Efficient Compromising*

by Tilman Börgers[†] and Peter Postl[‡]

December 2004

*We are grateful to Mark Armstrong, Eddie Dekel, Rafael Hortala-Vallve and Peter Norman for comments. Tilman Börgers' research was financially supported by the ESRC through the "Centre for Economic Learning and Social Evolution" (ELSE) at University College London.

[†]Department of Economics, University College London, Gower Street, London WC1E 6BT, United Kingdom, t.borgers@ucl.ac.uk.

[‡]Department of Economics, University of Birmingham, Birmingham B15 2TT, United Kingdom, p.postl@bham.ac.uk.

Abstract

Two agents have to choose one of three alternatives. Their ordinal rankings of these alternatives are commonly known among them. The rankings are diametrically opposed to each other. Ex-ante efficiency requires that they reach a compromise, that is choose the alternative which they both rank second, if and only if the sum of their von Neumann Morgenstern utilities from this alternative exceeds the sum of utilities when either agent's most preferred alternative is chosen. We assume that the von Neumann Morgenstern utilities of the middle ranked alternative are independent and identically distributed, privately observed random variables, and ask whether there are incentive compatible decision rules which elicit utilities and implement efficient decisions. We show that no such decision rules exist if the distribution of agents' types has a density with full support. We also study the problem of finding second-best decision rules in our set-up, and explain how this problem differs from more familiar second-best problems. Finally, we give some numerical insights into the nature of second-best rules. For a variety of distributions of types, second-best rules involve very little inefficiency.

1. Introduction

You and your partner disagree about which restaurant to go to. You prefer the Italian restaurant over the English restaurant, and the English restaurant over the Chinese restaurant. But your partner has exactly the opposite preferences. Should you compromise by going to the English restaurant, or should you go to a restaurant that one of you likes best? The answer to this question presumably depends on how strongly each partner prefers his favorite restaurant over the compromise, and how strongly he prefers the compromise over the bottom ranked alternative. Is there a way of finding out the partners' strengths of preference, or will they, for example, necessarily pretend to have a lower valuation of the compromise than they really have? This is the question which this paper addresses.

Of course, we first need to say what we mean by "strength of preference". One interpretation could be that the strength of preference is equal to the amount of money that an agent is willing to pay in order to obtain one outcome rather than another. If this were what we have in mind, then one could try to elicit the strength of the partners' preferences by introducing a mechanism that obliges any partner whose favorite restaurant is chosen to pay compensation to the other.

Here, we want to abstract from such side payments as they seem inappropriate in many situations. Spouses, for example, rarely pay money to each other to resolve conflicts. When initially conceiving of this paper, we had another situation in mind in which money payments are typically not made: voting. Voting rules might try to elicit, in some sense, the "strength of preference" for candidates, yet voters are typically not asked to offer payments together with their votes. The problem that we study here is a simplified version of the problem of designing voting rules that elicit strengths of preferences without side payments.

If side payments are ruled out, what do we mean by "strength of preferences", and how can we elicit them? We mean in this paper by "strength of preference" the von Neumann Morgenstern utility of alternatives. If we evaluate different mechanisms from an *ex ante* perspective (Holmström and Myerson (1983)), then von Neumann Morgenstern utilities have to be taken into account when resolving conflicts. How can we elicit von Neumann Morgenstern utilities truthfully? By exposing agents to risk. Agents' choices among lotteries indicate their von Neumann Morgenstern utilities. If agents play a game with incomplete information, then they are almost always automatically exposed to risk. Their choices can then reveal their utilities.

We develop this theme in a simple stylized example with two agents and three alternatives. We assume that it is commonly known that the agents' rankings of the alternatives are diametrically opposed. Their von Neumann Morgenstern utilities for the alternatives are, however, not known. To implement efficient decisions, these utilities need to be elicited, as it is optimal to implement the compromise if and only if the sum of the agents' utilities of the compromise is larger than the sum of their utilities when either agent's most preferred alternative is chosen.

Our first main result is that this decision rule, to which we refer as first-best, is not incentive compatible, and can therefore not be implemented, if the distribution of von Neumann Morgenstern utilities has a density with full support. We complement this observation with a study of second-best decision rules. We explain that the structure of the second-best problem in our context is different from that in other, more familiar settings, and that an analytical approach appears difficult. We then report numerical results about second-best decision rules. These indicate that the amount of inefficiency that second-best involves is surprisingly small.

One motivation for our paper is that mechanisms for efficient compromising are potentially relevant to many areas of conflict, such as labor relations or international negotiations. A second motivation was already mentioned above: we are interested in the application of the theory of Bayesian mechanism design to voting. The current study is a first and limited step into that direction. Traditionally, the literature on voting has either studied strategic behavior under specific voting rules, or the design of voting rules using solution concepts that rely on weak informational assumptions, such as dominant strategies (Gibbard (1973), Satterthwaite (1975), Dutta, Peters and Sen (2004)), or undominated strategies (Börgers (1991)). Our purpose here is to explore the theory of voting with stronger informational assumptions, which are, however, frequently made in other areas of incentive theory. A third motivation for this paper is that it is a case study in Bayesian mechanism design without transferrable utility. Much of the literature on Bayesian mechanism design has relied on the assumption of transferrable utility. It seems worthwhile to explore what happens if this assumption is relaxed.

It turns out that the setting that we study, although formally without transferrable utility, is closely related to models of mechanism design for public goods with transferrable utility as studied by d'Aspremont and Gérard-Varet (1979), Güth and Hellwig (1986), Rob (1989), and Mailath and Postlewaite (1990). These papers all consider settings in which there are two goods, a public good, and "money." Agents' preferences are assumed to be additive in the quantity of the public good that is provided and "money." In our setting there is no "money." However, for each agent the probability with which their most preferred alternative is chosen serves in some sense as “money.” The public good is the probability with which the compromise is implemented. Agents “pay” for an increased probability of the compromise by giving up probability of their most preferred alternative. Agents’ preferences are additive in the “real good” and “money” because they are von Neumann Morgenstern preferences over lotteries, which are additive in probabilities.

The details of the analogy between our work and the literature on mechanism design for public goods will be explained later. Two points deserve emphasis. Firstly, an important difference between our work and the established public goods literature is that agents, in our model, face a budget constraint, which is absent from traditional models. The budget constraint arises from boundaries on the amount of probability which agents can surrender: for instance, it cannot be larger than one.

The second difference is that our model does not feature individual rationality constraints. Most, though not all, of the previous literature on public goods has postulated an individual rationality constraint (see the discussion in Hellwig (2003)). Although in our setting there is no “outside option” which would guarantee agents a minimum utility, a lower boundary for agents’ expected utility nevertheless easily follows from the facts that there is only a finite number of allocation decisions, and that there is an upper boundary for the “payments” which agents can make. Thus the budget constraint has a similar effect as an individual rationality constraint.

In the light of the above discussions, it becomes intuitively plausible that it is not possible to implement the first best in our setting. Analogous results have been obtained for the public goods setting by Güth and Hellwig (1986), Rob (1989), and Mailath and Postlewaite (1990). The analysis of the second best in our setting is more involved than in the established public-goods literature because of the difficulty involved in taking account of the implicit budget constraint. For this reason our analysis of second-best rules is restricted to numerical results.

This paper is organized as follows. In Section 2 we introduce our model. Section 3 explains the analogy between our setting and the public goods problem. In Section 4 we characterize incentive compatible decision rules. Section 5 proves the impossibility of implementing first-best decision rules. Section 6 explains the difficulties with an analytical approach to second best. In Section 7 we report numerical results about second-best rules. Section 8 concludes.

2. The Model

There are two agents $i = 1, 2$ who must collectively choose one alternative from the set $\{A, B, C\}$. Agent 1 prefers A over B, and B over C. Agent 2 prefers C over B, and B over A. These preferences are common knowledge among the two agents.

Each agent i has a von Neumann Morgenstern utility function $u_i : \{A, B, C\} \to \mathbb{R}$. We normalize utilities so that $u_1(A) = u_2(C) = 1$ and $u_1(C) = u_2(A) = 0$. These features of the von Neumann Morgenstern utility functions are common knowledge among the two agents.[1]

For $i = 1, 2$ we write t_i for $u_i(B)$. We refer to t_i as *player i's type*. We assume that t_i is a random variable which is only observed by agent i. The two players' types are stochastically independent, and they are identically distributed with cumulative distribution function G. We assume that G has

[1] The normalization of agents' utilities that we have introduced in this paragraph will be discussed further towards the end of this section.

support $[0,1]$, that it has a continuous derivative g, and that $g(t) > 0$ for all $t \in (0,1)$. The joint distribution of (t_1, t_2) is common knowledge among the agents.

Definition 1 *A* decision rule f *is a function* $f : [0,1]^2 \rightarrow \Delta(\{A, B, C\})$ *where* $\Delta(\{A, B, C\})$ *is the set of all probability distributions over* $\{A, B, C\}$*.*

We write $f_A(t_1, t_2)$ for the probability which $f(t_1, t_2)$ assigns to alternative A, and we define $f_B(t_1, t_2)$ and $f_C(t_1, t_2)$ analogously.

Given any decision rule, denote for every $t_i \in [0,1]$ by $p_i(t_i)$ the probability that agent i's most preferred alternative is implemented, conditional on agent i's type being t_i, i.e.:

$$p_1(t_1) = \int_0^1 f_A(t_1, t_2) g(t_2) dt_2 \quad \text{and} \quad p_2(t_2) = \int_0^1 f_C(t_1, t_2) g(t_1) dt_1.$$

Denote by $q_i(t_i)$ the probability that the compromise is implemented, conditional on agent i's type being t_i, i.e. for $i = 1, 2$:

$$q_i(t_i) = \int_0^1 f_B(t_1, t_2) g(t_j) dt_j \quad \text{where} \quad j \neq i.$$

Finally, we denote by $U_i(t_i)$ agent i's expected utility, conditional on being type t_i, that is:

$$U_i(t_i) = p_i(t_i) + q_i(t_i) t_i.$$

We restrict attention to decision rules for which the integrals $p_i(t_i)$ and $q_i(t_i)$ exist for every $i = 1, 2$ and every $t_i \in [0,1]$.

We evaluate decision rules using a utilitarian welfare criterion.[2] Welfare is defined as the ex-ante expected utility of an agent who does not know whether he will be agent 1 or 2, and who does not know his type. We assume that the probability of being either agent 1 or agent 2 is equal to 1/2, and

[2]We comment on the welfare criterion in the last paragraph of this section.

that the prior probability of types that is used for welfare calculations is again described by G. We can then omit the probability weights for agents when calculating ex-ante expected utility and simply consider a non-weighted sum, as in the following definition.

Definition 2 *The* ex-ante expected utility *associated with decision rule* f *is:*

$$\int_0^1 U_1(t_1)g(t_1)dt_1 + \int_0^1 U_2(t_2)g(t_2)dt_2.$$

The expression in this definition can equivalently be written as:

$$1 + \int_0^1 \int_0^1 f_B(t_1,t_2)(t_1+t_2-1)g(t_1)g(t_2)dt_1dt_2.$$

In this sum we might as well omit the initial constant 1, which is what we shall do in future.

From the last formula it is obvious that the decision rules f that maximize ex-ante expected utility among all decision rules are those that are *first-best* in the sense of the following definition.

Definition 3 *A decision rule* f *is called* first-best *if with probability 1 we have:*

$$t_1 + t_2 > 1 \Rightarrow f_B(t_1,t_2) = 1 \text{ and}$$

$$t_1 + t_2 < 1 \Rightarrow f_B(t_1,t_2) = 0.$$

Note that there are many first-best decision rules. The reason is firstly that Definition 3 requires the listed conditions to be true *with probability 1*, but not *always*. The reason is secondly, and more importantly, that the above definition does not restrict the probabilities with which alternatives A and C are chosen if the compromise is *not* implemented.

In the next section we shall introduce incentive compatibility. Our interest will be in those decision rules that maximize ex-ante expected utility among all incentive compatible rules. In the following definition we use the term *incentive compatible* in the sense that will be defined in the next section.

Definition 4 *A decision rule f is called* second-best *if it yields the largest ex-ante expected utility among all incentive compatible decision rules.*

We emphasize two aspects of our framework. The first is that the framework is ex-ante symmetric with respect to agents. This is intended to reflect that ex-ante there is no known reason to systematically bias the decision rule in favor of one of the agents. This seems the most interesting scenario when designing rules that are meant to be used in a large variety of circumstances.

The second aspect of our model that we emphasize is the normalization of von Neumann Morgenstern utilities so that the utility of the most preferred alternative is 1, and the utility of the least preferred alternative is 0. Because we construct an ex-ante symmetric model, we apply this normalization to both agents. An implication of this is that welfare as defined above, i.e. the sum of agents' utilities, is unchanged if probability is shifted from alternative A to alternative C. The only welfare relevant decision is whether the compromise B is implemented.

While our normalization allows us to focus on the incentive issues involved in efficient compromising, it is clearly restrictive.[3] Moreover, one of our main results, i.e. that no first-best rule is incentive compatible, is non-trivial only

[3] Our model can be generalized somewhat without change to our main conclusions. We can allow both agents' vectors of von Neumann Morgenstern utilities to be multiplied by a constant τ_i, provided that the distribution of both τ_is is independent of t_1 and t_2, and that τ_1 and τ_2 are independent. A first-best rule would then condition on the τ_is, but no incentive compatible rule can do so because the values of the τ_is do not affect agents'

because we have chosen a set-up in which it is welfare irrelevant whether A or C is chosen. If the choice of A and C were welfare-relevant, say in a model in which agents' ex-ante utilities of A, B and C were continuously distributed random variables with some generic joint distribution, there would typically be only a single first-best decision rule, and it would be straightforward to check whether this rule is incentive-compatible. In our set-up, by contrast, there are many first-best rules, and it will require some work to find out whether any of them is incentive-compatible.

We argue, however, that our framework is no more restrictive than what is commonly assumed in mechanism design. Consider the allocation of a single indivisible good to a number of possible buyers with privately observed preferences. A commonly made assumption is that buyers' preferences are affine in the quantity of the good (either zero or one) and money, and that the coefficient applied to money is 1 for all types of all agents. This assumption implies that for ex-ante utilitarian welfare the allocation of money is irrelevant, and only the allocation of the good matters.

The assumption that the coefficient applied to money payments is 1 for all types of all agents is analogous to our normalization of all types' utilities for the alternatives A and B. If this assumption were not made, and if hence the allocation of money were welfare relevant, then there would typically only be a single first-best allocation rule, and it would be entirely straightforward to check whether this rule is incentive compatible.[4]

---

interim incentives. (For a more thorough discussion of the argument of the last sentence, though in a different setting, see Hortala-Vallve (2004, Theorem 3).) The analysis of this paper can be interpreted as second-best analysis that incorporates in the set-up the fact that it is impossible to condition on the τ_is.

[4]Indeed, in most cases, the first-best would *not* be incentive compatible. Thus, one of the standard results of single-unit auction theory would be overturned

We finally note that the utilitarian welfare criterion that we use here is implicit in the mechanism design literature. In the single unit auction context, for example, “allocating the good to the buyer who values it most” is necessary and sufficient for optimality from an ex-ante utilitarian stand point, but it is certainly not necessary for efficiency from an ex-post viewpoint. Giving the good and all the money always to the same bidder is another example of an ex-post efficient rule.[5]

3. Analogy with the Public Goods Problem

There is a close analogy between our model and models typically considered in the theory of Bayesian mechanism design for non-excludable public goods (d'Aspremont and Gérard-Varet (1979), Güth and Hellwig (1986), Rob (1989), Mailath and Postlewaite (1990)). We can view the probability with which the compromise is chosen in our framework as the quantity of a public good without exclusion that is consumed by both agents. Each agent's private type determines the agent's valuation of the public good. Agents pay for the public good with a reduced probability of their most preferred alternative.

To make this analogy more precise let us define somewhat arbitrarily the outcome in which each of the two extreme alternatives A and C is chosen with probability 0.5 as the *default* outcome. For every agent i define $m_i(t_1, t_2)$ to be the difference between the default probability of this agent's most preferred alternative, and the probability with which the agent's most preferred alternative is chosen by a given decision rule if the types are (t_1, t_2):

$$\begin{aligned} m_1(t_1, t_2) &\equiv 0.5 - f_A(t_1, t_2) \\ m_2(t_1, t_2) &\equiv 0.5 - f_C(t_1, t_2) \end{aligned}$$

[5] This allocation is also interim efficient and it is interim incentive compatible.

for all $(t_1, t_2) \in [0, 1]^2$. We can think of $m_i(t_1, t_2)$ as the *payment* made by agent i if types are (t_1, t_2). The probability of the compromise is then:

$$f_B(t_1, t_2) = m_1(t_1, t_2) + m_2(t_1, t_2)$$

for all $(t_1, t_2) \in [0, 1]^2$. We can think of this probability as the quantity of a public good that is produced if types are (t_1, t_2). The above equation shows that the public good is produced with a one-to-one technology where the quantity produced equals the sum of agents' payments. The quantity of the public good can obviously not be more than one, and we might model this by assuming that the public good's marginal costs rise to infinity once the quantity exceeds one.

Our model is then isomorphic to the traditional set-up for Bayesian mechanism design for non-excludable public goods, except that we have to respect a budget constraint: For every $i \in \{1, 2\}$ and every $(t_1, t_2) \in [0, 1]^2$ we must have:

$$m_i(t_1, t_2) \in [-0.5, +0.5].$$

Otherwise $f_A(t_1, t_2)$ or $f_C(t_1, t_2)$ would be larger than one or smaller than zero. This implicit ex-post budget constraint of individual agents is a first feature that distinguishes, to our knowledge, our set-up from all public-good models that have been investigated in the literature.

A second feature that distinguishes our set-up from the traditional public goods set-up is that there is no individual rationality constraint in our model. In the public goods context, and in other related contexts, one is often interested in characterizing all decision rules that are incentive compatible and individually rational.[6] But in our model there is no natural role for individual rationality.

[6] An exception is d'Aspremont and Gérard-Varet (1979).

The two differences between our context and the traditional set-up neutralize each other to some extent. Specifically, even though there is no individual rationality constraint, there is a lower boundary for the interim expected utility of the agents because there is only a finite number of alternatives, and agents cannot be asked to pay more than their budget allows. We shall return to this point in Section 5.

4. Incentive Compatibility

Because types are privately observed, a decision rule can be implemented in practice if and only if it is incentive compatible.[7]

Definition 5 *A decision rules f is* incentive compatible *if for $i = 1, 2$ and for any types $t_i, t_i' \in [0, 1]$:*

$$p_i(t_i) + q_i(t_i)t_i \geq p_i(t_i') + q_i(t_i')t_i.$$

The proof of the following Lemma is familiar from the literature on Bayesian incentive compatibility. Therefore, we omit it.

Lemma 1 *A decision rule f is incentive compatible if and only if for $i = 1, 2$ we have:*

(i) q_i is monotonically increasing in t_i;

(ii) for any two types $t_i, t_i' \in [0, 1]$ with $t_i < t_i'$:

$$-t_i'(q_i(t_i') - q_i(t_i)) \leq p_i(t_i') - p_i(t_i) \leq -t_i(q_i(t_i') - q_i(t_i)).$$

[7] The following definition implicitly assumes that the mechanism which is used to implement the decision rule is a direct one. By the *revelation principle* this is without loss of generality.

The first item in this Lemma says that the probability of the compromise, conditional on an agent's type, increases as this agent's utility of the compromise increases. Where is this probability taken from? The second item in Lemma 1 indicates that some of the probability has to be taken from the probability assigned to the agent's most preferred alternative. It is intuitive that the probability of the most preferred alternative must decrease. If the additional probability for the compromise were only taken from the agent's least preferred alternative, then the agent would have an incentive to report a higher utility for the compromise than he actually has. The agent has to *pay* for a higher probability of the compromise with a lower probability of his most preferred alternative.

The inequality in the second item in Lemma 1 provides a lower and an upper boundary for the change in the probability of the most preferred alternative. Both of these boundaries are negative. The boundaries are such that among two types the higher type prefers to pay the price and obtain a higher probability of the compromise, whereas the lower type prefers not to pay the price.

The next lemma describes incentive compatibility in terms of properties of the interim expected utility. The result is standard in related settings[8], and therefore we again omit the proof.

Lemma 2 *A decision rule f is incentive compatible if and only if for every agent $i = 1, 2$:*

(i) q_i is monotonically increasing in t_i;

(ii) for every $t_i \in [0, 1]$ such that q_i is continuous at t_i:

$$U_i'(t_i) = q_i(t_i).$$

[8] See, for example, Section 5.1.1 of Krishna (2002).

We can use the differential equation of Lemma 2 to obtain a formula that links the interim expected probabilities of each agent's favorite alternative to the interim expected probabilities of the compromise. This is done in Lemma 3. To solve the differential equation, we have to take the value at some boundary point as given. We choose here the highest type, i.e. $t_i = 1$, rather than, as is convention, the lowest type, $t_i = 0$, because in the proof of Proposition 1 the current formulation will be more useful. Apart from this modification, the proof is again standard, and therefore omitted.

Lemma 3 *A decision rule f is incentive compatible if and only if for every agent $i = 1, 2$:*

(i) q_i is monotonically increasing in t_i;

(ii) $p_i(t_i) = p_i(1) + q_i(1) - q_i(t_i)t_i - \int_{t_i}^{1} q_i(s_i)ds_i$ for all $t_i \in [0, 1]$.

5. Impossibility of Implementing First-Best Rules

We can now build on the characterization of incentive compatible rules, and show that:

Proposition 1 *No first-best decision rule is incentive compatible.*

We shall prove Proposition 1 by showing that a first-best decision rule that is incentive compatible would have to have the property that the ex-ante probability of the compromise, and the ex-ante probabilities of alternatives A and C, as implied by incentive compatibility, add up to more than one. This then contradicts the definition of decision rules.

If our set-up is interpreted as a public goods set-up, as indicated in Section 3, our result shows that the contributions which individuals are willing to make under incentive compatibility are not enough, from an ex-ante point of view, to cover the total resources required to produce the first-best quantity of the public good. This is also the reason why the first best cannot be implemented in standard models of incentives in public goods provision (for example: Güth and Hellwig (1986)). However, as argued above, our set-up differs from the most common set-up in that we have no individual rationality constraint. If there is no individual rationality constraint in the public goods framework, then the first-best can be implemented (d'Aspremont and Gérard-Varet (1979)). We obtain a different result because, as explained in Section 3, our agents face individual budget constraints. These budget constraints imply lower boundaries for the utility of each type, even if no individual rationality is required.

Despite of the differences between our model and the public goods model, the proof of Proposition 1 that we provide below parallels the modern approach to proving impossibility results in the field of mechanism design. For example, it is analogous to Milgrom's (2004, p.79) version of the proof of the Myerson-Satterthwaite (1983) impossibility theorem. We begin the proof by arguing that Lemma 3 implies that all incentive compatible first-best decision rules imply the same ex-ante probabilities for the three alternatives. We then construct one particular incentive compatible first-best decision rule for our problem, namely a Vickrey-Clarke-Groves (VCG) mechanism. We show for this decision rule that the ex-ante probabilities of the three alternatives add up to more than one. It then follows that the same has to be true for *all* incentive compatible decision rules.

An important difference between the structure of our proof and similar proofs of earlier impossibility results in Bayesian mechanism design is that in earlier proofs individual rationality is used to select the mechanism on which to focus among all conceivable VCG-mechanisms. In our proof, the VCG-mechanism on which we focus is determined by the condition that the highest type, $t_i = 1$, has to expect the compromise with probability 1, and all other alternatives with probability zero. Thus, we use efficiency, and this agent's "budget constraint" to select the appropriate VCG-mechanism.

Proof: The proof is indirect. Suppose there were a first-best decision rule that is incentive compatible. Then $q_i(t_i) = 1 - G(1 - t_i)$ for $i \in \{1, 2\}$ and almost all $t_i \in [0, 1]$. We want to use Lemma 3 and infer the functions p_i. For this we need to know $p_i(1) + q_i(1)$. Because $q_i(t_i) = 1 - G(1 - t_i)$ holds only for *almost all* $t_i \in [0, 1]$, we cannot assume that it holds for $t_i = 1$. However, interim expected utility U_i is continuous because, by Lemma 3, it is an integral. For almost all types interim expected utility is at least $q_i(t_1)t_i = (1 - G(1 - t_i))t_i$. By continuity, therefore, the expected utility of type $t_i = 1$ has to be one.

We can now apply Lemma 3. Because sets of measure zero don't affect the value of the integral, we can deduce $p_i(t_i) = 1-(1-G(t_i))t_i-\int_{t_i}^{1}(1-G(s_i))ds_i$ for all $t_i \in [0, 1]$. This implies that the value of $\int_0^1 p_i(t_i)g(t_i)dt_i$ is the same for all first-best, incentive compatible decision rules.

The idea of the proof is now to show that the interim probabilities implied by first best and incentive compatibility add up to more than one. We show this by considering the following decision rule, where we ignore for the moment that the components of this rule do not add up to one for every type vector. The function f_B is the first best rule of Definition 3. The functions f_A and f_C are defined as follows.

$$f_A(t_1, t_2) = (1 - f_B(t_1, t_2))(1 - t_2) \quad \text{for all } (t_1, t_2) \in [0, 1]^2;$$

$$f_C(t_1, t_2) = (1 - f_B(t_1, t_2))(1 - t_1) \quad \text{for all } (t_1, t_2) \in [0, 1]^2.$$

We assume that players evaluate outcomes under this rule by the expected utility calculation indicated in Section 2, ignoring the fact that the components of the decision rule do not always add up to one.

This rule is incentive compatible. This follows from the fact that it is a weakly dominant strategy for each player to report his true type. To see that truth telling is weakly dominant, consider, say, player 1, and assume player 1's true type is t_1. Suppose player 2's reported type is t_2. Assume first that t_2 is such that $t_1 + t_2 > 1$. If player 1 reports his true type, he receives utility t_1. If he reports a type τ_1 such that $\tau_1 + t_2 < 1$, then player 1's utility becomes under the above rule: $1 - t_2$. Player 1 will prefer to report his true type because $t_1 > 1 - t_2 \Leftrightarrow t_1 + t_2 > 1$, by assumption. Now suppose alternatively that player 2's reported type is some t_2 such that $t_1 + t_2 \leq 1$. Then, if player 1 reports his true type, he gets: $1 - t_2$. If, alternatively, he pretends to have a type τ_1 such that $\tau_1 + t_2 > 1$, then he receives utility t_1. Player 1 prefers to report his true type because $1 - t_2 \geq t_1 \Leftrightarrow t_1 + t_2 \leq 1$, by assumption.

The interim expected values of f_A and f_C implied by the above decision rule have to satisfy condition (ii) of Lemma 3. This is because the fact that the values of f_A, f_B and f_C add up to one plays no role in the proof of Lemma 3. Therefore, the values of $q_i(t_i)$ for $i \in \{1, 2\}$ and $t_i \in [0, 1]$ that are implied by the above decision rule must be the same as for any first-best, incentive compatible decision rule.

We complete the proof by showing that for the above decision rule the sum of the expected values of $q_i(t_i)$ (for arbitrary but fixed $i \in \{1, 2\}$), $p_1(t_1)$ and $p_2(t_2)$ is greater than one. This sum is equal to the expected value of the sum $f_A(t_1,t_2)+f_B(t_1,t_2)+f_C(t_1,t_2)$. Calculating this sum yields:

$$f_A(t_1,t_2)+f_B(t_1,t_2)+f_C(t_1,t_2)=\begin{cases}1 & \text{if} \quad t_1+t_2\geq 1\\ \\ 2-t_1-t_2 & \text{if} \quad t_1+t_2<1.\end{cases}$$

Because the bottom line is strictly larger than one, and because we have assumed that G has support $[0,1]$ it is obvious that the ex-ante expected value of $f_A(t_1,t_2)+f_B(t_1,t_2)+f_C(t_1,t_2)$ is greater than one.

Q.E.D.

6. Second-Best Rules: Problems with an Analytical Characterization

We now turn to second-best rules. A popular approach to characterizing second-best mechanisms proceeds by writing the maximization problem that defines second-best rules so that only directly welfare-relevant variables appear as choice variables. In our model the directly welfare-relevant variables are the probabilities of the compromise, $f_B(t_1,t_2)$. Thus, we might seek to eliminate from the problem the variables $f_A(t_1,t_2)$ and $f_C(t_1,t_2)$ which are needed to maintain incentives, but do not directly enter the welfare function. To do so, we need a characterization of all functions f_B that can be part of an incentive compatible decision rule. In a supplement to this paper we prove the following lemma, that goes some way towards such a characterization.[9]

[9] The supplement is available from: `http://www.ucl.ac.uk/∼ uctpa01/Papers.htm`. We prove Lemma 4 under the additional assumption that $p_i(1)=0$ for all $i=1,2$. We show that this assumption is without loss of generality in the sense that every incentive compatible decision rule that does not satisfy the assumption can be replaced by one that does, that is also incentive compatible, and that yields the same ex-ante welfare.

Lemma 4 *Consider a function* $\hat{f}_B : [0,1]^2 \to [0,1]$. *For* $i = 1, 2$ *define the interim expected value of* $\hat{f}_B$ *to be:* $\hat{q}_i(t_i) \equiv \int_0^1 \hat{f}_B(t_i, t_j) g(t_j) dt_j$ *for all* $t_i \in [0,1]$, *where* $j \neq i$. *If there is an incentive compatible decision rule* $f = (f_A, f_B, f_C)$ *such that* $f_B = \hat{f}_B$, *then for* $i = 1, 2$:

(i) $\hat{q}_i(t_i)$ *is monotonically increasing in* t_i;

(ii)

$$\int_0^1 \int_0^1 \hat{f}_B(t_1, t_2) \left(t_1 + \frac{G(t_1)}{g(t_1)} + t_2 + \frac{G(t_2)}{g(t_2)} - 1 \right) g(t_1) g(t_2) dt_1 dt_2$$

$$= \quad \hat{q}_1(1) + \hat{q}_2(1) - 1$$

The first condition of this lemma is the same as the first condition in Lemmas 1-3. A few lines of calculations show that the second condition is equivalent to the requirement that the ex-ante probabilities of the agents' preferred alternatives, as implied by the incentive compatibility condition (ii) in Lemma 3, add up to one minus the probability of the compromise. If we adopt the public goods interpretation of our model, we can think of this condition as an ex-ante balanced budget constraint: ex-ante, in expected terms, the contributions to the public good have to cover the costs of producing the public good.

In the theory of public goods conditions that are analogous to conditions (i) and (ii) in Lemma 4 are not only necessary, but also sufficient for an allocation rule to be part of an incentive compatible scheme (for example: Theorem 1 in Mailath and Postlewaite, 1990). The reason for this is that whenever ex-ante budget balance is satisfied by an incentive compatible decision rule, one can construct a payment scheme that is ex-post budget balanced and incentive compatible, and that supports the same allocation rule. This argument does not apply in our setting. If we mimic the standard construction of ex-post budget balanced rules (as described, for example, in the proof of Lemma 3 in Cramton, Gibbons and Klemperer (1987)), then we violate the budget constraints of individuals in our model. That is, under the standard construction individuals would be asked to give up so much probability of their preferred alternative that this probability would become negative. Thus, although ex-ante budget balance is necessary, it is not sufficient for a rule f_B to be part of an incentive compatible decision rule in our setting.

There are at least three directions our investigation could take at this point. Firstly, we could seek to introduce further conditions on f_B so that ex-post budget balance can be achieved. For a simpler setting than ours, Border (1991) has found such conditions. However, generalizing his results to our context seems hard.

A second approach would be to ignore initially that condition (ii) is only necessary, but not sufficient, and to solve the second-best optimization problem using only conditions (i) and (ii) from Lemma 4. We would then have to verify that the solution can be implemented using functions f_A and f_C such that f_A, f_B and f_C add up to one for every type pair. Quite apart from the problems involved in this latter step, we note that constraint (ii) would be awkward to handle as there is a choice variable on the right hand side.

A third approach would be to abandon the idea of focusing on directly welfare-relevant variables, and to optimize simultaneously over f_A, f_B and f_C. This problem would involve a continuum of constraints. In a context in which we don't want to rule out that f_B is a step function, and hence don't want to assume that f_B is continuous, it seems hard to obtain analytical insights into the solution to such a problem.

Here, we shall pursue none of the three possible avenues listed above. As numerical methods for analyzing second-best problems are readily available, it seems natural to investigate properties of second-best rules with the computer.

7. Second-Best Rules: Numerical Results

We now present numerical results about second-best decision rules. To make numerical calculations possible, we discretize the type space. Instead of having types continuously distributed in the interval $[0, 1]$, we assume that both players' types are drawn from the set $\{0, 1/20, 2/20, 3/20, \ldots, 19/20, 1\}$.[10] We maintain the assumption that the model is symmetric, and hence that both players' types follow the same distribution on this set.

For finite type spaces the problem of finding a second-best decision rule is a linear programming problem. The choice variables are the probabilities of the three alternatives for each possible pair of types. The objective function as well as the constraints are linear in these probabilities. We have used two software packages to solve this problem.[11]

[10] We have chosen the discretization as fine as was possible with the computing facilities available to us.

[11] For the computations reported in this section we used version 2.0.1 of the statistical programming language R (R Development Core Team, 2004). We also used version 1.1.5 of the R package "lpSolve" (Berkelaar and others, 2004). R and the R package lpSolve can be obtained from `http://www.r-project.org/`. The package lpSolve is due to Sam Buttrey. It is an interface to version 5 of the linear program solver Lp_solve. This is open source software that is documented at: `http://groups.yahoo.com/group/lp_solve/`. Lp_solve uses the revised simplex algorithm. We have also checked all calculations using the implementation of the interior point algorithm that is available in Mathemat-

Two features of our numerical work should be pointed out. Firstly, we only consider symmetric decision rules, that is, rules such that $f_A(t,t') = f_C(t',t)$ for all (t,t'). For any decision rule that is not symmetric there is a symmetric rule that yields the same ex-ante welfare. This symmetric rule can be constructed by an initial random draw that assigns the names "agent 1" and "agent 2" with equal probability to each of the two agents.

A second feature of our numerical work is that it only considers local incentive constraints, i.e. we only consider the constraint that no type has any incentive to imitate either of their two neighboring types. It is easily seen that in our model, as in other contexts, local incentive constraints imply global incentive constraints.

We begin with the case that all types have the same probability. Figure 1 indicates for this example the probability of the compromise B under the second-best decision rule. The figure shows a grid representing the 21^2 possible pairs of types. For each grid point we have drawn a square that is centered on the grid point. The color of this square reflects the probability with which B is chosen by the second-best rule. If the square is white, the probability of B is 0. If the square is black, then the probability of B is 1. If the color is grey, the probability is between 0 and 1. The darker the grey the larger the probability of B.

We also indicate in Figure 1 the diagonal of the unit square which connects the points $(0,1)$ and $(1,0)$. All first-best decision rules assign probability 1 to the compromise B if the types are above this diagonal, and probability 0 if the types are below this diagonal. Probabilities on the diagonal have no

---

ica 5.0.1.0. The R code that we used to generate the five figures in this section, and the Mathematica code that we have used to verify our calculations, are available from: `http://www.ucl.ac.uk/~ uctpa01/Papers.htm`.

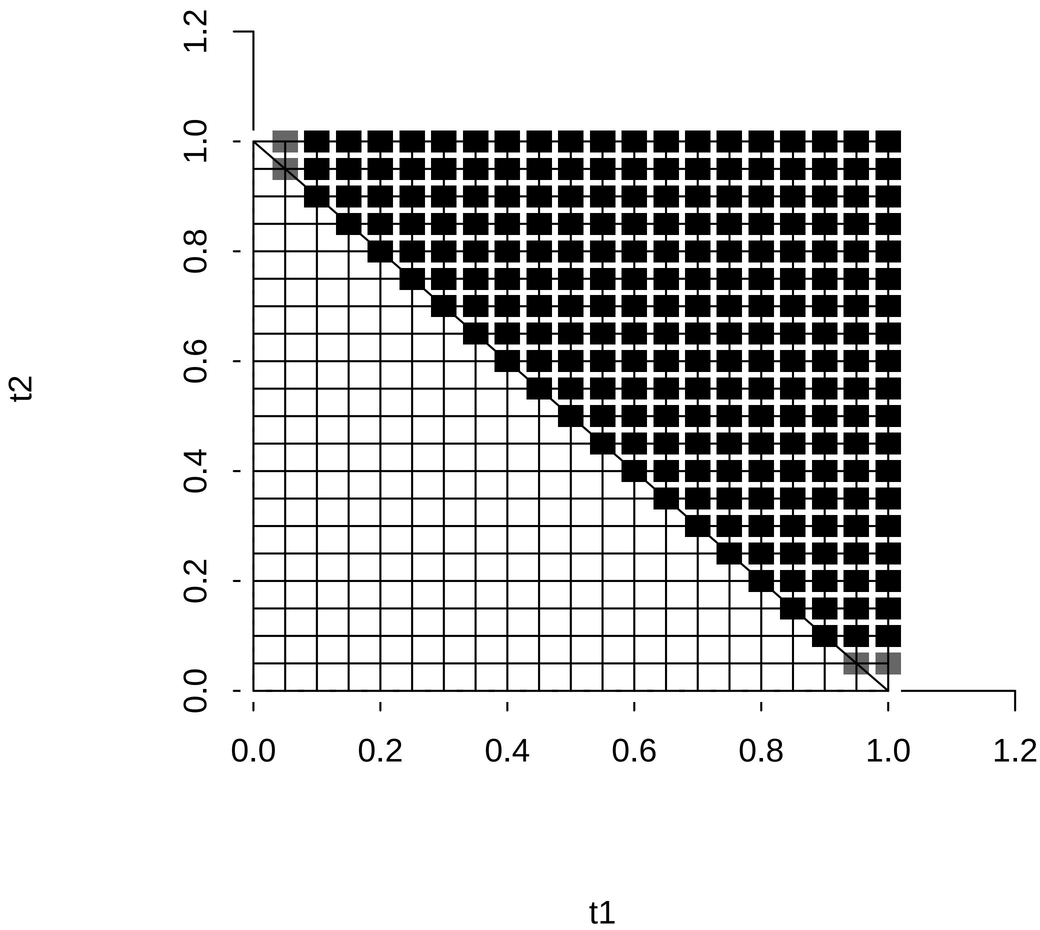

Figure 1: Probability of the compromise under the second-best rule for uniform type distribution

efficiency implications.

The main point to notice about Figure 1 is that the second-best rule is remarkably similar to the first-best rules. The second-best differs from the first-best only for the type pairs $(1, 1/20)$ and $(1/20, 1)$. First-best chooses the compromise with probability 1 for these type. In the second-best rule, the compromise is implemented with probability 0.5767.

We note that deviations from first-best occur only if the sum of the two types is relatively close to one. Moreover, such deviations occur only if the differences between the types are relatively extreme. Finally, we note that the ex-ante probability with which the second-best rule differs from first-best is very small. It is $2 \cdot (1/21) \cdot (1/21) \cdot (1 - 0.5767)$ which is less than 1/500.

The efficiency of the rule displayed in Figure 1 is surprising if one compares it to the efficiency of the second-best rule in the corresponding public goods problem. As a standard of comparison consider the public goods model with two agents where the public good can be produced in any quantity between zero and one, marginal costs equal 1, and valuations are uniformly distributed on $[0, 1]$. Assume also that there is an individual rationality constraint that requires interim expected utilities to be non-negative. Then the public good is implemented under the second-best decision rule if and only if the sum of the valuations is above 1.226, which implies that inefficiency occurs with approximate probability 0.2.[12]

This difference between our model and the public goods model is due to the absence of an individual rationality constraint in our model. To illustrate

[12] This is the result of an analytical calculation with a continuous type space. We have verified that the numerical solution of a discretized version of the public goods model yields almost the same conclusion. For this verification we have used the same discretization and similar R code as we used for the analysis of the compromise problem.

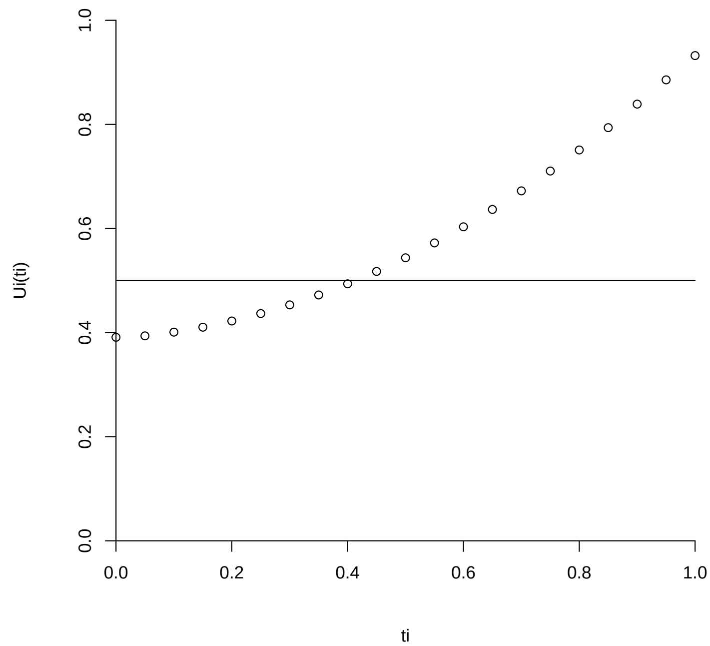

Figure 2: Expected utility under the second-best rule for uniform type distribution

this we display in Figure 2 agents' expected utility under the second best rule as a function of their type. If our model is given a public goods interpretation, as in Section 3, then individual rationality requires that all types' expected utility be at least 0.5. As Figure 2 shows, this condition is violated by the second best rule for all types less than or equal to 0.4. It is this violation of individual rationality that permits the second best rule to come as close to efficiency as displayed in Figure 1.

We next show in Figure 3, still for the case of the uniform distribution, the probability with which agent 1's preferred alternative A is chosen by a second-best rule. The method that we use for the graphical representation of these probabilities is the same as in Figure 1. All first-best rules have the feature that the probability of alternative A is zero above the diagonal. Below the diagonal, the probability of alternative A is indeterminate in first-best. In the second-best rule, alternative A receives positive probability if types are $(1/21, 1)$, even though this probability is zero in all first-best rules. This is a reflection of the same deviation from first-best that we already described when discussing the probability of the compromise B.

The probability assigned to A by the second-best rule for type pairs below the diagonal seems to follow no particular pattern. These probabilities are chosen by the optimization routine so as to provide at the interim stage incentives for agents 1 and 2 to report their valuations of the compromise truthfully. Beyond this, these probabilities have no efficiency implications.[13]

The main insights that we have found in the uniform distribution example concern the probability of the compromise B rather than the probabilities of the extreme alternatives A and C. To test the robustness of these conclusions, we have considered the second-best rule for 100 other probability distributions over the type space. We have generated these probability distributions by assigning to each type randomly a weight between 0 and 1, where these weights were drawn independently with uniform distribution over the interval [0,1]. We then normalized these weights by dividing them all by their sum so that they could be regarded as probabilities.

For each second-best rule we have calculated the overall ex-ante probability with which an alternative is implemented that could not be implemented by a first-best rule. We first show in Figure 4 a histogram of these probabilities. On the vertical axis we show absolute frequencies. The main point

[13] The matrix of probabilities that the second-best rule assigns to alternative A does not seem to be unique. We obtained different results using R and using Mathematica. The results reported in the paper were calculated using R. The probabilities of alternative B were independent of the software used.

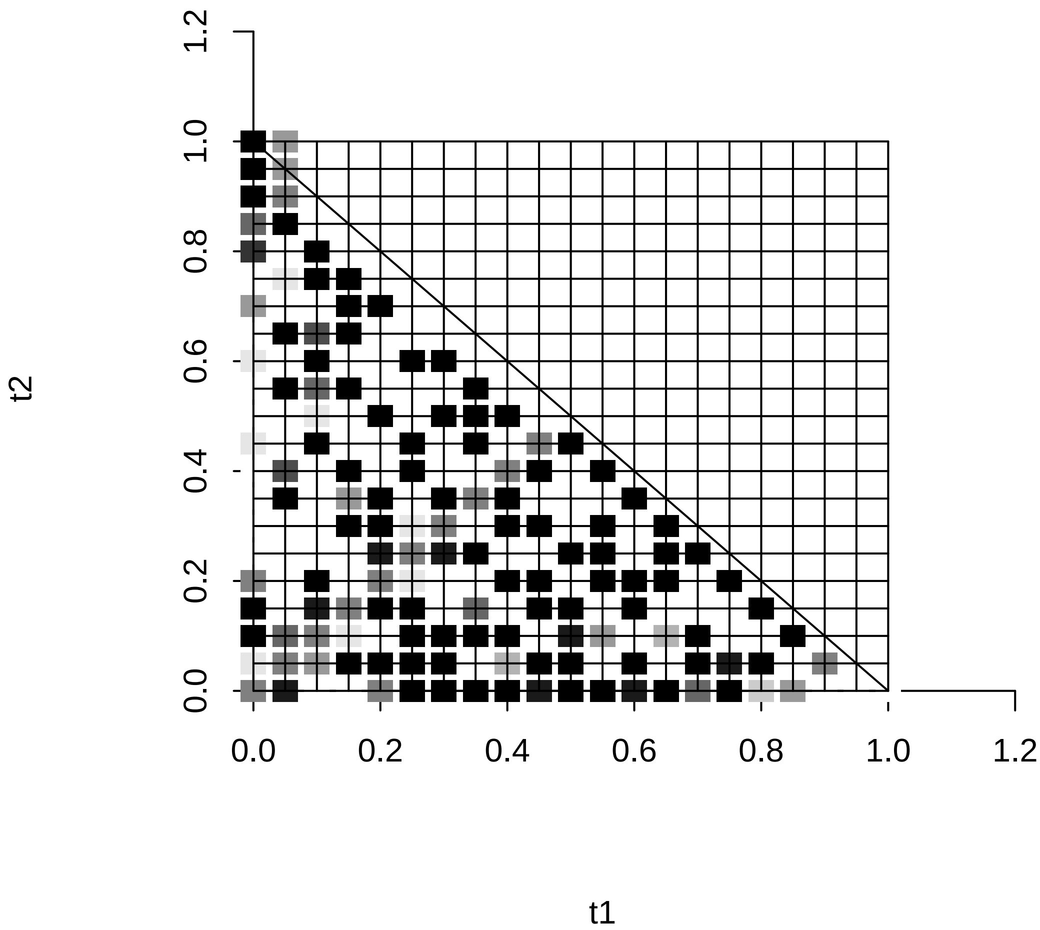

Figure 3: The probability of alternative A under the second-best rule for uniform type distribution

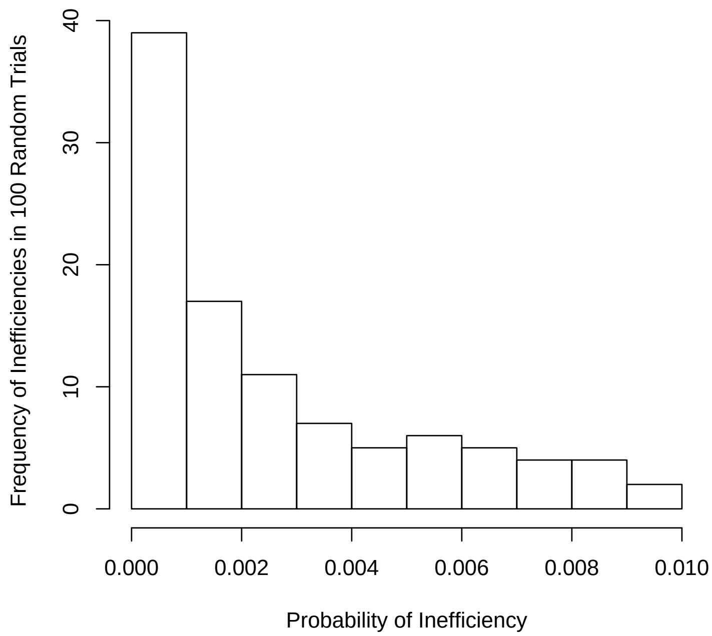

Figure 4: The frequency distribution of ex-ante probabilities of deviations from first-best for 100 randomly selected type distributions

to notice is that in all cases this probability was below 1/100. Thus, the observation that we made in the case of uniform distribution that deviations from first-best occur with low probability seems to hold quite generally.

Next, we indicate in Figure 5 for each pair of types the average probability with which the outcome under second-best deviates from first-best for this pair of types. The average is taken over our 100 randomly selected examples. For grid points that are marked in white the average probability of deviations is zero. For grid points that are marked in black the average probability of deviations is above 0.2. If the average probabilities of deviations was between 0 and 0.2, we have indicated the value by choosing an appropriate level of grey, where darker grey corresponds to higher probabilities.

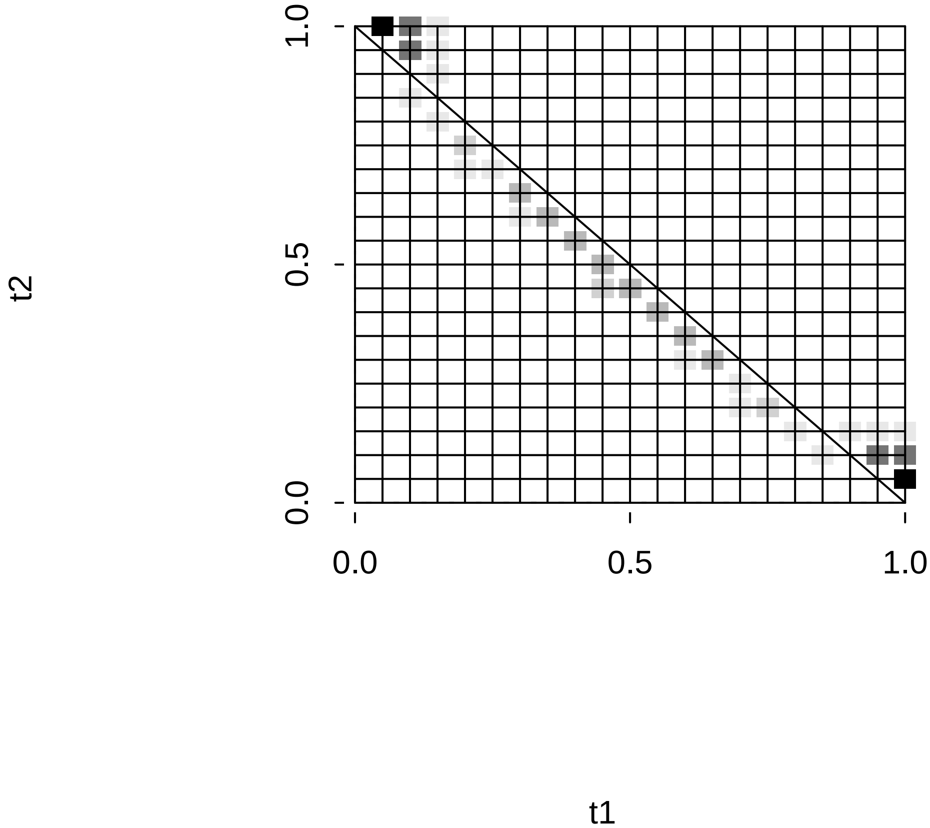

Figure 5: Average probabilities for deviations from first-best for 100 randomly selected type distributions

As in the uniform example, we find that deviations from first-best occur only if the sum of the players' types is relatively close to 1. The pairs of types for which deviations from first-best are most likely are those two pairs of types for which such deviations occur in the example of uniform distribution: $(1/20, 1)$ and $(1, 1/20)$. Deviations from first-best occur typically when the resulting welfare losses are small. Moreover, deviations from first-best are somewhat more likely when agents' valuations of the compromise are very different from each other, and seem somewhat less likely in the centre of Figure 5 where the two agents' valuations are close together.

8. Conclusion

For a simple compromise problem with non-transferrable utility we have shown the impossibility of implementing the first best, and we have determined for some particular examples a second-best decision rule. In future research we plan to extend our work to a scenario in which agents' rankings of the alternatives as well as their von Neumann Morgenstern utilities are privately observed. We suspect that in this setting second-best decision rules can only be determined numerically. We also plan to examine in more detail the robustness of the decision rules that we obtain in our simple Bayesian setting, and to compare these decision rules to decision rules which are optimal if informationally less demanding concepts of implementation are considered.

References

Berkelaar, M. and others (2004), *lpSolve: Interface to Lp_solve v. 5 to solve linear/integer programs*, R package, version 1.1.5.

Border, K. (1991), Implementation of Reduced Form Auctions: A Geometric Approach, *Econometrica* 59, 1175-1187.

Börgers, T. (1991), Undominated Strategies and Coordination in Normalform Games, *Social Choice and Welfare* 8, 65-78.

Cramton, P., R. Gibbons, and P. Klemperer (1987), Dissolving a Partnership Efficiently, *Econometrica* 55, 615-632.

D'Aspremont, C. and L.-A. Gérard-Varet (1979), Incentives and Incomplete Information, *Journal of Public Economics* 11, 25-45.

Dutta, B., H. Peters and A. Sen (2004), Strategy-Proof Cardinal Decision Schemes, mimeo., University of Warwick.

Gibbard, A. (1973), Manipulation of Voting Schemes: A General Result, *Econometrica* 41, 587-601.

Güth, W. and M. Hellwig (1986), The Private Supply of a Public Good, *Journal of Economics*, Supplement 5, 121-159.

Hellwig, M. (2003), Public-Good Provision with Many Participants, *Review of Economic Studies* 70, 589-614.

Holmström, B., and R. Myerson (1983), Efficient and Durable Decision Rules with Incomplete Information, *Econometrica* 51, 1799-1820.

Hortala-Vallve, R. (2004), Qualitative Voting, mimeo., London School of Economics.

Krishna, V. (2002), *Auction Theory*, San Diego: Academic Press.

Mailath, G. and A. Postlewaite (1990), Asymmetric Information Bargaining Problems with Many Agents, *Review of Economic Studies* 57, 351-368.

Milgrom, P., *Putting Auction Theory to Work*, Cambridge: Cambridge University Press, 2004.

Myerson, R., and M. Satterthwaite (1983), Efficient Mechanisms for Bilateral Trading, *Journal of Economic Theory* 29, 265-281.

R Development Core Team (2004), *R: A language and environment for statistical computing,* R Foundation for Statistical Computing, Austria: Vienna, `http://www.R-project.org`.

Rob, R. (1989), Pollution Claim Settlements Under Private Information, *Journal of Economic Theory* 47, 307-333.

Satterthwaite, M. A. (1975), Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions, *Journal of Economic Theory* 10, 187-217.