Welfarist Evaluations of Decision Rules under Interstate Utility Dependencies

Claus Beisbart and Stephan Hartmann

2008

Currently under revision for *Social Choice and Welfare*.

Abstract

We provide welfarist evaluations of decision rules for federations of states and consider models, under which the interests of people from different states are stochastically dependent. We concentrate on two welfarist standards; they require that the expected utility for the federation be maximized or that the expected utilities for people from different states be equal. We discuss an analytic result that characterizes the decision rule with maximum expected utility, set up a class of models that display interstate dependencies and run simulations for different dependency scenarios in the European Union. We find that, under positive correlations, the welfare distribution tends to be less sensitive to the choice of the decision rule, whereas it can be important under negative correlations. The results that Beisbart and Bovens (SCW 29, p. 581, 2007) have found for two types of models without interstate dependencies are relatively stable. There are exceptions, though, under which the way the welfare distribution is shaped by a decision rule is significantly affected by dependencies.

1 Introduction

Recent times have seen a growing interest in welfarist evaluations of decision rules (see Beisbart et al. 2005, Coelho 2005, Barberà & Jackson 2006 and Beisbart & Bovens 2007, for instance). Welfarist approaches evaluate alternative decision rules for federations of states (counties, constituencies, ...) by comparing the resulting welfare distributions. The aim of this paper is to extend the welfarist approach by relaxing crucial assumptions on which much work so far is based. We consider federations for which the interests of the people from different states are stochastically dependent.

In order to understand the welfarist approach, consider the following problem: A *federation* of states has a *decision board* where *representatives* vote on proposals drafted by some committee. Each representative represents a state. The decisions of the board are taken according to a *decision rule*. The problem is: Which decision rule can be recommended for the federation?

This problem has many real world applications. Recently, the Council of Ministers in the European Union (CM, for short; see Felsenthal & Machover 2000 for an overview) has received much attention in political as well as in academic circles. In the CM, representatives of the EU states vote on proposals drafted by the European Commission. The question is: Which decision rule is best for the EU?

In order to judge a decision rule better than another, we need a standard of evaluation. *Welfarist* evaluations of decision rules rely on the welfare distribution that results from a decision rule. They are consequentialist, and ask the question: cui bono – to whose benefit?

People can benefit from decision rules in the following way (cf. Beisbart et al. 2005, Sec. 1): If a single proposal passes, some people will take profits, whereas other people might have to take some costs. For instance, if import tariffs are imposed on sugar beets in the EU, sugar farmers in the EU will take profits, whereas sugar consumers in the EU will have to pay. Now, whether the proposal about import tariffs passes, depends on the decision rule. The crucial question is whether states that vote in favor of the tariffs have sufficiently many votes to push it through. In this way, the decision rule indirectly affects the welfare distribution of the people in the EU.

Of course, it doesn't make sense to assess a decision rule on the base of a single proposal. Rather, the question is, how the welfare distribution is affected in the long run, on the base of many proposals that are not yet known in detail. We will thus evaluate alternative decision rules assuming a probability model for proposals, and the welfare of a person will be quantified using an expected utility.

In this paper we will consider two welfarist *standards* for decision rules (cf. Beisbart & Bovens 2007, p. 582):

- $\mathcal{U}$ Decision rule R is pro tanto better than rule S, if the expected utility for an average person in the EU is larger (see Dancy 2004, Ch. 2 for the notion of the "pro tanto").

- $\mathcal{E}$ Decision rule R is pro tanto better than rule S, if there is more equality in the distribution of the expected utilities for different people across the EU.

Whereas, under $\mathcal{U}$, the best decision rule maximizes expected utility, under $\mathcal{E}$, the best decision rule minimizes inequality.

Very recently, decision rules have been evaluated using these standards. Beisbart et al. (2005) compare a few alternative decision rules regarding $\mathcal{U}$ for the CM. Barberà & Jackson (2006) generally specify the decision rule that maximizes expected utility. Beisbart & Bovens (2007) consider both $\mathcal{U}$ and $\mathcal{E}$. However, all of these works rely on the assumption that the utilities from proposals for people from different states are stochastically independent. It is thus excluded that people from Spain, say, are typically proposed similar utilities as people from Portugal are.

Assuming stochastic interstate independence was certainly justified for developing the welfarist approach. But, as a matter of fact, the assumption is not very plausible for many real world applications. We will therefore relax the assumption. Our approach is explorative, i.e. we will not search for an empirically adequate model of some decision board. Rather, we will set up a general framework for modeling dependencies, consider a variety of *dependency scenarios* and check how the rankings of alternative decision rules are affected. Beisbart & Hartmann (2006) also introduce dependencies, but they assume a multivariate Gaussian for the utilities of the people, which does not seem to be very realistic, since higher-order correlation effects are neglected. Also, they only compare the performances of a few decision rules.

We will focus on the CM as an example. However, the methods in this paper can be applied to any decision board. Also, our results indicate effects to be expected for other boards as well.

Alternative approaches to assess decision rules focus on the distribution of influence that political agents have in virtue of their votes. Influence is often quantified in terms of the Banzhaf measure of voting power (see Felsenthal & Machover 1998 for an introduction and Felsenthal & Machover 2000 for an application to the CM). The Banzhaf measure for a voter equals the probability that her vote is pivotal. For calculating the Banzhaf measure one assumes that the votes of different people are stochastically independent. Thus, a stochastic independence assumption is crucial for the so-called standard voting power approach as well. The independence assumptions in the welfarist and in the standard voting power frameworks are related, since there is an analytical connection between expected utilities and voting power (Beisbart et al. 2005, Appendix and Beisbart & Bovens 2007; for the conception of voting power see also Laruelle & Valenciano 2005). So our work is also interesting for the voting power approach. For probability models with dependent votes see Good & Mayer (1975), Chamberlain & Rothschild (1981) and Gelman et al. (2004).

The plan of the paper is as follows: In Sec. 2, we briefly outline the welfarist framework. In Subsec. 2.2, a general analytic result regarding expected utility is stated. This result, however, does not take us very far in practice. We thus turn to models in Sec. 3. The expected utilities under these models cannot be calculated analytically any more, so we carry out simulations (Sec. 4). Our results are described in Secs. 5 and 6. We draw our conclusions in Sec. 7.

2 The welfarist framework

In order to evaluate decision rules on welfarist grounds, we need to know how decision rules affect the welfare distribution (cf. Schweizer 1990 and Beisbart et al. 2005 for the framework).

2.1 Basic definitions

Assume that the federation has n people. Label the people in the federation with numbers ranging from 1 to n. Let there be N states in the federation and label them from 1 to N. In the following, we will use lowercase letters for labels of people, whereas uppercase letters will be used for label of states. For $I = 1, .., N$, let K_I be the set of labels for people from the Ith state. Set $n_I \equiv \# K_I$. Of course, $\sum_I n_I = n$.

Consider now a single proposal. The proposal can either be accepted or rejected. Without loss of generality, we assume that citizen i receives zero utility, if the proposal is rejected, and a utility v_i, if the proposal is accepted for every i. We assume interpersonal comparability of utility scales and form the average utility for citizens from state I from the proposal:

$$\overline{v}_I = \frac{1}{n_I} \sum_{i \in K_I} v_i \ . \tag{1}$$

The average utility from the proposal accross the federation, call it $\overline{v}$, is then

$$\overline{v} = \frac{1}{n} \sum_{I=1}^{N} n_I \overline{v}_I \ . \tag{2}$$

We model proposals in a probabilistic way and think of the v_is as values of random variables V_i. Similarly, $\overline{v}_I$ is a value of random variable $\overline{V}_I$ for $I = 1, .., N$, and $\overline{v}$ is a value of random variable $\overline{V}$. The random variables V_i have $p(v_1, .., v_n)$ as their joint probability density. In our analysis we take p to be exogeneously fixed. p induces a joint probability density over the random variables $\overline{V}_I$ $(I = 1, .., N)$, call it $p\left(\overline{V}_1, .., \overline{V}_N\right)$.

The vote of state I's representative is modeled as another random variable Λ_I. Λ_I takes the value $\lambda_I = 1$, if the representative votes in favour of the proposal (votes yes, for short). It takes the value $\lambda_I = 0$, if the representative votes against the proposal (votes no, for short). Vectors $(\lambda_1, .., \lambda_N)$ are called voting profiles.

How do the representatives vote and what is their strategy? In this paper we assume that a representative will vote yes, iff the average utility for her state I, $\overline{v}_I$ is positive. That is, $\lambda_I = \theta\left(\overline{v}_I\right)$, where θ is a variation of the well-known

Heaviside step function:[1]

$$\theta(x) \equiv \begin{cases} 0 & \text{i}f \quad x \leq 0 \\ 1 & \text{i}f \quad x > 0 \ . \end{cases} \tag{3}$$

If the utilities $\overline{v}_I$ are known prior to the decision and if the decision rule is monotonic[2], this choice of strategies constitutes a Nash equilibrium (Beisbart & Bovens 2007, Sec. 2).[3]

A decision rule R can be thought of as a function D^{R} that maps voting profiles into $\{0,1\}$: $D^{\mathrm{R}} : \{-1,1\}^N \to \{0,1\}$. We set D^{R} at 1 for acceptance, and at 0 for rejection. There are $2^{\left(2^N\right)}$ decision rules possible.

We introduce other sets of random variables U_i^{R}, $\overline{U}_I^{\mathrm{R}}$ and $\overline{U}^{\mathrm{R}}$. They denote a utility that is received, once a decision has been taken following rule R. For instance, U_i^{R} is the utility for person i, after a decision has been taken. In more detail,

$$U_i^{\mathrm{R}} = D^{\mathrm{R}} \times V_i \ , \quad \overline{U}_I^{\mathrm{R}} = D^{\mathrm{R}} \times \overline{V}_I \ , \quad \overline{U}^{\mathrm{R}} = D^{\mathrm{R}} \times \overline{V} \ . \tag{4}$$

In the following, we will often suppress the upper label R.

We can now calculate the expected utilities that a person i derives from a decision rule R under the probability distribution $p\left(v_1, .., v_n\right)$ over proposals, call it $E\left[U_i^{\mathrm{R}}\right]$. It reads:

$$E\left[U_i^{\mathrm{R}}\right] = \int \mathrm{d}v_1 \cdots v_n \ p\left(v_1, .., v_n\right) v_i D^{\mathrm{R}}\left(\theta\left(\overline{v}_1\right), .., \theta\left(\overline{v}_N\right)\right) \ , \tag{5}$$

where the $\overline{v}_I$s depend on the v_js as specified in Eq. (1). In a similar way, we obtain for $E\left[\overline{U}_I^{\mathrm{R}}\right]$

$$E\left[\overline{U}_I^{\mathrm{R}}\right] = \int \mathrm{d}\overline{v}_1 \cdots \mathrm{d}\overline{v}_N \ p\left(\overline{v}_1, .., \overline{v}_N\right) \overline{v}_I D^{\mathrm{R}}\left(\theta\left(\overline{v}_1\right), .., \theta\left(\overline{v}_N\right)\right) \tag{6}$$

and for $E\left[\overline{U}^{\mathrm{R}}\right]$:

$$E\left[\overline{U}^{\mathrm{R}}\right] = \frac{1}{n} \sum_{I=1}^{N} n_I E\left[\overline{U}_I^{\mathrm{R}}\right] \ . \tag{7}$$

[1]We thus assume that the representative of state I votes no, if $\overline{v}_I = 0$. Of course, this assumption is in a way arbitrary, since voting yes in this case would not make a difference for state I. Nevertheless, our choice does not make a difference, if $\overline{v}_I = 0$ has zero probability measure. That this is so is a very natural assumption.

[2]A rule R is monotonic, if $D^{\mathrm{R}}\left(\lambda_1, .., \lambda_N\right) \geq D^{\mathrm{R}}\left(\lambda'_1, .., \lambda'_N\right)$, in case $\lambda_I \geq \lambda'_I$ for $I = 1, .., N$. A monotonic rule induces a simple voting game, which is characterized by set-theoretic monotonicity (Felsenthal & Machover 1998, Def. 2.1.1 on p. 11).

[3]Barberà & Jackson (2006) are more general at this point by not assuming any specific strategy.

2.2 An analytical result for expected utility

Let us now focus on expected utility. Proposals come in following a fixed probability distribution. The best thing one can do in terms of expected utility is the following: Proposals with $\overline{v} > 0$ are accepted and proposals with $\overline{v} < 0$ are rejected. The result is

$$E\left[\overline{V}\theta\left(\overline{V}\right)\right] \ , \tag{8}$$

where the expectation value is calculated on the base of $p\left(\overline{v}_1, .., \overline{v}_n\right)$. This is the maximum utility that one can draw from the proposals (Barberà & Jackson 2006, p. 325).

However, this procedure does not follow a decision rule, since no vote is taken. A decision rule looks at the votes and not at the utilities. Decisions according to a rule lose information present in the utilities (ib.). The expected utility for a decision rule will therefore in general be smaller than $E\left[\overline{V}\theta\left(\overline{V}\right)\right]$.

This can be illustrated as follows. For an arbitrary rule R, we calculate $E\left[\overline{U}^{\mathrm{R}}\right]$ by conditioning on all possible voting profiles $(\lambda_1, .., \lambda_N)$:

$$E\left[\overline{U}^{\mathrm{R}}\right] = \sum_{\lambda_1, ..\lambda_N} D^{\mathrm{R}}\left(\lambda_1, .., \lambda_N\right) E\left[\overline{V}\middle|\theta(\overline{V}_1) = \lambda_1, .., \theta(\overline{V}_N) = \lambda_N\right] p\left(\lambda_1, .., \lambda_N\right) \ , \tag{9}$$

where $p\left(\lambda_1, .., \lambda_N\right)$ is the probability for a particular voting profile $(\lambda_1, .., \lambda_N)$. It can be calculated from the joint probability distribution over the $\overline{V}_I$s. If $\overline{V}$ can take positive and negative values with finite probabilities, each, $E\left[\overline{V}\middle|\theta(\overline{V}_1) = \lambda_1, .., \theta(\overline{V}_N) = \lambda_N\right]$ in Eq. (9) can be expressed as

$$\begin{aligned} &E\left[\overline{V}|\theta\left(\overline{V}_1\right) = \lambda_1, .., \left(\overline{V}_N\right) = \lambda_N \wedge \overline{V} > 0\right] p\left(\overline{V} > 0|\lambda_1, .., \lambda_N\right) \\ +&E\left[\overline{V}|\theta\left(\overline{V}_1\right) = \lambda_1, .., \left(\overline{V}_N\right) = \lambda_N \wedge \overline{V} < 0\right] p\left(\overline{V} < 0|\lambda_1, .., \lambda_N\right) \ . \end{aligned} \tag{10}$$

Here, the first addend is positive, whereas the second is negative. Looking at the value of $\overline{V}$ itself always picks the positive term for every voting profile. But a decision rule has to yield either acceptance or rejection for each particular voting profile. If the rule yields acceptance for a voting profile $(\lambda_1, .., \lambda_N)$, then the second, negative addend in Eq. (10) is bought in as well. If the rule yields rejection, then the positive contribution from the first addend is missed. Thus, in either case, utilities are lost.

Obviously, the best *decision rule* R that one can think of, yields acceptance if $E\left[\overline{V}\middle|\lambda_1, ..\lambda_N\right]$ (or Eq. 10) is positive, and it yields rejection if this term is negative. This follows from the proof that Barberà & Jackson (2006) give for their Theorem 1.

As Barberà & Jackson (2006) show, the decision rule that maximizes expected utility can be described in a very elegant way, if the $\overline{V}_I$s are stochastically independent. Call this rule the BJ solution. In case of dependent utilities $\overline{V}_I$, this is in general not possible any more. The optimal decision rule might just be a list that assigns each voting profile 1 or 0 according to what we have said. Call this the generalized BJ solution.

It can be doubted whether the generalized BJ solution is helpful in practice. First, it can be very difficult to obtain sufficient data for estimating the $E\left[\overline{U}\middle|\lambda_1, ..\lambda_N\right]$s. Second, if the generalized BJ solution cannot be cast in simple terms, voting itself will become difficult. There will be no way to keep in mind the voting rule, so one can only look up the result from the list, once the votes are taken. The result might come as a surprise for the representatives themselves. If the generalized BJ solution is not monotonic – and this is very well possible – then $\lambda_I = \theta\left(\overline{v}_I\right)$ might not be a Nash equilibrium any more, and the representatives might have to reconsider their strategies.

For these reasons, we will not consider this analytical result further. Instead, we confine ourselves to a family of voting rules that can be cast in simple terms. In this paper we consider weighted voting rules (cf. the notion of a weighted voting game, Def. 2.3.14 in Felsenthal & Machover 1998, pp. 29–30). Under a weighted voting rule, the representative of state I has a weight w_I for $I = 1, .., N$, and a proposal is accepted, iff the sum of weights associated with the yes-votes exceeds a certain threshold of acceptance t. A weighted voting rule R can be represented as $D^{\mathrm{R}} = \theta\left(\sum_I \lambda_I w_I - t\right)$. In the following, we restrict ourselves to weights that only depend on population size n_I. Moreover, it is assumed that the weights depend on population in the following way: (Bovens & Hartmann 2007, Eq. 1):

$$w_I \propto (n_I)^{\alpha} \ . \tag{11}$$

Here, the exponent α is a measure of degressive proportionality. $\alpha = 1$ yields proportional weights, $\alpha = 0$ yields equal weights, and $\alpha \in (0, 1)$ parameterizes weightings in between. Without loss of generality, we will also assume that the voting weights are normalized: $\sum_I w_I = 1$. Our question is then: Which pairs of numbers (α, t) do best wrt to $\mathcal{U}$ and $\mathcal{E}$, respectively? And is there an (α, t)-rule that one can recommend on both counts?[4]

3 Models for proposals

We will also confine ourselves to a particular class of probabilistic models for proposals. They are inspired by Crain et al. (1993).[5]

[4]Note that two pairs $(\alpha, t) \neq (\alpha', t')$ yield the same decision rule, if they induce the same function D (cf. Barberà & Jackson 2006, p. 324). Nevertheless, pairs (α, t) will sometimes loosely be called rules.

[5]The model by Crain et al. (1993) is about fractions of votes. We are concerned with utilities instead. This is of significant advantage. For the model by Crain et al. (1993) leads to finite (but minimal) probabilities that the fraction of votes for one candidate is larger than 1. The reason is that Gaussian probability densities for fractions of votes are assumed. Our model, which is about utilities, does not suffer from such a problem.

For setting up the class of models, we partition the federation into *groups of states* (abbreviated as „G“, labeled with uppercase Greek letters Γ, Δ). That is, each state is member of exactly one group. The groups need not be organized in any way. Rather the idea is that people from the same group have dependent utilities, even if they are from different states.

The utilities from proposals for person i, i.e. the values v_i of the random variables V_i are now determined as a sum of stochastically independent addends:

$$v_i = \chi_\Gamma^{\mathrm{G}} + \epsilon_I^{\mathrm{S}} + \epsilon_i^{\mathrm{P}} \; . \tag{12}$$

Here, the first addend, χ_Γ^{G}, is common to the people from the same group, the second addend, ϵ_I^{S}, is common to the people from the same state, and the third addend is specific of the individual person i. In Eq. (12) we assume that i lives in state I ($i \in K_I$), and state I is in group Γ. Let us now fix the details.

- χ_Γ^{G} is a *group utility tendency.* We assume a joint probability distribution over the χ_Γ^{G}s, $p^{\mathrm{G}}(\chi_1^{\mathrm{G}}, \chi_2^{\mathrm{G}}, ...)$. This probability distribution will always assumed to be a multivariate Gaussian. Under this multivariate Gaussian, each χ_Γ^{G} has zero mean. The standard deviations for the different χ_Γ^{G}s are identically set at σ^{G}, and the covariances are identically set at ϱ^{G}. This leaves us with a covariance matrix with $\left(\sigma^{\mathrm{G}}\right)^2$ as diagonal elements, and ϱ^{G} as off-diagonal elements. Since the covariance matrix has to be positive semidefinite, we obtain the following constraints on ϱ^{G}: $\varrho^{\mathrm{G}} / \left(\sigma^{\mathrm{G}}\right)^2 \in [-1/\left(N_G - 1\right), 1]$, where N_G is the number of groups.

- $\chi_\Gamma^{\mathrm{G}} + \epsilon_I^{\mathrm{S}}$ is a *national utility tendency.* ϵ_I^{S} is randomly drawn from a Gaussian with zero mean and a standard deviation σ_I^{S}. It is independent from χ_Γ^{G}. The national utility tendencies are thus scattered around the group utility tendencies χ_Γ^{G}. For simplicity, we will assume that the ϵ_I^{S} follow the same probability density, i.e. the σ_I^{S}s are identical for all states I: $\sigma_I^{\mathrm{S}} = \sigma^{\mathrm{S}}$.

- ϵ_i^{P} encodes a utility contribution that is completely personal. The utilities for citizens from state I are scattered around the national utility tendency of their state. ϵ_i^{P} is supposed to be the value of a random variable that is independent from any other random variable. For simplicity, we assume that ϵ_i^{P} is normally distributed with zero mean and a standard deviation $\sigma_i^{\mathrm{P}} = \sigma^{\mathrm{P}}$ that is constant throughout the federation.

Beisbart & Bovens (2007) focus on models with no interstate dependencies ($\sigma^{\mathrm{G}} = \varrho^{\mathrm{G}} = 0$) and consider two types of federations. Each state in a federation might be an *aggregate*, i.e. the utilities of people from the same state are independent. Alternatively, each state in a federation might be an *interest group*, i.e. the utilities of people from the same state are perfectly correlated. Beisbart & Bovens (2007) do not consider interstate dependencies. Thus their *aggregate model* is a special model in our class, where $p^{\mathrm{G}}\left(\chi_1^{\mathrm{G}},..\right)$ factorizes into $\delta\left(\chi_1^{\mathrm{G}}\right) \times \delta\left(\chi_2^{\mathrm{G}}\right)$ ... and where $\sigma^{\mathrm{P}} = 0$ and $\sigma^{\mathrm{S}} \neq 0$. Their *interest group model* corresponds to $p^{\mathrm{G}}\left(\chi_1^{\mathrm{G}},..\right) = \delta\left(\chi_1^{\mathrm{G}}\right) \times \delta\left(\chi_2^{\mathrm{G}}\right)$ ..., $\sigma^{\mathrm{S}} = 0$ and $\sigma^{\mathrm{P}} \neq 0$.

In this paper, we will start with the aggregate and the interest group models and switch on interstate dependencies. In this way modifications of the aggregate/interest group model arise. For clarity, the aggregate/interest group model defined above will often be called *default* aggregate/interest group model henceforth. Each specific model that we consider can be fixed by taking four decisions:

- Chose $\sigma^{\mathrm{S}} = 0, \sigma^{\mathrm{P}} = 1$ (for modifications of the aggregate model) or $\sigma^{\mathrm{S}} = 1, \sigma^{\mathrm{P}} = 0$ (for modifications of the interest group model).

- Fix a partition of states.

- Fix the correlation matrix for $p^{\mathrm{G}}\left(\chi_1^{\mathrm{G}},..\right)$, i.e. fix σ^{G} and ϱ^{G}.

Our model and the assumptions are constrained by the following idea: The marginal probability densities for the personal utilities V_i are identical. The underlying idea is as follows: In fixing the probability model for proposals, there is no reason to assume that the proposals are biased towards any person in the federation.

Clearly, for some real world federations, unbiasedness is a realistic assumption. However, in examining generic scenarios with dependencies, it is reasonable to start with the assumption of unbiasedness.

Symmetry. Start with a rule (α, t) and suppose that the threshold can not be exactly met by a coalition. That is, there is no subset $\mathcal{I} \subseteq \{1, .., N\}$ with $\sum_{I \in \mathcal{I}} w_I = t$. Then the rule $(\alpha, 1-t)$ yields exactly the same welfare distribution as (α, t). This result generalizes Proposition 1 in Beisbart & Bovens (2007), and the proof parallels their proof.

Suppose now that $N > 1$, $t \neq 0, 1$ and $\alpha \geq 0$. Consider the set of $(n_1, .., n_N)$-vectors with the following property: $n_I \neq 0$ for $I = 1, .., N$ and there is a subset $\mathcal{I} \subseteq \{1, .., N\}$ with $\sum_{I \in \mathcal{I}} (n_I)^{\alpha} = t$ (i.e. there is a coalition that exactly meets the threshold). It can be shown that this set has zero measure in $\mathbb{R}^N$. As a consequence, we do not expect that a threshold t can be exactly met in a given federation. Therefore, we expect that t and $(1-t)$ produce the same welfare distributions for most federations, unless $t = 0, 1$. We will therefore only consider thresholds $t \geq .5$.

Scaling. Our class of models have the following scale invariance. If we multiply all model parameters (i.e., σ^{G}, ϱ^{G},σ^{S} and σ^{P}) by some $\alpha > 0$, our results scale linearly with α as well. So effectively, one model parameter drops out.

4 How results are obtained

In order to obtain results, we carry out simulations. Utility vectors are randomly drawn, the voting profile is determined, and the collective decision is taken. If there is acceptance, the proposed utilities are distributed; if there is rejection, they are not so distributed. For saving computation time, we do not simulate the utilities of different persons i, i.e. the v_is. Rather, the $\overline{v}_I$s are drawn randomly. Given values of the group utility tendencies, $\chi_1^{\mathrm{G}}, ..$, the random variables $\overline{V}_I$ are independent. The conditional probability density for values $\overline{v}_I$ of $\overline{V}_I$, $p\left(\overline{v}_I|\chi_1^{\mathrm{G}}, \chi_2^{\mathrm{G}}, ..\right)$ is a Gaussian with mean χ_Γ^{G} and with a standard variation $\sqrt{\left(\sigma^{\mathrm{S}}\right)^2 + \frac{1}{n_i^{\mathrm{S}}}\left(\sigma^{\mathrm{P}}\right)^2}$.

In our simulations, we obtain the $E[\overline{U}_I]$s, from which the average expected utility, $E[\overline{U}]$, is determined. For applying our egalitarian standard, we quantify the distance from equality in the $E[\overline{U}_I]$s by calculating a weighted scatter (cf. Beisbart & Bovens 2007). The idea is that the inequality is larger, if a larger state is an outlier wrt the $E[\overline{U}_I]$s. If f_I is the population fraction of state I, then our measure of inequality In is defined via

$$\mathrm{In}^2 = \sum_I f_I E[\overline{U}_I]^2 - E[\overline{U}]^2 \, . \tag{13}$$

Note that this is in general not the expectation value of the scatter in the u_is. This expectation value is difficult to obtain.

We checked the reliability of our simulations by comparing to analytical results for a federation of 15 states under the default aggregate and interest group models. For such a federation, it is still possible to obtain analytic results by summing up 2^{15} addends (Beisbart et al. 2005, Appendix). We find that 2m realizations do reasonably well for 15 states. We thus run 2m realizations of votes for our results.

For each model considered, we scan the (α, t) space. We often work with a spacing of .02 along the α- and .02 along the t-dimension. We do never consider $\alpha > 1$ or $\alpha < 0$. Obviously, our results have only the accuracy of our grid, and it is possible, that a rule that has $\alpha > 1$ is optimal wrt U or E. Moreover, errors due to the finite number of simulations might affect our results.

In the following, results for $E[\overline{U}]$ and the measure of inequality are always normalized by

$$E\left[\overline{V}\theta\left(\overline{V}\right)\right] \, . \tag{14}$$

As we showed in Subsec. 2.2, this quantity is an upper limit on $E\left[\overline{U}\right]$. We call the normalized expected utility efficiency, and the normalized measure of inequality normalized spread (of utilities).

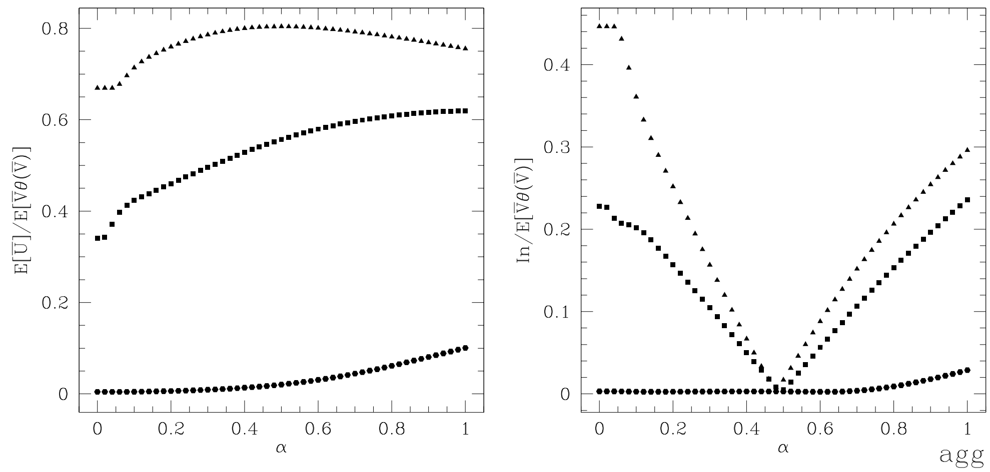

Figure 1: The welfare distribution for the default aggregate model for the EU (27 members). Efficiency (left column) and normalized spread (right column) are shown as function of the α-parameter for different thresholds: Hexagons: $t = .8$; pentagons: $t = .6$; squares: $t = .5$. The results are based on simulations with 2m realizations.

5 Results for models without interstate dependencies

We will now present results. We start with very simple models and move on to more complicated models. Let us first mention results for the default aggregate interest group models (Beisbart & Bovens 2007). Models of this kind do not yet display interstate dependencies ($\sigma^{\mathrm{G}} = \varrho^{\mathrm{G}} = 0$). Interstate dependencies are discussed in the next section.

5.1 The default aggregate model

Under the default aggregate model, utilities from proposals are independent for any two persons, regardless whether they are from the same state or not. Thus, $\sigma^{\mathrm{S}} = 0$, whereas $\sigma^{\mathrm{P}} \neq 0$. Because of the scale invariance, we can set σ^{P} at 1. The default aggregate model is the fixed-size-block model with one person per block and no bias in Barberà & Jackson (2006), pp. 326, 328–9.

Under the default aggregate model, the distributions for the average utilities $\overline{V}_I$ from proposals are different for the different states I. They are Gaussians that peak at 0 and that have a standard deviation of $\sigma^{\mathrm{P}}/\sqrt{n_i}$. Thus, smaller states have a larger spread in their average utility from proposals.

Results for are shown in Fig. 1. The efficiency (left panel) and the normalized spread (right panel) are shown for various decision rules in the (α, t)-space. We plot the efficiency and the normalized spread as functions of α. Different point styles designate different thresholds.

The rule ($\alpha = .5, t = .5$), also known as Penrose 50, yields maximum expected utility for the federation. Penrose 50 does not only maximize expected utility in the constrained class of (α, t)-rules, but also in the class of all decision rules (Barberà & Jackson 2006, pp. 327 and 329 f.). For a qualitative understanding, note that the quantities $E\left[\overline{V}_I|\overline{V}_I > 0\right]$ and $E\left[\overline{V}_I|\overline{V}_I < 0\right]$ are proportional to the standard deviation of $\overline{V}_I$. As a consequence, people from smaller states tend to benefit more than people from larger states, given they benefit. Equally, they suffer more than larger ones, given they suffer. Consequently, what citizens from the smaller states get, makes a larger difference to the expected utility $E[U]$ than under the default interest group model, e.g. Therefore, a decision rule has to strike a fair compromise between larger and smaller states, if it is to maximize expected utility. For two simple mathematical arguments why Penrose 50 maximizes expected utility for the default aggregate model, see Beisbart & Bovens (2007), Sec. 5.

Our measure of inequality has a local minimum around $\alpha = .5$ for thresholds not too far from $t = .5$, which is also the global minimum in our range of α-values. The reason why there is not perfect equality for equal weights ($\alpha = 0$) is again that people from smaller states tend to benefit more, given they benefit, etc. Thus, under equal weights, their $E\left[\overline{U}_I\right]$s are larger than those of larger states. Thresholds that are very close to 0 or 1 produce almost flat curves close to zero.

Remarkably, under the default aggregate model, one rule – Penrose 50 – is almost optimal on both of our welfarist standards: It yields maximum expected utility and does astonishingly well in terms of equality, given the large range of In-values in the right panel of Fig. 1. Note, however, that, on our scans and simulations, Penrose 50 is not strictly the best rule regarding equality. Curves for high thresholds yield lower minima on our simulations.

5.2 The default interest group model

Under the interest group model, on a particular proposal, people from the same state have exactly the same utility. Utilities for people from different states are independent and follow the same probability density. The interest group model originates, if $\sigma^{\mathrm{P}} = 0$. Without loss of generality, σ^{S} can be set at one. Note that, under the interest group model and all of its modifications in this paper, the $\overline{V}_I$s follow the same probability density. Our default interest group model is the fixed-number-of-blocks model with one block per state and no bias in Barberà & Jackson (2006), pp. 326, 328–9.

Results are shown in Fig. 2. For the thresholds shown, the expected utility mostly goes up, as α increases (left panel).[6] The reason is that, under the default

[6] Here and at similar occurrences, "mostly" or "for most parts" means: in a large fraction

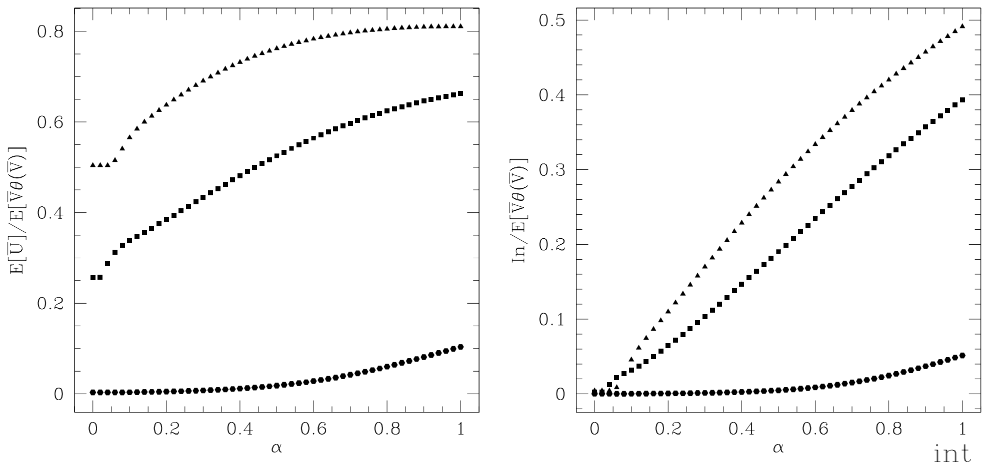

Figure 2: The default interest group model. For point styles see the caption of Fig. 1

interest group model, each person has the same benefits from proposals, if she benefits. Since there are more people in larger states, the net utility is higher, if larger states have higher weights, as long as $\alpha \leq 1$. Among the thresholds shown in this Figure, $t = .5$ does best. The maximum expected utility is obtained at for a very high α and a threshold $t = .5$. Indeed, the rule $(\alpha = 1, t = .5)$ represents the BJ solution.

Regarding our measure of inequality (right panel), $\alpha = 0$ is optimal for any threshold t – at this point the states have the same weights and thus according to the default interest group model the same expected utilities. As α increases, the differences between the different states' weights become larger, and thus their expected utilities do so as well (cf. Proposition 1 on p. 331 in Barberà & Jackson 2006). For $t \geq .5$, larger thresholds yield less inequality – the higher the threshold, the more states have to agree with the outcome of the decision, and thus the smaller are the differences between them.

Altogether, for the default interest group model, our welfarist standards pull into different directions. Maximizing expected utility requires one to adopt high values of α and thresholds close to .5. Minimizing inequality requires one to adopt low values of α and thresholds close to 1 or 0.

of the $\alpha \in [0, 1]$ range, the function under consideration does something, as we move from one α to the next in our spacing. There might be exceptions, but that there are is not implied. Note that really the curves for efficiencies and normalized spread are step functions. It is also possible that the functions jump back and forth on small scales, if they are sufficiently resolved.

5.3 The transition between the default aggregate and the default interest group model

Clearly, the default aggregate and interest group models are idealizations. Under realistic circumstances, citizens from one state will have some similarities in interests, such that the utilities from proposals are dependent. But we do not expect full correlations; rather the utilities from proposals will fluctuate within one state. Thus, a model between the default aggregate and interest group models seems more realistic.[7]

Within our class of models, a transition between the default aggregate and the interest group models can be effected. Either we start with the default interest group model ($\sigma^{\mathrm{P}} = 0$) and let intrastate fluctuations grow ($\sigma^{\mathrm{P}} > 0$). As $\sigma^{\mathrm{P}}/\sigma^{\mathrm{S}}$ approaches infinity, we end up with the aggregate model (in different units of utilities). Alternatively, we can start with the default aggregate model ($\sigma^{\mathrm{S}} = 0$), and let the σ^{S} parameter grow. Beisbart & Bovens (2007) investigate a different transition between the models – they have a different class of models that connects both models. They assume that the joint probability density for the people's utilities from proposals in one state follows a multivariate Gaussian. Correlations are then entirely due to two-points effects. Our class of models is more realistic, since they allow for non-trivial higher-order correlations, under which all people in the federation are affected at the same time.

Suppose now, we set both σ^{P} and σ^{S} at non-zero values and consider a fixed voting rule. We know that the $E\left[\overline{U}_I\right]$s only depend on the probability density for the $\overline{V}_I$s. It can be shown that, in the moments of this probability density, terms with a σ^{P} are damped with a factor of $\frac{1}{\sqrt{n_I}^{\nu}}$, where $\nu \geq 1$ is a natural number. Thus, unless $\sigma^{\mathrm{P}}/\sqrt{n_I} \gtrsim \sigma^{\mathrm{S}}$, the results are pretty much the same as for the default interest group model. Put differently, the aggregate model is unstable; already a σ^{S} that is significantly smaller than σ^{P} can push the results to something like the results for the interest group model. This is also confirmed by means of simulations.

We thus follow Beisbart & Bovens (2007) in concluding that the interest group model is the generic model to look at. Nevertheless, in the following, we will consider modifications of both types of models. But we will not investigate in detail how the transitions between the modifications of the aggregate and the interest group models look like. That is, we either set σ^{S} or σ^{P} at zero.

[7] Apart from the default aggregate and interest group models and transitions between them, there is a third model defined by $\sigma^{\mathrm{P}} = \sigma^{\mathrm{S}} = 0$. But such a model is not interesting: Every person in the federation is always proposed the same utility $u_i \equiv u$.

partition	$\alpha = 0$	$\alpha = .5$	$\alpha = 1$
P0	1	1	1
P1	.85/.15	.66/.34	.46/.54
P2	.48/.52	.60/.40	.73/.26

Table 1: Basic characteristics of the three partitions P0–P2 that we investigate below. The numbers are the fractions of weights that a group has for a given α. Thus, in the first column, the fractions of the numbers of states are given. In the last column, the fractions of populations are given. For P1, the first number refers to the group of small states. For P2, the first number refers to the group of western states.

6 Results for models with interstate dependencies

We now switch on dependencies between the utilities that are proposed for citizens from different states. For this, we partition the EU into groups. On each proposal, citizens from states of the same group get the same group utility tendency χ_Γ^{G}.

For simplicity, under each partition, we will consider only one value of σ^{G}. For each partition and each σ^{G}, three values of ϱ^{G} are investigated, viz. $\varrho^{\mathrm{G}} = 0$, the maximal and the minimal possible value of ϱ^{G} . We consider three different partitions. Under the partition P0, the whole federation forms one big group. Under P1, the four largest states form one group, the other states form another group. Under P2, the groups have nothing to do with population: There are two groups that contain the more western/eastern states (the western/eastern states, for short). The eastern states comprise Cyprus, Estonia, Slovenia, Latvia, Lithuania, Finland, Slovakia, Bulgaria, Austria, Hungary, the Czech Republic, Greece, Romania and Poland.

We list important characteristics of the groups under the partitions in Table 1. Results for the decision rules that do best in terms of $\mathcal{U}$ and $\mathcal{E}$, respectively, are listed in Table 2, where the different partitions and models are considered.

6.1 Modifications of the aggregate model with dependencies

We first consider modifications of the aggregate model. Qualitatively, the picture is as follows: The utilities that people are proposed have a group-specific and a personal component. States do not have any significance regarding the utilities proposed. They can be thought of as random parts of groups.

In the following models, σ^{G} is always set at 2×10^{-4}. The idea is that the utility contribution from the group utility tendency χ_Γ^{G} fluctuates significantly

agg, default, $\mathcal{U}$	$\alpha = .5, t = .5$		
agg, default, $\mathcal{E}$	high t		
agg with $\sigma^{\mathrm{G}} = 2 \times 10^{-4}$	$\varrho^{\mathrm{G}} = 0$	$\varrho^{\mathrm{G}} = \left(\sigma^{\mathrm{G}}\right)^2$	$\varrho^{\mathrm{G}} = -\left(\sigma^{\mathrm{G}}\right)^2$
P0, $\mathcal{U}$	$\alpha = .52, t = .5$	–	–
P0, $\mathcal{E}$	high t	–	–
P1, $\mathcal{U}$	$\alpha = .7, t = .5$	$\alpha = .52, t = .5$	$\alpha = .9, t = .5$
P1, $\mathcal{E}$	high t	high t	high t
P2, $\mathcal{U}$	$\alpha = .5, t = .5$	$\alpha = .52, t = .5$	$\alpha = .78, t = .5$
P2, $\mathcal{E}$	high t	high t	high t
int, default, $\mathcal{U}$	$\alpha = 1, t = .5$		
int, default, $\mathcal{E}$	high t		
int with $\sigma^{\mathrm{G}} = .4$	$\varrho^{\mathrm{G}} = 0$	$\varrho^{\mathrm{G}} = \left(\sigma^{\mathrm{G}}\right)^2$	$\varrho^{\mathrm{G}} = -\left(\sigma^{\mathrm{G}}\right)^2$
P0, $\mathcal{U}$	$\alpha = .66, t = 0.5$	–	–
P0, $\mathcal{E}$	high t	–	–
P1, $\mathcal{U}$	$\alpha = .88, t = .5$	$\alpha = .62, t = .5$	$\alpha = 1, t = .5$
P1, $\mathcal{E}$	$\alpha = .76, t = .92$	high t	high t
P2, $\mathcal{U}$	$\alpha = .74, t = .5$	$\alpha = .62, t = .5$	$\alpha = 1, t = .5$
P2, $\mathcal{E}$	high t	high t	high t

Table 2: A summary of the decision rules that do best on our standards $\mathcal{U}$ and $\mathcal{E}$, respectively. Upper half: aggregate model and its modifications. Lower half: interest group model and its modifications. For the default models, $\sigma^{\mathrm{G}} = \varrho^{\mathrm{G}} = 0$; for the other models, three parameter choices are considered. For the default models and standard $\mathcal{U}$, the BJ solution is shown, otherwise the results rely on simulations that scan the (α, t) space. Only thresholds up to $t = .98$ were considered. For high t, the results for alternative values of α are very close and probably not accurate enough as to allow for a meaningful comparison. Thus no optimal α is given. Note that thresholds t and $(1 - t)$ give almost always the same results; thus, if (α, t) is optimal in some respect, (α, t) is also very likely to be optimal in the same respect.

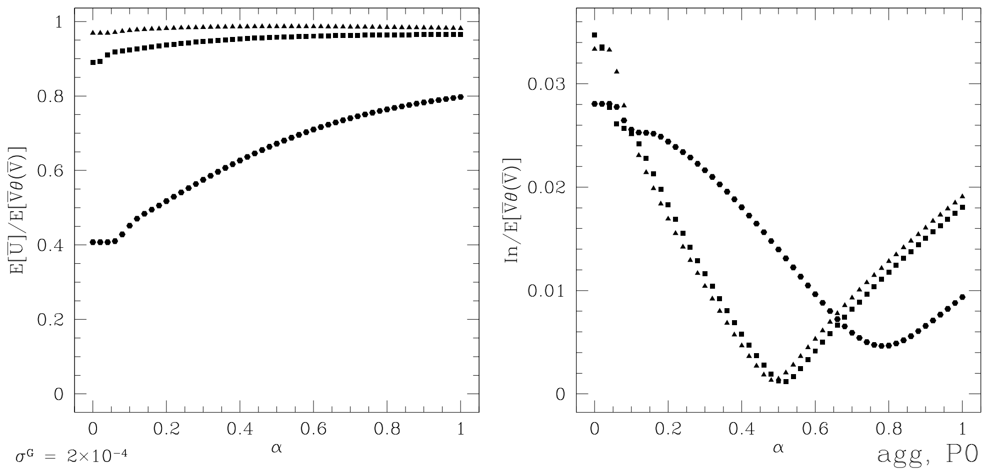

Figure 3: Results for a modification of the aggregate model under partition P0 (the states in the EU form one big group). Efficiencies (left panel) and normalized spread (right panel) are shown as functions of α.

less than the contribution for individual persons. For Ireland, the standard deviation in group utility tendency is 40% of the standard deviation of $\overline{V}_I$ under the default aggregate model. We also looked at slightly different values σ^{G} and found qualitatively similar results.

6.1.1 Partition P0

Under P0, there is only one group. There are thus no off diagonal elements ϱ^{G}. Only σ^{G} can be varied.

Under P0, proposals have a utility socket that changes from proposal to proposal, but is the same for each person from each state for a specific proposal. Results are shown in Fig. 3. If we compare to our default aggregate model, we observe that the decision rules are more efficient for all rules shown. The high efficiencies can be explained using Eq. (10): If the utilities $\overline{V}_I$ for states are positively correlated, voting profiles under which most states cast the same vote carry large probabilistic weight. Profiles with many states voting yes yield acceptance under most of the rules considered here, and indicate a $\overline{v} > 0$. Thus, losses due to the decision rules are small. Likewise, profiles with many states voting no carry large probabilistic weight, yield rejection under most of the rules considered here, and indicate a $\overline{v} < 0$. Thus, losses due to missed positive utilities are small.

For many rules, the normalized spread in utilities is smaller than under the default aggregate model (right panel). Part of the explanation is that the normalizing $E\left[\overline{V}\theta\left(\overline{V}\right)\right]$ is larger now, but this does not fully account for the new results for the normalized spread. As far as the ranking of rules is concerned, there is not much change. A rule close to Penrose 50 maximizes expected utility, and, for thresholds $t = .5, .6$, there is a local minimum in the spread around $\alpha = .5$, which is also the global minimum for these curves and the α-range considered. We conclude that the recommendations for the default aggregate model prove stable under the modification that we have considered. In particular, a rule close to $(\alpha = .5, t = .5)$ does very well on both welfarist standards.

6.1.2 Partition P1

Under P1, there are two groups, where one group comprises the four largest states (large states, for short, in this subsection and in Subsec. 6.2.2). This group has more population, but less states than the other group with the small states (see Table 1 for details).

In the first row of Fig. 4, we set σ^{G} at 2×10^{-4} and ϱ^{G} at 0. Utilities for people from different groups are thus independent. Again, the effects are not too dramatic. The efficiencies go up in comparison to the default aggregate model for all rules shown. Decision rules with very high thresholds still yield significantly lower expected utilities than a threshold of .5. The curve for $t = .5$ has still a maximum, which is now at $\alpha = .7$. This is also the overall maximum in the (α, t)-space.

The normalized spreads cover approximately the same range of values as for $\sigma^{\mathrm{G}} = 0$. There is still a significant local minimum in all curves shown. The minima are shifted, comparing to the default aggregate model. Since the peak in the expected utility and the minimum in the measure of inequality shift in the same direction, we can come up with a political recommendation that does very well on both welfarist standards: a rule in in the vicinity of $(\alpha = .6, t = .5)$.

For the second row of Fig. 4 ϱ^{G} is set at its maximal possible value. The utilities for people from different groups are now positively correlated. The results resemble the results that we showed for partition P0. This is no surprise, since in both cases, positive dependencies between utilities for people from different states abound.

The case of $\varrho^{\mathrm{G}} < 0$ is more interesting, since the groups compete with each other: If people from one group are proposed some good, people from the other group will typically be proposed some costs. Results for the minimal possible value of ϱ^{G} are shown in the third row of Fig. 4. Low values of α produce very small efficiencies. For higher values of σ^{G} and for $\varrho^{\mathrm{G}} = -\left(\sigma^{\mathrm{G}}\right)^2$, $E\left[\overline{U}\right]$ is even below zero for some rules. The reason is that, under P1, for small values of α, the group with fewer people (the group with the small states) has more weights than the thresholds $t = .5, .6, .8$ require. Proposals that are accepted thus typically put benefits (costs) on the people from the small (large) states. Since, altogether, more people live in the group with the large states, the net expected utility, $E\left[\overline{U}\right]$, is negative for $t = .5, .6, .8$. For the same reason, the spread in the utilities is very large for rules with thresholds $t = .5, .6$ and $\alpha \approx 0$. As α increases, things mostly become better, though, for all thresholds shown and on both standards. We have

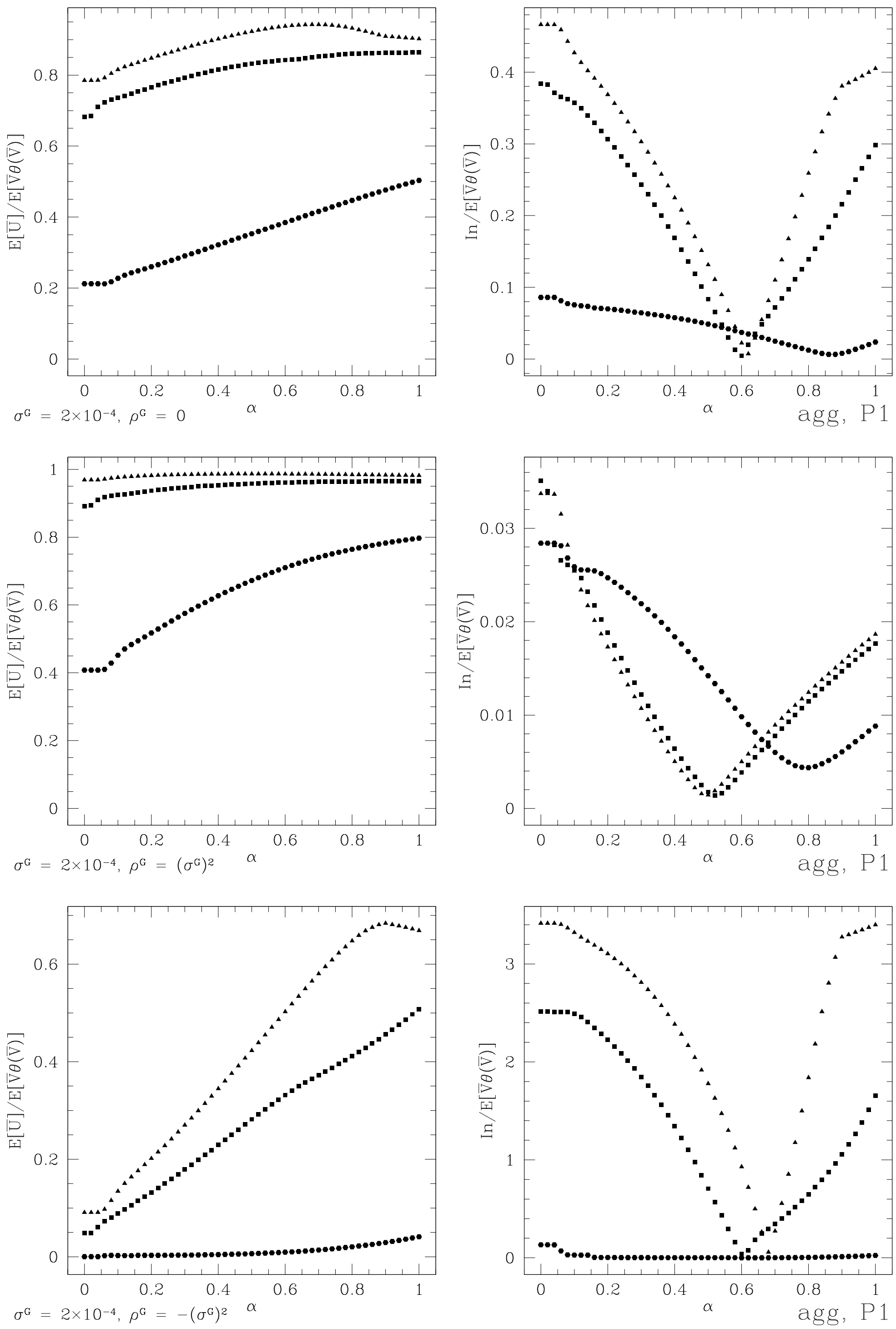

Figure 4: Results for modifications of the aggregate model under partition P1 (two groups with large and small states, respectively).

a maximum expected utility around ($\alpha = .9, t = .5$) (see Table 2). The measure of inequality for $t = .5$ has a very sharp local minimum, which is also the global minimum of this curve in $\alpha \in [0, 1]$. Unfortunately, it is not at .9, which means that there is no rule that is very good regarding both welfarist standards.

In order to check the stability of our results we also consider variations of partition P1. If the group with the small states is split into two subgroups, where one subgroup contains states with less than 10m people, the results are similar as under P1. Under a second variation, call it P1', Italy is counted as a small state such that the group with the large states does not have the majority of the people any more. For the minimal possible ϱ^{G}, the results are qualitatively different between P1 and P1'. For instance, whereas, under P1, equal weights ($\alpha = 0$) yield less efficiency than proportional weights ($\alpha = 1$) for $t = .5, .6, .8$, it is exactly the other way round under P1'. We conclude that, under anticorrelated groups, where group membership is defined in terms of population, it is very important to know which group has the majority of the population.

6.1.3 Partition P2

Under partition P2, group membership and size are unrelated. As one can see from Tab. 1, the group with the western states has about 75% of the EU population, but less states than the other group.

The most noteworthy results can be seen from Fig. 5. In the first row, $\sigma^{\mathrm{G}} = 2 \times 10^{-4}$ and $\varrho^{\mathrm{G}} = 0$. The efficiencies are quite high. The curve for $t = .5$ is still larger than the other curves for any value of α shown. The overall maximum is still the Penrose 50 rule.

The measure of inequality is very different from everything we have observed for the aggregate model and its modifications so far. For small values of α, the normalized spread is small; indeed, for each t shown, it is smaller than under P1. Then, for most parts of the curves $t = .5, .6$, the inequality increases as a function of α.

Why don't the expected utilities $E[U_I]$ almost equalize around $\alpha = .5$ under this model for $t = .5, .6$ as they do under the default aggregate model? The reason is as follows: For each group, there is an α-value in the range $\alpha \in [.5, .7]$ at which the $E[U_I]$s come very close for states I *within the group*. But for these values of α, people from the western states are better off than those from the eastern ones, and the $E[U_I]$-value at which the expected utilities for people from the western states come very close is significantly higher than it is for the people from the eastern states. Part of the reason should be that, for $\alpha \geq .5$, the western states have significantly more votes than the eastern states do.

In order to obtain a low measure of inequality, small values of α are required – at least for $t = .5, .6$. Thus, for this partition and this model, it is very difficult to be very good on both welfarist standards at the same time.

If an additional $\varrho^{\mathrm{G}} > 0$ is introduced, the efficiencies are often higher, but there is not much qualitative change comparing to $\varrho^{\mathrm{G}} = 0$. However, if the ϱ^{G}-parameter is sufficiently high, the measure of inequality changes qualitatively. The curves for $t = .5, .6$ display a local minimum close to .5 again, which is also the global minimum in the range $\alpha \in [0, 1]$ each time. This makes it easier to be very good on both welfarist standards at the same time.

The last row shows results for the minimal possible $\varrho^{\mathrm{G}} < 0$. In this case, the curves for efficiency start at comparably low values. As under P1, voting rules with a small α are not efficient, since, under P2, the group with the larger population has less states, and the majority of people will often have to pay for the benefits of people from the eastern states. As soon as α is sufficiently high, this changes very quickly. However, as α increases, the inequality mostly increases as well, at least if $t = .5., 6$. So it is again not possible to be very good on both welfarist standards at the same time.

6.2 Modifications of the interest group model with dependencies

Let us now consider modifications of the interest group model. The qualitative picture is as follows: On each particular proposal, people from the same state are proposed the same utility. This utility has two parts that are specific of the group and the state, respectively. As an example, we will set $\sigma^{\mathrm{G}} = .4$. Thus, σ^{G} is 40% of the standard deviation for $\overline{V}_I$ for all states I under the default interest group model. For Ireland, the change in the standard deviation is exactly the same as under the modifications of the aggregate model. The idea is that the utility contributions from the group utility tendencies $\chi_{\Gamma}^{\mathrm{G}}$ are smaller than the contributions for the national utility tendencies, but that both contributions are of the same order of magnitude. The standard deviation for the proposed utilities $\overline{V}_I$ is then $\sqrt{1 + .4^2}\sigma^{\mathrm{S}} \approx 1.08\sigma^{\mathrm{S}} = 1.08$. We also checked results for slightly different values of σ^{G} and found qualitatively similar results.

6.2.1 Partition P0

Results for the partition P0 are shown in Fig. 6. From the left hand panel, we observe that the efficiencies are larger under this model than under the default interest group model for all rules shown. The explanation why the efficiencies are high is the same as under the P0 modification of the aggregate model that we have considered.

As under our aggregate model modification P0, the threshold $t = .5$ yields always higher expected utility than the other thresholds shown. As functions of α, the $E\left[\overline{U}\right]$ curves are flatter for $t = .5, .6$. The overall maximum for the expected utility is around $(\alpha \approx 0.66, t = 0.5)$ (and not at $\alpha = 1, t = .5$ any more), but there is not a pronounced peak.

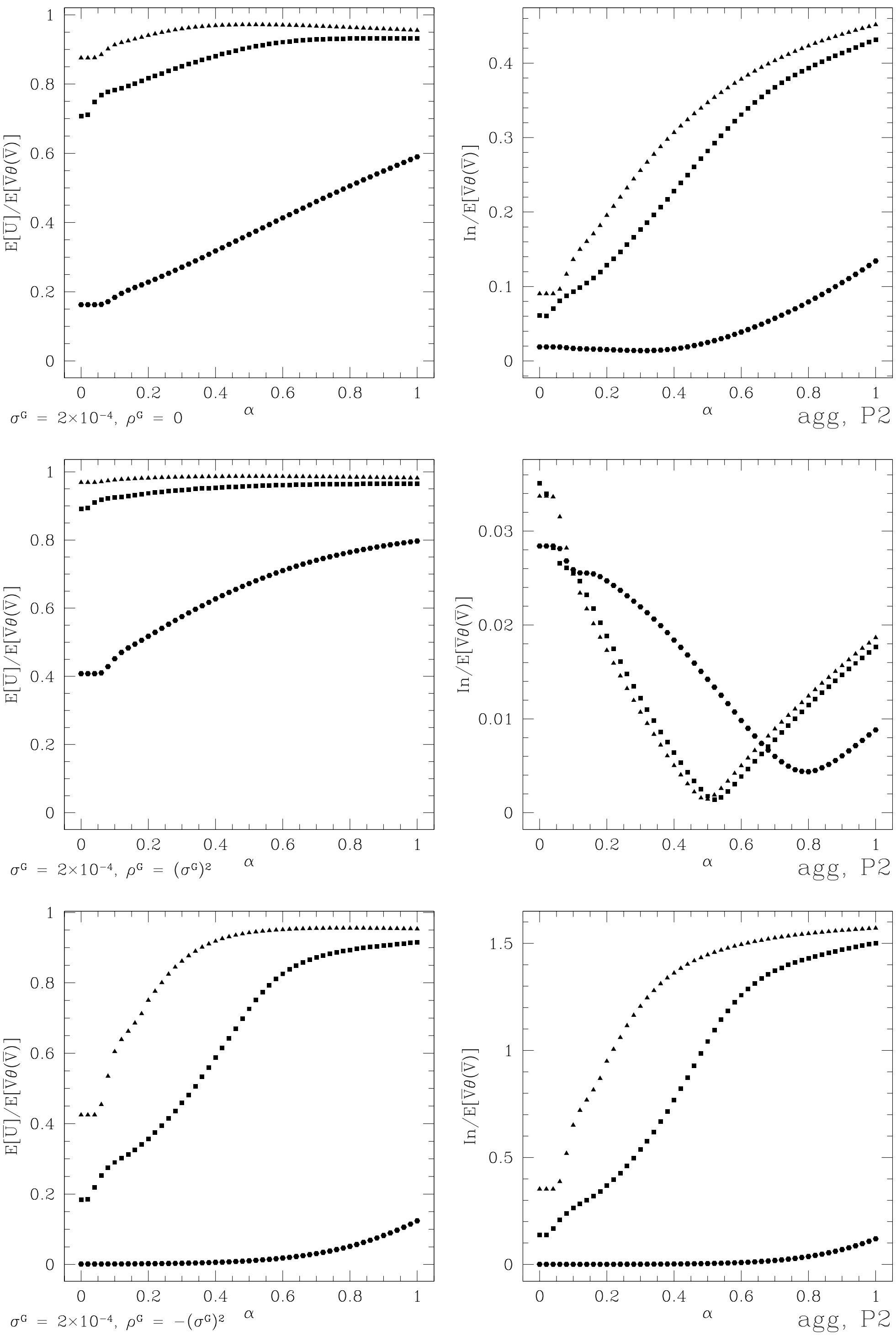

Figure 5: Results for modifications of the aggregate model under partition P2 (western vs. eastern states).

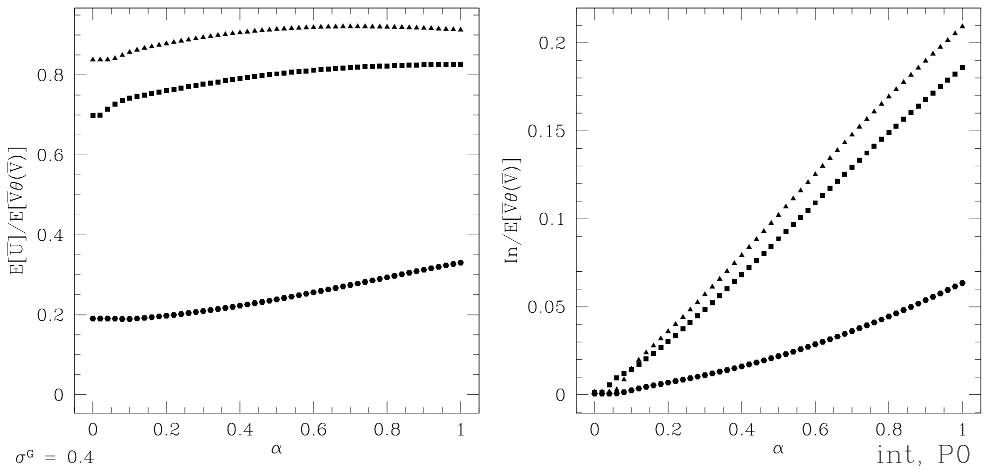

Figure 6: Results for a modification of the interest group model under P0 (the states in the EU form one big group).

The normalized spread for the most part becomes larger, as α increases. Because of the symmetry of the model, $\alpha = 0$ still yields minimum (and zero) inequality, independently from the value of t. All in all, a reasonable welfarist compromise may consist in $(\alpha = 0, t = .5)$ for this model. This yields zero inequality and a relatively high expected utility.

6.2.2 Partition P1

We now turn to modifications of the interest group model, under which large and small states form a group, each (P1, states from the group with larger states are again called large states, etc. in this subsection). Results for the partition can be seen from Fig. 7.

For $\sigma^{\mathrm{G}} = .4$ and $\varrho^{\mathrm{G}} = 0$ (first row), all voting rules shown are again more efficient than under the default interest group model. Also, for $t = .5, .6$ the α-dependence of the curves is not as strong any more as under the default interest group model. The location of the maximum efficiency has shifted from $(\alpha = 1, t = .5)$ (default) to $(\alpha = .88, t = .5)$.

The results for the normalized spread are completely different from the results under the default interest group model. Equality is not established at $\alpha = 0$ any more. Rather, each curve for $t \in \{.5, .6, .8\}$ has a local minimum around $\alpha = .5$. For higher values of σ^{G} and $\varrho^{\mathrm{G}} = 0$, the minimum shifts further to the right. As far as α is concerned, it is easier to do reasonably well on both standards than it is under the default interest group model, since the maxima in the $E\left[\overline{U}\right]$-$\alpha$ curves and the minima in the In-α curves are closer. Unfortunately, the standards pull into different directions regarding the threshold.

Let us partly explain some of these results for $t = .5, .6$: For small α-values,

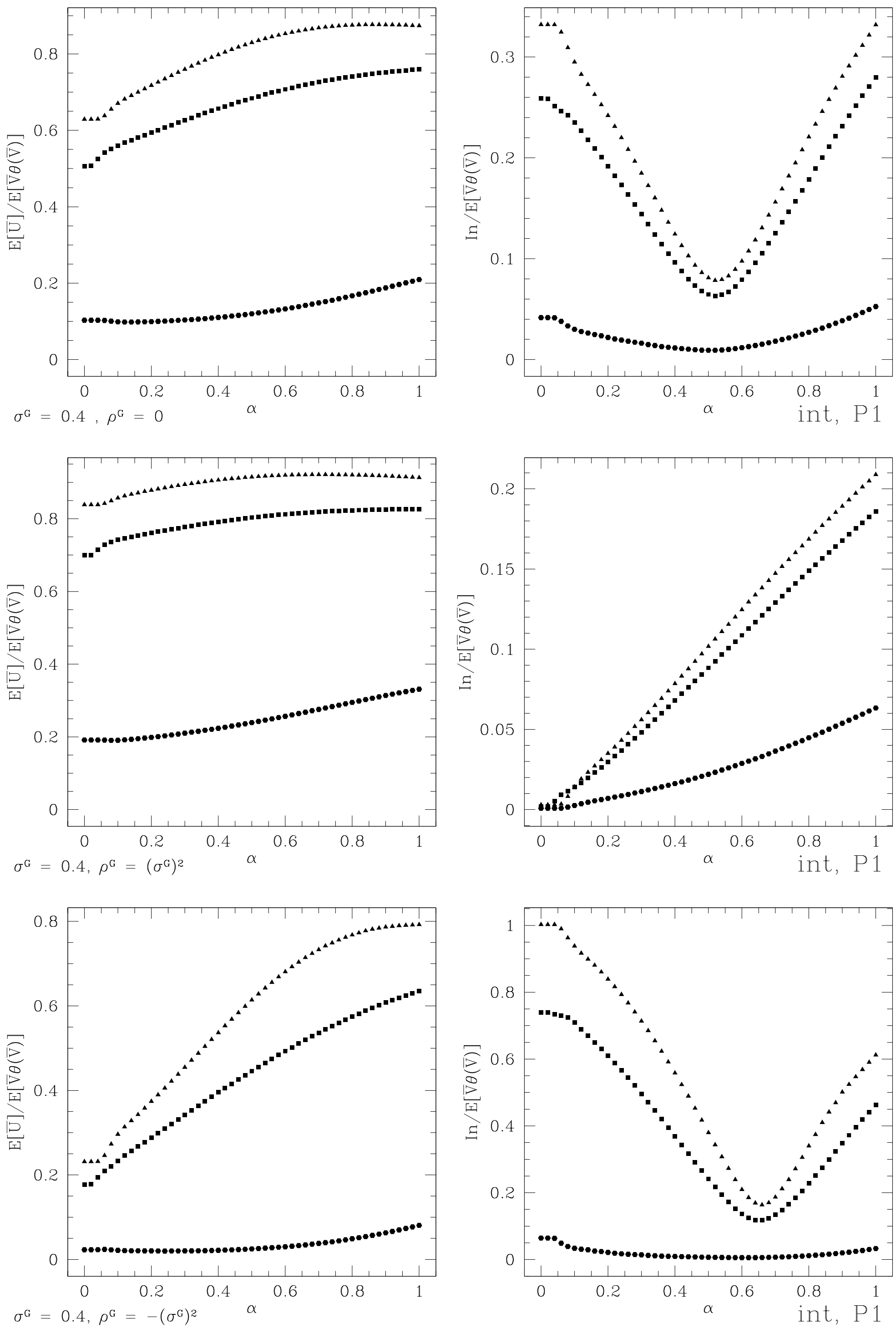

Figure 7: Results for modifications of the interest group model under P1 (large vs. small states).

the sum of weights from the group with the small states exceeds the threshold. Thus, under the probability model under consideration, the people from small states are likely to get what is in their interest. The votes of the large states, on the opposite, will often not make a difference for the outcome of the vote, and people from the large states are not that likely to get what is in their interest. This should at least be part of the reason why people from the large states receive significantly less benefits, and the spread in expected utilities is comparably high. As α increases, the large states are assigned higher weights. As a consequence, as α increases, the benefits increase for people from the group with the large states, whereas they decrease for people from many of the small states, at least if α is sufficiently large. Since the large states have the majority of the population, this is good for the whole federation: $E\left[\overline{U}\right]$ grows, as α increases, and at the same time inequality goes down. But for very high values of α, the situation for $\alpha = 0$ is reversed: Any person from any large state take more profits than any person from any small states does.

Let us now switch on a maximal possible $\varrho^{\mathrm{G}} = .16$ (second row). The efficiency curves for $t = .5, .6$ are very flat. Efficiencies are higher for $\varrho^{\mathrm{G}} = .16$ than for $\varrho^{\mathrm{G}} = 0$ for each rule shown. As a function of α, the normalized spread starts with small values and mostly becomes larger for all t-values shown. Altogether, this situation is quite favorable to the federation, and the choice of a decision rule is not that critical any more because the flatness of some efficiency curves. Note that there is qualitative similarity with the results for the P0 modification of the interest group model that we have considered.

Things are different, if we keep σ^{G} at .4 and set ϱ^{G} at the minimal possible value. As the third row in Fig. 7 shows, the efficiencies are lower than under $\varrho^{\mathrm{G}} = 0$ for all rules shown, whereas the normalized spread is often very high. For the low efficiencies at $\alpha = 0$, the same explanation as under the corresponding P1 modification of the aggregate model applies. Rules with $\alpha \approx .7$ do comparably well on both welfarist standards for many thresholds t. But the choice of a threshold is still dilemmatic.

Our results for P1 are again quite stable, if the group with the small states is split into two groups as before (cf. end of Subsec. 6.1.2). There is more change in the results, if we compare to a partition under which the group with the large states does not have the majority of the people (P1'), particularly for the minimal ϱ^{G}. One difference is efficiency, where the curves for $t = .5, .6$ are not that steep functions of α any more.

6.2.3 Partition P2

Consider now modifications of the interest group model under partition P2, where the western and the eastern states form groups, each (see Tab. 1 for the partition and Fig. 8 for results).

If the utilities for people from different groups are independent ($\varrho^{\mathrm{G}} = 0$, first

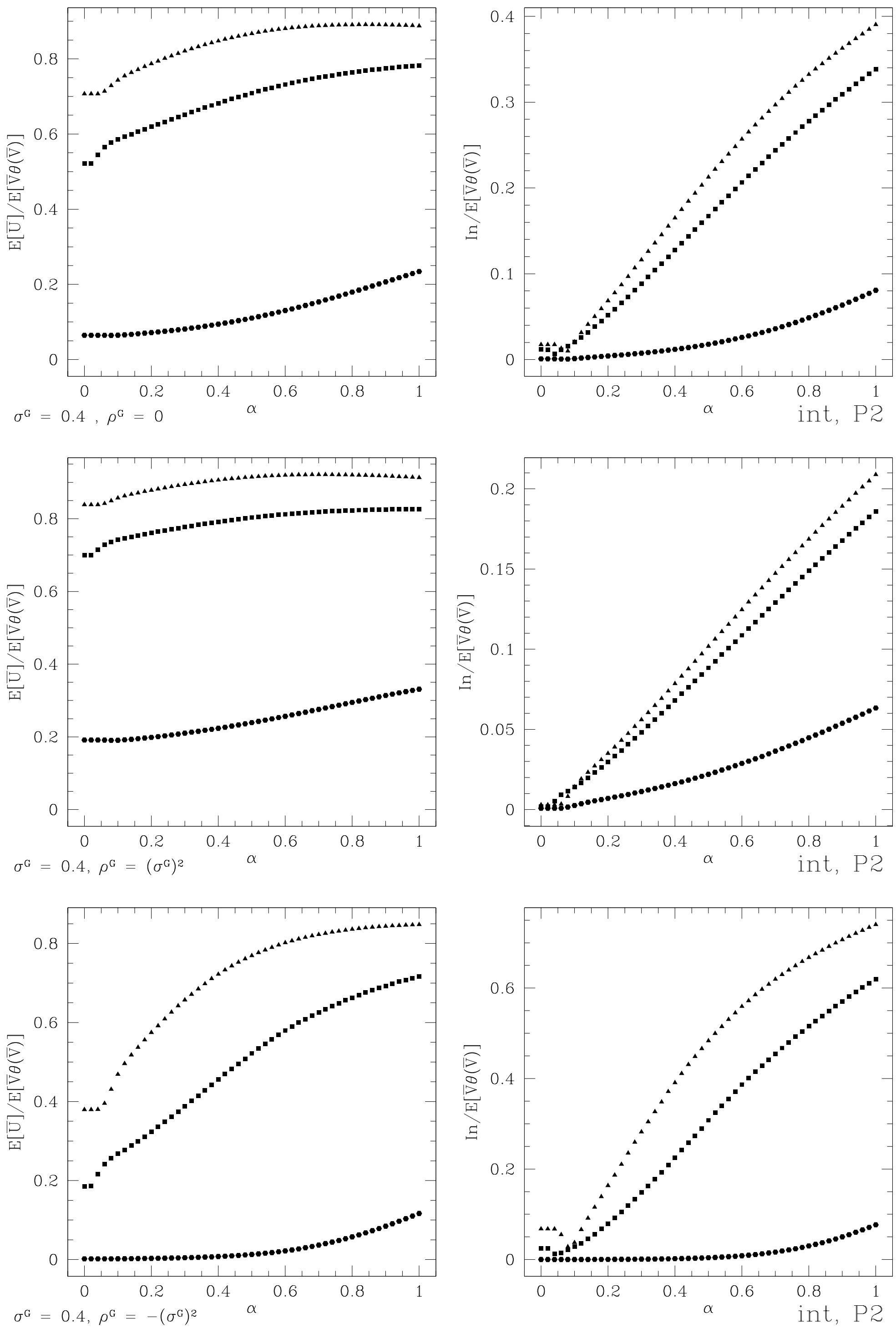

Figure 8: Results for modifications of the interest group model under P2 (western vs. eastern states).

row), the efficiencies are higher than they are under the default interest group model for all rules shown. At the same time, the efficiency curves are flatter for $t = .5, .6$, and the maximum of the $t = .5$ curve has shifted to the left. The maximum efficiency is now found at $(\alpha = .74, t = .5)$. But since this curve is comparably flat, the efficiency is not that sensitive to α any more.

The shape of the In vs. α curves reminds one of the results for the default interest group model. However, there is not perfect equality at $\alpha = 0$ any more at least for $t = .5, .6$. The reason is as follows: For $\alpha = 0$, the group with the eastern states has more weights in the assembly, for which reason people from the eastern states take slightly larger profits than people from the western states do – hence the inequality. For the most part, In is a growing function of α for all thresholds shown. Thresholds around .5 are worst for inequality for most values of α.

If we consider both welfarist standards, the choice of a decision rule is a bit easier than it is under the default interest group model. Particularly, the efficiency under $t = .5$ is not that sensitive to the α-value any more. But one has still to weigh between the different standards.

If positive correlations between utilities for different groups are introduced and if $\varrho^{\mathrm{G}} > 0$ is set at its maximal possible value (second row), the curves for the efficiencies become even higher. Also, for $t = .5., 6$ they are very flat. The local minima of In for $t = .5, .6$ seem still to be at $\alpha \neq 0$, but they are less pronounced. Under this model, $t = .5$ and that α-value at which In has a minimum in the $t = .5$ curve, define a rule that does reasonably well on both standards: The expected utility is comparably high, and inequality is very small.

It is very different for anticorrelations between the groups. In the third row in Fig. 8, we consider the minimal ϱ^{G} compatible with $\sigma^{\mathrm{G}} = .4$. The curves for efficiency cover a broad range of efficiency values. For each t, they mostly increase as a function of α. The overall maximum in our scans is obtained at $(\alpha = 1, t = .5)$. For larger σ^{G} and a minimal ϱ^{G}, even negative expected utilities are possible. Unfortunately, for each t, the measure of inequality is also mostly an increasing function of α with a minimum close to $\alpha = 0$, so minimizing In and maximizing efficiency is not possible at the same time.

In order to scrutinize the transition between the modifications of the interest and the aggregate models, we set $\sigma^{\mathrm{P}} = 1$ and $\sigma^{\mathrm{S}} = .01$. Roughly following the procedure how we set σ^{G} before, we determine σ^{G} such that for Ireland, $\sigma^{\mathrm{G}} \approx .4\sqrt{(\sigma^{\mathrm{S}})^2 + \frac{1}{n_I}(\sigma^{\mathrm{P}})^2}$. As expected, the results that we get are very similar to our results for the interest group model both under the partitions P1 and P2.

7 Conclusions

In this paper, we have considered federations in which representatives vote on proposals that affect the people represented. As an example, we have focused on the Council of Ministers in the European Union, where representatives of the different states vote on proposals drafted by the European Commission. We have evaluated alternative decision rules from a welfarist perspective. The idea is that a decision rule is better than another, if the resulting welfare distribution is better. Alternative decision rules result in different welfare distributions, since they let different proposals pass.

We have considered two welfarist standards, abbreviated as $\mathcal{U}$ and $\mathcal{E}$. $\mathcal{U}$ demands that the expected utility for the whole federation be maximum. $\mathcal{E}$ demands that the inequality in the welfare distribution be minimal. In this paper we have only considered inequalities in the expected utilities that people from different states would have on average.

In order to work out welfarist evaluations of decision rules, one has to model how the proposals affect the welfare of the people concerned. Under certain assumptions (interpersonal comparability of utility scales) proposals can be assigned utility vectors, such that the components capture the benefits or costs for the individual people. A probability distribution over the utility vectors is assumed.

So far, most work in this area (Barberà & Jackson 2006 and Beisbart & Bovens 2007, e.g.) has been assuming that the utilities from proposals are stochastically independent for people from different states. This is not a realistic assumption, though. The aim of this paper was therefore to work out the welfarist approach under the assumption of dependent utilities.

Following Barberà & Jackson (2006), we identified the decision rule that maximizes expected utility for the federation in analytical terms, even if the utilities of people from different states are dependent. However, in general, this decision rule cannot be cast in simple terms. It is thus not suitable for practical purposes, or so we have argued.

We have therefore restricted ourselves to a family of weighted rules and a specific class of probability models. The family of weighted rules can be parameterized using the threshold of acceptance, t, and the degree of degressive proportionality, α. The models in our model class assign each person a utility that has three parts: One is common to all people in a group of states, one is common to all people in a state only, one is perfectly personal. In this way dependencies arise.

Results have been obtained by means of simulations. We have started from the (default) aggregate and the interest group models (Beisbart & Bovens 2007) and introduce dependencies between states. Under the default aggregate model, all utilities are independent, and the rule ($\alpha = .5, t = .5$) (Penrose 50) is very good on both counts. Under the default interest group model, the utilities for people from the same state are perfectly correlated, and the standards pull into different directions: Maximizing expected utility requires $t = .5$ and $\alpha = 1$, whereas equalizing the expected utilities for the different states requires a very low α and favors thresholds far away from .5.

We have then modified the aggregate and the interest group models by assuming interstate utility dependencies. Our dependencies are such that the marginal probability densities are identical for all people in the federation. The idea is thus, that the proposals are not biased towards any person – at least in a stochastic sense. Of course, the results for the modifications of the aggregate and the interest group models depend on the details of the model parameters. Still, a few general things can be said.

First, interstate utility dependencies are important. If they are switched on, the impact that alternative decision rules have on the welfare distribution can change qualitatively. For a clear example compare the default aggregate model (Fig. 1) and our P2 modification with a minimal ϱ^{G} (third row of Fig. 5). Whereas, under the P2 modification, α-values close to zero minimize inequality for the t-values shown, the curves for inequality have a minimum around $\alpha = .5$ on the default aggregate model.

Second, the following questions are important for the results

- Regarding the partition of the states into groups: Is group membership related to population size?
- If the federation is partitioned into two groups of states: Does the group with the higher population have more weights than the threshold requires?
- Are there positive or negative correlations between people from different groups?

Third, regarding dependencies, there is the following trend: If there are enough positive dependencies (i.e. under P0, $\sigma^{\mathrm{G}} > 0$ and under P1 or P2, $\sigma^{\mathrm{G}} > 0$ and ϱ^{G} maximal), the welfare distribution is less sensitive to the choice of a decision rule in the following sense: There is a broader range of rules that are close to optimal in terms of expected utility, particularly for modifications of the aggregate model. At the same time, the expected utilities and the efficiencies (i.e. suitable normalized expected utilities) are higher than under no dependencies. This makes the choice of a decision rules easy and in a way less important. The choice of a decision rule is often more important for negative correlations.

Fourth, one can ask, whether the recommendations are stable that we mentioned for the default aggregate and interest group models above. The results for the aggregate model are fairly stable for the partitions that we called P0 and P1 above. That is, one can always find a decision rule that does well in terms of both standards. The α-value of this rule is sometimes larger than .5, the threshold can be set at .5. However, the results are not stable under partition P2, where $\mathcal{E}$ requires a very low α close to zero.

The recommendations for the interest group model also remain rather stable, which means that the choice of a decision rule is still dilemmatic: $\mathcal{U}$ requires one to chose something like $t = .5$ and a high α, whereas $\mathcal{E}$ demands a threshold far away from .5 and a small α. Exceptions are the P1 modifications with $\varrho^{\mathrm{G}} \leq 0$ that we have considered.

In spite of these results, much work remains to be done within the welfarist approach to decision rules. One task is to find empirically adequate probability models for proposals in real world federations. Maybe, such models can be constrained by data on past voting profiles. Another task are budget constraints. In this paper, we have assumed that the utilities from proposals have zero means. But we do not exclude proposals that put a very high utility on the federation. This may seem unrealistic, since the budget of the federation is constrained. Our models with negatively dependent groups come closest to a federation with a budget constraint. But budget constraints require further examination. Further restrictions of the welfarist approach so far are discussed in Beisbart et al. (2005), Sec. 5. Our hope is that these restrictions can be overcome in future research.

References

Barberà, S. & Jackson, M., *On the Weights of Nations: Assigning Voting Weights in a Heterogeneous Union*, Journal of Political Ecomomy **114** (2006), 317–339.

Beisbart, C. & Bovens, L., *Welfarist Evaluations of Decision Rules for Boards of Representatives*, Social Choice and Welfare **29** (2007), 581–608.

Beisbart, C. & Hartmann, S., *Welfarism and the Assessment of Social Decision Rules*, in: *Proceedings of the 1st International Workshop on Computational Social Choice* (Endriss, U. & Lang, J., eds.), 2006, pp. 35–48.

Beisbart, C., Bovens, L., & Hartmann, S., *A Utilitarian Assessment of Alterantive Decision Rules in the Council of Ministers*, European Union Politics **6** (2005), 395 – 418, Appendix online under http://www.uni-konstanz.de/eup/iss_64.htm.

Bovens, L. & Hartmann, S., *Welfare, Voting and the Constitution of a Federal Assembly*, in: *Reasoning, Rationality and Probability* (Galavotti, M. C., Scazzieri, R., & Suppes, P., eds.), CSLI Publications, Stanford, 2007.

Chamberlain, G. & Rothschild, M., *A Note on the Probability of Casting a Decisive Vote*, Journal of Economic Theory **25** (1981), 152–162.

Coelho, D., *Maximin Choice of Voting Rules for Committees*, Economics of Governance **6** (2005), 159–175.

Crain, W. M., Messenheimer, H. C., & Tollison, R. D., *The Probability of Being President*, The Review of Economics and Statistics **75** (1993), 683 – 689.

Dancy, J., *Ethics without Principles*, Oxford University Press, Oxford, 2004.

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

Felsenthal, D. S. & Machover, M., *Enlargement of the EU and Weighted Voting in its Council of Ministers*, Voting Power Report 01/00, London School of Economics and Political Science, Centre for Philosophy of Natural and Social Science, London; downloadable from http://www.lse.ac.uk/vp, 2000.

Gelman, A., Katz, J. N., & Bafumi, J., *Standard Voting Power Indexes Don't Work: An Empirical Analysis*, British Journal of Political Science **34** (2004), 657 – 74.

Good, I. J. & Mayer, L. S., *Estimating the Efficacy of a Vote*, Behavioral Science **20** (1975), 25–33.

Laruelle, A. & Valenciano, F., *Assessing Success and Decisiveness in Voting Situations*, Social Choice and Welfare **24** (2005), 171 – 197.

Schweizer, U., *Calculus of Consent: A Game-theoretic Perspective*, Journal of Institutional and Theoretical Economics **146** (1990), 28–54.