissive rules are more desirable. If, on the contrary, the proposals tend to yield positive utility, then selectivity is less of a concern and more permissive rules are more desirable.

So far we have restricted ourselves to the case of $\sigma = 1$. Remarkably, all of our results carry over to arbitrary values of σ. This can be shown by means of a scaling argument. Let us start with arbitrary values of μ and σ and the corresponding expected utilities for the different rules. If we multiply both μ and σ with a positive constant α, the corresponding expected utilitities $E[U]$ are multiplied by α as well. As a consequence the order of the decision rules is unaffected. Moreover, if we plot $E[U]/\sigma$ as a function of the rescaled mean, μ/σ, the resulting curves convey full information for all normal models. In the following two subsections we will therefore consider the rescaled quantities $E[U]/\sigma$ and μ/σ only.

Expected utilities for single countries

From the perspective of a single country, what matters is not so much the expected utility for the federation, but rather the expected utility for its own

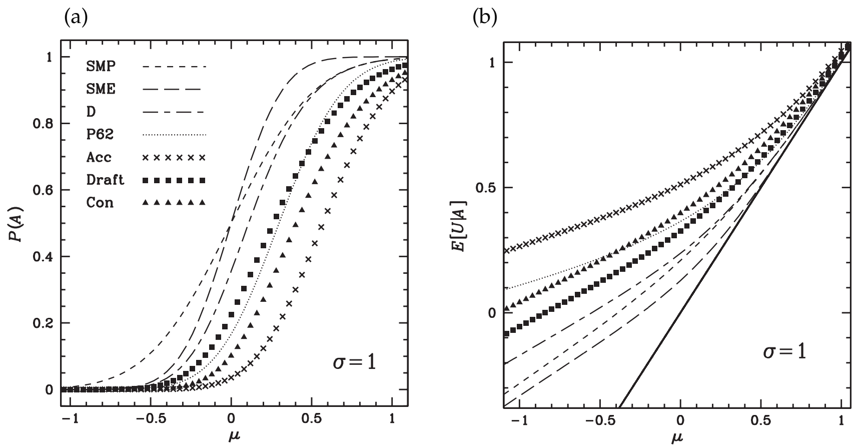

Figure 3 The probability that a motion is accepted, $P(A)$, (panel a) and the conditional expected utility, $E[U|A]$, (panel b) as functions of μ.

people, $E[U_i]/\sigma$. As an example consider Spain and Poland's position on the Constitution around the Brussels summit in December 2003. Both countries refused to vote for the Constitution because they thought any decision rule on the table would disadvantage them compared with Acc.

But is it really true that some countries, Spain and Poland for instance, would be worse off, if the EU were to switch from Acc to Con or Draft? In order to check this, we turn to single countries and consider our model for the most extreme countries in terms of population sizes – Malta and Germany – and for a middle-sized country – Poland. In Figure 4 we show the expected utilities $E[U_i]/\sigma$ as a function μ/σ. There are at least three lessons to be learned here.

First, larger countries do better than smaller countries on any rule (except on the SME, on which all countries do equally well), since they are more probably on the winning side due to their greater weights.

Secondly, as expected, the SMP, which assigns greater weights to larger countries, does best for Germany, whereas Malta fares much better with the SME for a broad range of μ/σ values. SMP also produces a significant dip in Malta's expected utility for $\mu/\sigma < 0$. But, as already noted in the previous subsection, rule D provides a balance between SMP and SME and thus strikes a compromise between the interests of smaller and larger countries. It also avoids the big dip for Malta. Moreover, it outperforms all the political rules for all countries for quite a large range of μ/σ.

Thirdly, for $\mu/\sigma > 0$ the political rules perform poorly for *all* countries, because they impose high quotas. Historically, the EU tried to move from a

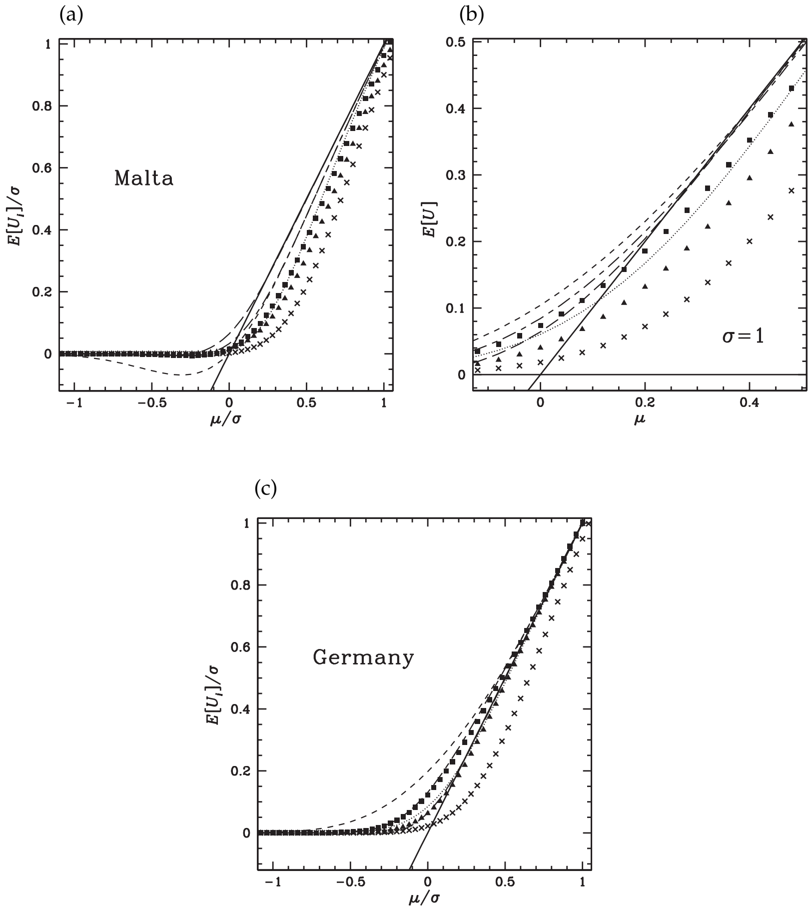

Figure 4 The expected utility $E[U_i]/\sigma$ for Malta (panel a), Poland (panel b) and Germany (panel c) as a function of μ/σ.

relatively restrictive rule, Acc, to a relatively permissive one, Draft, and then struck a compromise with Con. But the general trend is that more permissive rules also benefit small countries for positive μ/σ values, not to mention the expected utility of larger countries and of the whole EU. So, for some large range of positive μ/σ values, the Spanish and Polish concerns are not justified.

Pareto-dominance

Our results in the last subsection raise the question of whether there is a certain rule R that Pareto-dominates another rule S. R Pareto-dominates S should R yield expected utilities $E[U_i^R]$ that are at least as high as those from S for all countries $i(E[U_i^R] \geq E[U_i^S])$ and yield a higher expected utility for

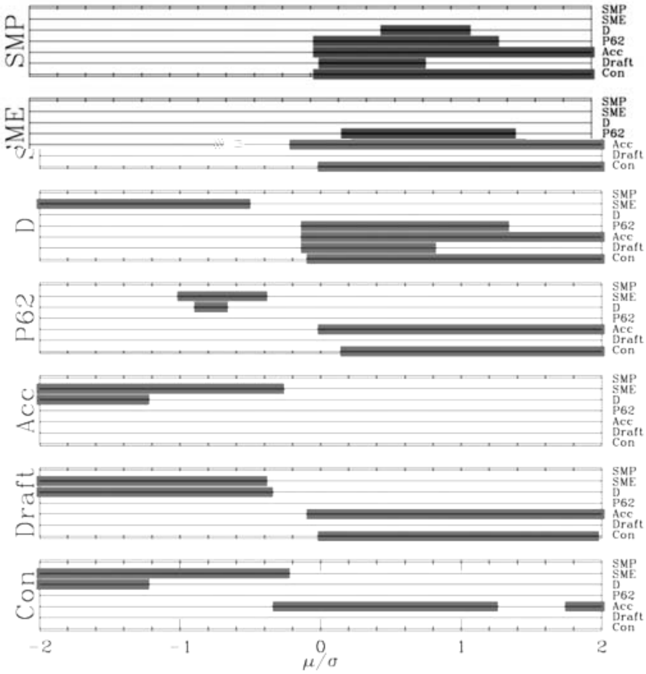

Figure 5 Pareto-dominance relations between the decision rules. In each panel we compare a fixed rule *R* (e.g. SMP in the top panel) with all rules *S* listed on the right-hand side of the panel. The grey intervals denote the range of μ/σ values in which *R* Pareto-dominates *S*.

some countries $j\,(E[U_j^R] > E[U_j^S])$. In Figure 5, we show the intervals for which one rule Pareto-dominates another. We see immediately that there is no rule that Pareto-dominates any other rule for all values of μ/σ. However, certain Pareto-dominance relations do hold for roughly the regions $\mu/\sigma > 0$ and $\mu/\sigma < 0$.

For positive values of μ/σ, political rules tend to be dominated by benchmark rules. The Con and Acc are 'heavily' Pareto-dominated, i.e. by a broad range of benchmark rules and in a broad region of μ/σ, whereas Draft is least heavily Pareto-dominated. Within the political rules, we witness a matching pattern in the Pareto-dominance relations: Draft Pareto-dominates both Con and Acc for a broad range of μ/σ values.

Amongst the negative values of μ/σ, the political rules Pareto-dominate the benchmark rule SME and, for a smaller range of μ/σ values, also D, which in turn Pareto-dominates SME.

Altogether, these results indicate once more that the political rules tend to be better for some range of negative μ/σ values, whereas for positive μ/σ the benchmark rules outperform the political rules.

Relaxing the assumption of normality

So far our results were based on the assumptions (a), (b) and (c) as stated in the section on the model. We now relax the normality assumption in (c) and consider arbitrary probability densities p for the utilities V_i. As we show in the Appendix, the expected utility $E[U]$ in our model depends on only two global characteristics of the utility distribution p. One characteristic is the probability p_+ that the utility for a given single country, V_i, is higher than zero. According to our model, p_+ equals the probability that a proposal is voted for by a given single country. The other characteristic compares the gains and losses to be expected for that country, *given that it votes in favour or against the proposal*. More specifically, let v_+ denote the mean of V_i given that V_i is positive:

$$v_+ = \int_0^\infty p(v)\, v\, dv \Big/ \int_0^\infty p(v)\, dv. \qquad (8)$$

Similarly, let v_- denote the mean of V_i, given that V_i is negative. Only the positive ratio $v_+/|v_-|$ affects the expected utilities $E[U_i]$, as we demonstrate in the Appendix.

Altogether, p_+ and $v_+/|v_-|$ are the only free parameters in our model, and they can be thought of as spanning a phase space. In Figure 6 we show a phase diagram, where each point in the phase space – that is, each full specification of the model – is marked with a different shade of grey indicating the decision rule that performs best for these values of the parameters. The performance is once again measured in terms of expected utility.

We observe connected regions in the phase diagram where some decision rules outperform all others. Moreover, all but one rule – D – appear in the diagram and do best for some values of the model parameters. In the region where $v_+/|v_-| > 1$, only SMP and SME can be seen. In this region the expected gains v_+ outweigh the expected losses $|v_-|$, and therefore highly permissive rules do better. The region where $v_+/|v_-| < 1$, on the other hand, is divided into smaller regions where mostly political rules maximize expected utility. In this case, more selective rules guarantee that more countries profit from a decision, and thus yield greater expected utility. Note that the phase transition curves, where one rule stops and another rule starts maximizing expected utility, are typically horizontal in this range such that we have a layered structure. The different layers are roughly ordered according to their acceptance quotas: decision rules with higher quotas are maximizers for smaller values of $v_+/|v_-|$. This illustrates once more how important the quota is.

The effects of different weightings can be observed in the upper half of the phase space, where $v_+/|v_-| > 1$. As we decrease p_+, we move from SME into SMP, i.e. the optimal weighting changes from equal weights to proportional weights.

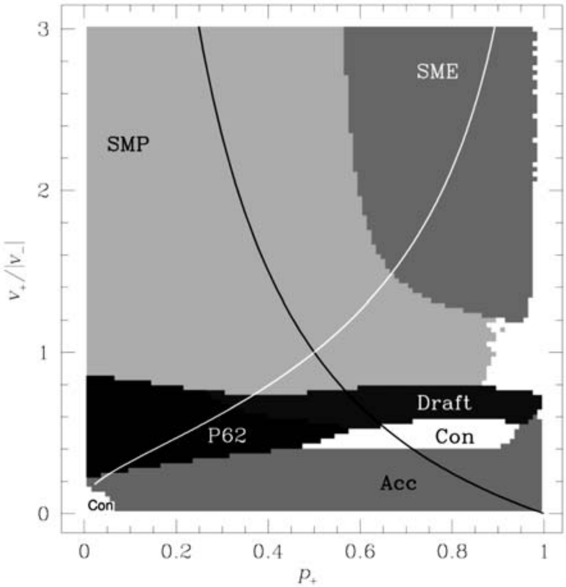

Figure 6 The phase diagram for the general models with independently and identically distributed utilities V_i.

Note that the phase diagram provides the optimal decision rules for the *open* interval $p_+ \in (0, 1)$ only. The differences in the expected utility of all decision rules converge to 0 as p_+ goes to 0 or 1. When $p_+ = 0$, every country votes against the proposal and hence the expected utility of all rules equalizes at 0. When $p_+ = 1$, every country votes for the proposal and hence the expected utility of all rules equalizes at v_+. Because of numerical round-off errors in our computation we do not get meaningful rankings for some values around $p_+ \approx 0.9$.

In the phase diagram, two lines are shown. They help us to choose realistic values for the model parameters. For values of the parameters on the black line, the mean utility $\bar{v}$ of a proposal is 0. This mean can be written as:

$$\bar{v} = \int_{-\infty}^{\infty} p(v)\, v\, dv = v_+ p_+ + v_- p_-. \tag{9}$$

The black line thus separates between regions that have positive – upper right half – and negative – lower left half – values of $\bar{v}$. Note that for the normal model $\bar{v} = \mu$.

The white line has the following significance. Whereas the generalized model has two free parameters, p_+ and $v_+/|v_-|$, the normal model is defined by a single parameter, μ/σ. Hence, the normal model creates a dependency between p_+ and $v_+/|v_-|$. The white line represents this dependency. In regions close to this line, the probability distribution is nearly single-peaked and symmetric. Asymmetric or multiple-peaked probability distributions, in contrast, yield p_+ values and $v_+/|v_-|$ values in the upper left or the lower right corner.

Discussion

What are the political consequences of our results? Can we recommend a decision rule on this basis?

Unfortunately, there is no decision rule that performs best independently of all model parameters. Within the normal models, there are roughly two regimes. For μ/σ values significantly smaller than 0, the political rules tend to outperform other rules in terms of expected utility because they are less permissive and thus in effect block proposals with too many negative utilities V_i. On the other hand, for $\mu/\sigma > 0$, benchmark rules tend to outperform the political rules in terms of expected utility and Pareto-dominance. The choice of a decision rule thus crucially depends on the value of μ/σ. If a typical EU proposal tends to make things worse, i.e. $\mu/\sigma < 0$, then we are in need of *defensive politics*, and for this purpose the more selective political rules are preferable. If they tend to make things better, i.e. $\mu/\sigma > 0$, then we can afford *offensive politics*, and for this purpose the less selective benchmark rules are preferable.

We take this μ/σ dependence to be an advantage of our model. μ/σ is an empirical parameter that describes how people in the EU countries would be affected by the proposals drafted by the Commission. It can thus, in principle, be determined on empirical grounds. In this sense, our approach is a posteriori. Thus our model reveals what kind of empirical input is needed for choosing the best decision rule. Yet some might want to object to this feature (Felsenthal and Machover, 2001: 13): if we are to choose among different decision rules under the Rawlsian veil of ignorance, how can we make use of empirical information? We disagree with this objection. The rationale of the Rawlsian veil is to prevent people from knowing the positions they will take up in society. The empirical information that we require is of a different kind. It is already neutral and does not presuppose a specific position in the world. Rather it spells out certain characteristics of the improvements that collective actions can bring. There is no reason to hide this information from people choosing decision rules.

Thus everything hinges on the choice of the model parameters. But what are realistic choices for μ/σ and, more generally, p_+ and $v_+/|v_-|$ for the current EU? Consider the following two arguments.

First, it does not seem unreasonable to take symmetric and single-peaked probability density functions over the utilities of a proposal as our default option – at least as long as there is no special reason that we should deviate from this assumption for a particular kind of proposal.

Secondly, no Commission would bother submitting a proposal unless it promotes the welfare of a majority of the countries, so it is fair to assume that $p_+ > 0.5$. On the assumption of normality, this implies $\mu/\sigma > 0$ and $v_+/|v_-| > 1$.[6] But you cannot always serve all countries' interests. Some countries are bound to lose from collective action in some domain, say agriculture, whereas they are likely to gain in another domain, say traffic. If we assume that the probability that a proposal would harm a particular country exceeds 16%, i.e. $p_- > 0.16$, then it can be shown that $\mu/\sigma < 1$ and that $v_+/|v_-| < 2.46$. In sum, we think that the proposals submitted by the Commission tend to improve upon the status quo but cannot provide radical improvements.

A realistic choice of the parameters would therefore lie on the white line and in the region $0.50 < p_+ < 0.84 = 1 - 0.16$ and $1 < v_+/|v_-| < 2.46$ (see Figure 6). In this region, SMP and SME yield the highest expected utility. If we can ascertain a particular value of p_+ within this region on empirical grounds, then we can determine a best decision rule. Fortunately, we can allow for some margin of error in our empirical estimates, because there are broad ranges of parameter values for which the best choice stays the same.

It is noteworthy that such simple rules perform best for reasonable choices of the parameters. Furthermore, it is remarkable that the Constitution (Con) is optimal for only a small region in the phase space. On the basis of these considerations, a common mistake in choosing decision rules can be identified: countries insist on higher quotas to prevent being too easily outvoted. Yet, our model shows that imposing higher quotas makes proposals fail that would have increased the utility of the average European. Moreover, even an average person in smaller countries would, in many circumstances, fare better with lower quotas. In the range $0 < \mu/\sigma < 2$, the benchmark rules Pareto-dominate many of the political rules.

There is something puzzling about this result. Why did the EU countries agree to a Constitution that does poorly in terms of expected utility and that is Pareto-dominated by many other rules for a substantial range of μ/σ? Why does the actual Constitution do worse than the original Draft Constitution? Questions such as this might motivate a different attitude to our model: instead of criticizing EU politics on the basis of our model, some might be inclined to criticize our model on the basis of the political agreements. Considering how surprising our results are, we should examine our model and our assumptions carefully.

The most conservative way to account for these problems from the standpoint of the model would be to question the setting of $0 < \mu/\sigma < 1$. For it is true that, for lower values of μ/σ, the political rules tend to do better than the benchmark rules. But we do not think that this is the correct explanation. We have shown that there are good reasons for setting $0 < \mu/\sigma < 1$. And, although political rules do yield a higher expected utility in the range of clearly negative μ/σ values, the differences between the rules under consideration are minimal. In addition, more plausible concerns about our model can be raised.

The identity assumption

The identity assumption (b) is suspect. We assumed that the utility distributions for the proposals were identical; for the normal models we characterized them by using the same μ/σ. But some countries may point out that their utility distributions are systematically *biased* relative to the utility distributions of other countries; that is, on the submitted proposals they systematically derive more or fewer benefits than others. Maybe these biases are correlated with the country size, or other characteristics such as infrastructure and geographical location come into play. These are once again empirical questions.

We can accommodate this point easily by slightly modifying our model. Allowing for different instead of identical utility distributions calls for more empirical input, but it does not fundamentally change the model. We have tested variations of our model where up to five big (small) states have significantly different utility distributions from the utility distributions of other nations. We assumed that the means of their Gaussians are systematically increased by 1σ or 2σ. We found that overall our results are not affected. There are, of course, some changes in the curves $E[U]$ vs. μ/σ, but the ranking between the different decision rules stays roughly the same within both regimes of μ/σ. This might, however, change if both normality *and* identity are dropped at the same time. In the future, we hope to study further such modifications of our model.

The independence assumption

The independence assumption (a) is suspect. Clearly it is not the case that the V_is are independent. For instance, Portugal and Spain's interests will typically be in line with each other and diametrically opposed to, say, the interests of Luxemburg. So proposals that will benefit Spain might typically harm Luxemburg. But, once again, the correlations between the utilities V_i for the various countries in the European Union are an empirical question. The correlations can be incorporated into our model by specifying a μ vector with a μ_i parameter for each country and a covariance matrix that characterizes a multivariate normal. Our analysis then can be carried out with an increased sensitivity to empirical data.

Misvotes

Countries may misjudge the utility that they receive from a proposal, which leads them to vote against a proposal that would benefit them or in favour of a proposal that would harm them. Misvotes are important for assessing decision rules, for it is a reasonable requirement that a decision rule should work not only in ideal circumstances, but also if our knowledge is less than perfect. This concern can readily be accommodated in the model. Assuming that countries are just as likely to overestimate as to underestimate their utility and that they do so to the same extent, we could insert an error term with a normal distribution. We plan to investigate such a refined model in our future work.

Proposals as a function of decision rules

So far we have assumed that the choice of a decision rule does not affect the probability distribution $p(v)$. But there is likely to be a feedback effect. Depending on the decision rule adopted, different proposals will be drafted. One reason is that the Commission only makes proposals that have a fair chance of being accepted, and whether a proposal will be accepted in turn depends on the decision rule adopted. So, the utility distribution that characterizes a proposal is a function of the decision rule and should be modelled as an endogenous variable: if we adopt a decision rule with a high quota of acceptance, then we would expect a higher μ/σ value for proposals than if we adopt a decision rule with a low quota of acceptance.

In response, we want to argue that we are not interested in the actual proposals that come to vote, but rather in opportunities to benefit from collective action at a more fundamental level. Whether there are such opportunities and the nature of these opportunities are not influenced by the choice of the decision rule. The decision rule will affect only the subset of opportunities that materialize in the form of *actual* proposals, but this is not what we are interested in.

Coalition formation

If a *series* of decisions are made with the decision rule, as there certainly are, then strategic moves and the formation of coalitions become possible. For instance, it may pay off for Germany to form a coalition with Austria. Both countries may agree upon the following policy: usually the countries vote in accordance with their utilities but, when a proposal is on the table that harms Austria very much, Germany will vote against it regardless of its own utility, and vice versa. Coalitions of this kind are common in politics, and game-theoretical models need to be invoked to study this kind of interaction. We plan to extend our model to a repeated game in which such strategies can be investigated together with their impact on expected utility.

A similar concern is that the decision on a single proposal is usually preceded by informal discussions and consultations during which the original proposal might be changed by the Commission. To account for this, we would have to extend our model into a sequential game. A simple way to do so is as follows: first, the Commission drafts a proposal according to our model, second the countries get a vote on it, but then there is a modification – if the proposal is rejected according to the decision rule, but the decision is pretty close, then the Commission redrafts the proposal in accordance with the voting profile. As an example, the Commission might just add to the utilities of countries that voted no and subtract from the utilities of the countries that voted yes. By modelling the work of the Commission in such a way, we can keep it as an exogenous factor. Then a second, final decision on the modified proposal is made according to the decision rule. Now extensions of this type will presumably yield qualitative differences to our model in the following way. If one regards our model as a game, it is a rather trivial game, because, for every country, voting according to its own utility is a weakly dominant strategy and these strategies yield a Nash equilibrium. This property will presumably be lost if the model is extended, and genuine strategic interaction might become important. In order to account for such effects we would have to find equilibrium solutions for the game in the extended model and to see how they affect the expected utility (Steunenberg et al., 1999). We admit that such strategic interactions are important for describing the CM's work. We also concede that single countries might evaluate different decision rules according to the strategic options they open up for them. But the crucial point for our work is whether taking into account strategic interactions will affect the expected utility for an average European person as we have calculated it. We suspect that such changes will not be very significant, but we want to check this in our future work.

The expected utility given acceptance

We noted that the curves for the expected utility of a proposal *given acceptance*, $E[U|A]$, are quite different from the curves for the expected utility of a proposal, $E[U]$. Furthermore, benchmark rules do rather poorly with respect to $E[U|A]$. So how can we be sure that $E[U]$ instead of $E[U|A]$ is the relevant measure of evaluation? In our opinion, the truth lies somewhere in the middle. Consider the following analogy. Suppose that you get 100 closed envelopes. Some envelopes contain debts, whereas others contain credits. You may either accept or reject an envelope and you need to develop some kind of algorithm about when you will accept and when you will reject. Now, clearly, what is of interest here is an algorithm that maximizes $E[U]$ and not an algorithm that maximizes $E[U|A]$. But now imagine that there is an infinite supply of envelopes and that, if you reject one, there will be a new envelope drawn at random from the supply. New envelopes will be drawn until you have accepted 100 envelopes. In this case, what you want to maximize is $E[U|A]$.

Let us now return to the proposals. Certainly when a proposal fails, a new proposal concerning the same item may be brought to the table. This seems to indicate that it is $E[U|A]$ rather than $E[U]$ that matters. But drafting a proposal and discussing it take some time. As a consequence, the supply is not infinite. So maybe it is fair to strike a compromise between maximizing $E[U \mid A]$ and maximizing $E[U]$. What kind of compromise one should strike is an empirical question. Even better, one might try to model the process of how proposals get drafted in more detail.

Do these concerns pose a serious threat to the prospects of our model? To some extent the answer is yes. But notice that other assessments of decision rules rely exactly on the same assumptions. We will explain this by considering the well-known Penrose measure. The Penrose measure equals the proportion of all possible voting profiles on which a country's vote is pivotal. A country's vote is pivotal in the case where, had the country voted differently, the outcome of the vote would have been different.

The formal relation between our model and the power measure account is as follows. As we explain at the end of the Appendix, the expected utilities of the single countries, $E[U_i]/\sigma$, are proportional to the Penrose measures for the Gaussian models with $\mu/\sigma = 0$. In this sense, the Penrose measures are contained in our framework for a special choice of the model parameters. (The Shapley–Shubik index, by the way, can also be interpreted as an expected utility; see Shapley, 1953.)

Therefore, it is justified to say that the Penrose measure presupposes a certain probabilistic model. This is a probability distribution according to which the countries vote in a probabilistically independent way and according to which each country has an equal chance of voting for and voting against a proposal; i.e. $p_+ = p_- = 0.5$ in our terminology. But why should this be the case? Our model is a bit more general in this respect. True, we adopt a fixed value for p_+ for all countries, but we do not assume that $p_+ = p_-$. Furthermore, the voting power approach is far from taking into account concerns similar to the last two concerns mentioned above.

Moreover, the Penrose measure does not take into account the expected benefits if one is on the winning side and the expected costs if one is on the losing side (Steunenberg et al., 1999: 342). Suppose, for instance, that there are several poor countries in the federation that have debts and an underdeveloped infrastructure. If a poor country comes out on the winning side and receives subsidies from the EU, it lacks the administrative structure to distribute them efficiently, so that the money usually gets lost; i.e. its benefits are small. On the other hand, if a poor country comes out on the losing side and has to pay into the federation, then it suffers under the weight of the debts; i.e. its costs are large. If this were the situation, then it would be reasonable to favour more selective decision rules than if benefits and costs are roughly symmetrical or if benefits are large and costs are small. However, since the Penrose measure abstracts from this information, it is not capable of differentiating between decision rules on the basis of this information.

Apart from the different degrees of generality, there is another fundamental difference between the power measure account and our utilitarian criterion. Whereas the voting power account appeals to equality and thus demands equalizing the voting powers of all citizens, our evaluation is based upon welfare and aims at maximizing expected utility for an average European person. In this way there is disagreement between both evaluations even at the point where the model assumptions coincide (the Gaussian model with $\mu/\sigma = 0$). We will illustrate this by discussing a lately developed argument based on voting power.

In a recent article, Kirsch et al. (2004; see also Slomczynski and Zyczkowski, 2004) argue in support of P62, saying that this voting rule provides equal voting power to each citizen in the European Union. Their argument was developed in the section on alternative decision rules. Now the P62 rule may well equalize voting power as measured by the Penrose measure, but where the assumptions coincide ($p_+ = 0.5$, $v_+/|v_-| = 1$) it is SMP that yields the highest expected utility, not the P62. Providing equal voting power is simply a very different desideratum from maximizing expected utility.

It is not our intention to argue against the voting power approach at this point. Thus, we will leave it open whether equalizing voting power or maximum welfare is the more valuable desideratum. Note, however, that on both desiderata the P62 is to be preferred to the Accession Treaty as well as to the current Constitution.

Let us finally point to some other characteristics of our model. Although we adopt quite simplistic assumptions, the features of our model are surprisingly rich. It is intriguing to observe how the ranking between the different decision rules is affected when the empirical input is changed. Moreover, a qualitative story can be told to make at least some of the results intuitively plausible. Examining the expected utilities of single countries, checking for Pareto-dominance and investigating the phase diagrams provide a qualitative understanding of how alternative decision rules in a federation affect the resulting utility distribution. In future research, we hope to investigate how robust our results are if the model assumptions are relaxed.

Notes

We acknowledge constructive criticism by three referees. We thank the Alexander von Humboldt Foundation, the German Federal Ministry of Education and Research and the Program for the Investment in the Future (ZIP) of the German Government for supporting this research.

1 The Appendix can be found on the EUP web page.
2 Here we have given a simpler, although equivalent, formulation of the decision rule implemented in the current Draft Constitution for 25 countries.

For future enlargements, (con.b) is replaced by a clause according to which at least 55% of the member countries have to vote in favour of the proposal.

3 Averaging utilities only means averaging over some population (we could work with the sum instead). It does not involve any expectation value yet. In the following we will sometimes speak of the utility of an average person. This is only an abbreviation for the average effect in terms of utilities for some group of persons. We also assume that we can always average over a fixed number of persons for each country; that is, the population sizes stay constant and are not affected by the adoption of the proposals.

4 See note 3.

5 More generally, we have $E[U] = E[U|A]P(A) + E[U|\neg A]P(\neg A)$, but here $E[U|\neg A] = 0$, since every country gets zero utility if the proposal is rejected.

6 Some might want to object that the Commission acts according to its own interests only, and that setting $\mu/\sigma > 0$ is too optimistic. But one of the Commission's most crucial interests will certainly consist in not being defeated in the CM. Moreover, even if μ/σ were merely larger than –0.1, our results would not essentially change.

References

Albert, Max (2003) 'The Voting Power Approach: Measurement without Theory', *European Union Politics* 4(3): 351–66.

Banzhaf, John F. (1965) 'Weighted Voting Doesn't Work: A Mathematical Analysis', *Rutgers Law Review* 19: 317–34.

Barberà, Salvador and Matthew O. Jackson (2004) 'On the Weights of Nations: Assigning Voting Power to Heterogeneous Voters', unpublished working paper, URL (consulted Sept. 2004): http://www.hss.caltech.edu/~jacksonm/weight.pdf.

Bilbao, Jesús M. et al. (2004) 'Letter to the Governments of the EU Member States', URL (consulted Sept. 2004): http://www.esi2.us.es/~mbilbao/pdffiles/letter.pdf.

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

Felsenthal, Dan S. and Moshé Machover (2001) 'Enlargement of the EU and Weighted Voting in its Council of Ministers', URL (consulted Sept. 2004): http://www.lse.ac.uk/Depts/cpnss/projects/VPPpdf/eubooktd.pdf.

Felsenthal, Dan S., Dennis Leech, Christian List and Moshé Machover (2003) 'In Defence of Voting Power Analysis: Responses to Albert', *European Union Politics* 4(4): 473–97.

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

Kirsch, Werner, Moshé Machover, Wojciech Slomczynski and Karol Zyczkowski (2004) 'Voting in the EU Council – A Scientific Approach', URL (consulted Sept. 2004): http://www.ruhr-uni-bochum.de/mathphys/publikationen/voting.pdf.

Laruelle, Annick and Mika Widgrén (1998) 'Is the Allocation of Voting Power among the EU States Fair?', *Public Choice* 96: 317–39.

Napel, Stefan and Mika Widgrén (2004) 'Power Measurement as Sensitivity Analysis – A Unified Approach', *Journal of Theoretical Politics* 16(4): 517–38.

Penrose, Lionel S. (1946) 'The Elementary Statistics of Majority Voting', *Journal of the Royal Statistical Society* 109: 53–7.
Penrose, Lionel S. (1952) *On the Objective Study of Crowd Behaviour*. London: H.K. Lewis & Co.
Schweizer, Urs (1990) 'Calculus of Consent: A Game-theoretic Perspective', *Journal of Institutional and Theoretical Economics* 146: 28–54.
Shapley, Lloyd S. (1953) 'A Value for *n*-Person Games', in Harold W. Kuhn and Albert W. Tucker (eds) *Contributions to the Theory of Games, Vol II*, pp. 307–17. Princeton, NJ: Princeton University Press.
Shapley, Lloyd S. and Martin Shubik (1954) 'A Method for Evaluating the Distribution of Power in a Committee System', *American Political Science Review* 48: 787–92.
Slomczynski, Wojciech and Karol Zyczkowski (2004) 'Rules Governing Voting in the EU Council – The Treaty of Nice, the Draft Treaty Establishing a Constitution of Europe and a Proposed Compromise Solution', URL (consulted Sept. 2004): http://www.phys.uni.torun.pl/~milosz/PDF/SZ2ang.pdf.
Steunenberg, Bernhard, Dieter Schmidtchen and Christian Kobold (1999) 'Strategic Power in the European Union: Evaluating the Distribution of Power in Policy Games', *Journal of Theoretical Politics* 11(3): 339–66.
Widgrén, Mika (1994), 'Voting Power in the EC Decision Making and the Consequences of Two Different Enlargements', *European Economic Review* 38: 1153–70.

About the authors

Claus Beisbart is a postdoctoral fellow in the Philosophy, Probability and Modeling Research Group, Center for Junior Research Fellows, University of Konstanz, PO Box M 682, D-78457 Konstanz, Germany.
Fax: +49 7531 88 4898
E-mail: Claus.Beisbart@uni-konstanz.de

Luc Bovens is Professor in the Department of Philosophy, Logic and Scientific Method, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK, and Director of the Philosophy, Probability and Modeling Research Group at the University of Konstanz.
Fax: +44 20 7955 6845
E-mail: Luc@bovens.org

Stephan Hartmann is Reader in Philosophy in the Department of Philosophy, Logic and Scientific Method, and Director of the Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK.
Fax: +44 20 7955 6845
E-mail: S.Hartmann@lse.ac.uk