Efficient Compromising*

Tilman Börgers† and Peter Postl‡

*We are grateful to Mark Armstrong, Eddie Dekel, Rafael Hortala-Vallve, Peter Norman, an associate editor, and two anonymous referees for comments.

†Department of Economics, University of Michigan, 611 Tappan Street, 337 Lorch Hall, Ann Arbor, MI 48109-1220, U.S.A., tborgers@umich.edu.

‡*Corresponding author.* Department of Economics, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom, p.postl@bham.ac.uk, Phone: +44 121 415 8215, Fax: +44 121 414 7377.

Abstract

Two agents must select one of three alternatives. Their ordinal rankings are commonly known and diametrically opposed. Efficiency requires choosing the alternative the agents rank second whenever the weighted sum of their von Neumann Morgenstern utilities is higher than under either agent's favorite alternative. The agents' utilities of the middle-ranked alternative are i.i.d., privately observed random variables. In our setup, which is closely related to a public goods problem where agents face liquidity constraints but no participation constraints, decision rules that truthfully elicit utilities and implement efficient decisions do not exist. We provide analytical and numerical results on second-best rules.

Keywords: Arbitration; Compromise; Mechanism design without transferrable utility.

JEL classification: C72; D70; D80.

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.

We need to say first 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 because they seem inappropriate in many situations. Spouses, for example, rarely pay money to each other to resolve conflicts. Another context in which money payments are uncommon is 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 preference," 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* or an *interim* perspective (Holmström and Myerson [10]), 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. Decision rules are evaluated using the ex ante Pareto criterion. This is equivalent to maximizing a weighted sum of ex ante expected utilities. Not taking into account incentive compatibility of truthful reporting of types, a rule is efficient if and only if it picks for every realization of von Neumann Morgenstern utilities an alternative that maximizes the weighted sum of the two agents' utilities.

For such a first-best decision rule to be implementable when von Neumann Morgenstern utilities are privately observed, the rule needs to be incentive compatible. Our first main result is that no first-best decision rule is incentive compatible 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, that is, decision rules that are efficient among all incentive compatible 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 a full analytical solution to the second-best problem appears difficult. We then report a mixture of partial analytical, and more complete numerical results about second-best decision rules. Our results indicate that the shape of second-best rules is different from the shape of second-best rules in more familiar settings, and that the amount of inefficiency that second-best rules imply 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 in 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 [7], Satterthwaite [18], Dutta, Peters and Sen [6]), or undominated strategies (Börgers [2]). 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 [5], Güth and Hellwig [8], Rob [17], and Mailath and Postlewaite [14]. 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 liquidity constraint, which is absent from traditional models. The liquidity 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 [9]). 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 liquidity 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 [8], Rob [17], and Mailath and Postlewaite [14]. 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 liquidity constraint. Our results on second-best rules indicate that the amount of inefficiency implied by second-best rules in our set-up is much smaller than the inefficiency of second-best rules in the corresponding public goods set-up. The reason is that the liquidity constraints implicit in our model are less restrictive than the individual rationality constraints present in the public goods model.

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 explores second-best in a special case: equal welfare weights and uniform type distribution. For this case we give a detailed presentation of numerical findings as well as some partial analytical results. In Section 7 we pursue the numerical approach in a more general context. Whereas the bulk of the paper is concerned with ex ante efficiency we briefly weaken the efficiency concept in Section 8 and consider interim efficiency. Section 9 concludes.

2. The Model

Two agents $i = 1, 2$ 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. We refer to alternative B as the "compromise" because it is the middle-ranked alternative for each of 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.

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 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* $f : [0, 1]^2 \to \Delta(\{A, B, C\})$ *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 f, we denote for every $t_i \in [0, 1]$ by $p_i(t_i)$ the probability that agent i's favorite alternative is implemented, conditional on agent i's type being t_i, i.e.:

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

We 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. Welfare is defined as the weighted sum of the agents' ex ante expected utilities.

Definition 2 *For any* $\lambda \in [0.5, 1)$ *the* λ-weighted ex ante welfare *associated with decision rule* f *is:*

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

In this definition we focus without loss of generality on the case that agent 1's weight λ is at least 0.5, and we rule out the trivial case in which $\lambda = 1$.

As noted by Holmström and Myerson [10], the set of all decision rules maximizing λ-weighted ex ante welfare for some λ is the same as the set of all ex ante efficient rules. As Holmström and Myerson suggest, we shall compare rules that are classically efficient, i.e. ex ante efficient among all feasible rules ("first-best"), and rules that are ex ante incentive efficient, i.e. ex ante efficient among all incentive compatible, feasible rules ("second-best"). While the focus of this paper is on ex ante efficiency, we shall briefly consider in Section 8 interim efficiency. As Holmström and Myerson [10] point out, this is equivalent to allowing the welfare weight attached to each agent i to depend on that agent's type t_i.

The expression in Definition 2 can equivalently be written as:

$$\int_0^1 \int_0^1 \left(\lambda f_A(t_1, t_2) + \left[\lambda t_1 + (1-\lambda) t_2\right] f_B(t_1, t_2) + (1-\lambda) f_C(t_1, t_2)\right) g(t_1) g(t_2) dt_1 dt_2.$$

From this expression it is obvious which decision rules f maximize λ-weighted ex ante welfare among all decision rules. We call such decision rules λ-weighted first-best rules.

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

- *If* $\lambda = 0.5$*:*

$$\begin{aligned} t_1 + t_2 > 1 \quad &\Rightarrow \quad f_B(t_1, t_2) = 1 \\ t_1 + t_2 < 1 \quad &\Rightarrow \quad f_B(t_1, t_2) = 0 \end{aligned}$$

- *If* $\lambda > 0.5$*:*

$$\begin{aligned} \lambda t_1 + (1-\lambda) t_2 > \lambda \quad &\Rightarrow \quad f_B(t_1, t_2) = 1 \\ \lambda t_1 + (1-\lambda) t_2 < \lambda \quad &\Rightarrow \quad f_A(t_1, t_2) = 1 \end{aligned}$$

If $\lambda > 0.5$ the first-best decision rule is uniquely determined except for a set of types of measure zero. By contrast, if $\lambda = 0.5$, there are many first-best decision rules, and these rules differ from each other on a set of types of positive probability measure. The reason is that, for $\lambda = 0.5$, Definition 3 does not restrict the probabilities with which alternatives A and C are chosen if the compromise is *not* implemented.

Because types are privately observed, in practice one can only implement incentive compatible rules.

Definition 4 *A decision rule 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 purpose of this paper is to study the potential discrepancy between λ-weighted first-best rules and incentive compatible rules. For this purpose we focus on λ-weighted second-best rules.

Definition 5 *A decision rule f is called* λ-weighted second-best *if it maximizes λ-weighted ex ante welfare among all incentive compatible decision rules.*

We now discuss some features of our model. We begin with the modeling of the utility functions. Our model implies that for each interim preference ordering that a player might have there is a unique type of that player with these preferences. The main implicit restriction is that we rule out multiple types that have the same interim preferences, i.e. whose von Neumann Morgenstern utility functions differ only by an affine transformation. From the ex ante point of view, it could be important to keep types with identical interim preferences in the model, because welfare maximization might assign different allocations to these types. However, if two types have identical interim preferences, they will make the same choice, provided that the optimal choice is unique. An incentive compatible mechanism will not be able to assign different outcomes to these types.[1] Therefore, not much is lost by assuming that for each interim preference ordering there is only one type that has this preference ordering.

[1]For a more thorough discussion of this point, though in a different setting, see Hortala-Vallve [11, Proposition 1].

The fact that for each agent the normalization of utilities is the same across different types is not restrictive. Differences in the normalization of utilities for different types would reflect that these types receive different weight in the decision maker's ex ante expected utility maximization. These weight differences can equivalently be expressed by the probability distribution of types.

The assumption that types are identically distributed for the two players can easily be relaxed. In fact, it is immediate that our theoretical analysis and results in Section 4 would remain unchanged. The numerical analysis would change, although the findings that we report in Sections 6 and 7 are robust in the sense that small changes to the distributions would not change the results by much.

It is also potentially important that types are independently distributed. In mechanism design with transferrable utility, models with types that are not independent sometimes have incentive compatible rules that achieve first-best (e.g. Crémer and McLean [4]). The constructions used in this context in the literature do not immediately extend to a setting without transferrable utility. We have not yet explored relaxations of the independence assumption in our model.

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 [5], Güth and Hellwig [8], Rob [17], Mailath and Postlewaite [14]). 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 favorite 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 favorite alternative, and the probability with which the agent's favorite alternative is chosen by a given decision rule if the types are (t_1, t_2):

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

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 can model this by assuming that the marginal cost rise to infinity if 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 liquidity 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 liquidity constraint of individual agents is a first feature that distinguishes, to our knowledge, our set-up from all public good models that have been studied in the literature.

A second feature that distinguishes our set-up from the traditional public goods set-up is the absence of individual rationality constraints 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.[2] But in the context of arbitration there is no natural role for individual rationality.

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.

4. Incentive Compatibility

In this section we translate standard characterizations of incentive compatible decision rules into our setting. Because the proofs of these results are familiar from the literature, we omit them.

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 $t_i < t_i'$:

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

The first item in this Lemma states 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 shows that some of the probability has to be taken from the probability assigned to the agent's favorite 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 favorite alternative.

[2] An exception is d'Aspremont and Gérard-Varet [5].

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.[3]

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:

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

We can use the differential equation in the second item 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 as given the value of the interim expected utility at some boundary point. We choose here the highest type, i.e. $t_i = 1$, rather than, as is convention in the literature, the lowest type, $t_i = 0$, because this turns out to be more useful in the proof of Proposition 2 below. Apart from this modification, the proof of Lemma 3 is again standard, and is 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 $t_i \in [0, 1]$.

[3]See, for example, Section 5.1.1 of Krishna [13] for a proof of a similar result.

5. Impossibility of Implementing First-Best Rules

For asymmetric welfare weights the impossibility of implementing the first-best rule is elementary. In this case there is essentially one λ-weighted first-best decision rule. Moreover, this decision rule never implements the most preferred action of the agent who has the lower welfare weight. But, as Lemma 1 revealed, this probability is the main instrument by which an agent can be given incentives to reveal truthfully their type.

Proposition 1 *No λ-weighted first-best decision rule is incentive compatible for $\lambda > 0.5$.*

Proof: By Definition 3 the λ-weighted first-best decision rule implies $q_2(t_2) = 1 - G\left(1 - \frac{1-\lambda}{\lambda}t_2\right)$ and $p_2(t_2) = 0$ for all $t_2 \in [0,1]$ if $\lambda > 0.5$. But this violates condition (ii) of Lemma 1. For any t_2, t_2' with $0 < t_2 < t_2'$ the left hand and the right hand sides of the inequality in condition (ii) of Lemma 1 are negative, but the expression in the center of that inequality is zero.

Q.E.D.

The case of symmetric welfare weights is more subtle. In this case, there are multiple λ-weighted first-best decision rules, and the interim probability of each agent's most preferred alternative may vary with that agent's type. Thus, an instrument for providing incentives is available for each agent. Yet these instruments are never flexible enough to make the first-best decision rule incentive compatible.

Proposition 2 *No λ-weighted first-best decision rule is incentive compatible for $\lambda = 0.5$.*

We shall prove Proposition 2 by showing that, if $\lambda = 0.5$, any first-best decision rule that is incentive compatible has 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. The same reasoning is also

behind the impossibility of implementing the first-best in standard models of incentives in public goods provision (for example: Güth and Hellwig [8]). 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 [5]). We obtain a different result because, as explained in Section 3, our agents face individual liquidity constraints. These liquidity constraints imply lower boundaries for the utility of each type, even if no individual rationality is required.

Despite 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 [15, p.79] version of the proof of the Myerson-Satterthwaite [16] impossibility theorem. We begin the proof by arguing that Lemma 3 implies that all incentive compatible first-best decision rules have 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 "liquidity 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 to 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 $q_i(t_i)t_i = (1 - G(1 - t_i))t_i$. By continuity, therefore, the expected utility of type $t_i = 1$ has to be equal to 1.

We can now apply Lemma 3. Because sets of measure zero do not 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, first bullet point. The functions f_A and f_C are defined as follows.

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

We assume that players evaluate outcomes under this rule by the expected utility calculation shown 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 that 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 1 for all type vectors plays no role in the proof of Lemma 3. Therefore, the values of $p_i(t_i)$ for $i \in \{1,2\}$ and $t_i \in [0,1]$ that are implied by the above decision rule must be the same as the ones associated with 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: Uniform Type Distribution and Equal Welfare Weights

Analytical characterizations of second-best rules are difficult to obtain. Consider, for simplicity, the case of equal welfare weights: $\lambda = 0.5$. We could try to mimic the typical approach to characterizing second-best mechanisms, which proceeds by writing the optimization problem that defines second-best rules so that only directly welfare-relevant variables appear as choice variables. In our model, when the agents have equal welfare weights, 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 the theory of public goods, allocation rules that can be part of an incentive compatible scheme are those for which the interim expected allocations of the public good are monotonically increasing, and for which the agents' ex ante payments, as implied by an incentive compatibility condition like condition (ii) in Lemma 3, add up to the ex ante expected quantity of the public good. That is, ex ante, in expected terms, the contributions to the public good have to cover the cost of producing the public good. These conditions are not only necessary, but also sufficient for an allocation to be part of an incentive compatible scheme (see, e.g., Theorem 1 in Mailath and Postlewaite [14]) because, 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, incentive compatible, and that supports the same allocation rule and the same interim expected utilities.

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 [3]), then we violate the individuals' liquidity constraints. That is, individuals would be asked to give up so much probability of their favorite 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.[4]

In this section we begin by presenting some numerical results about second-best decision rules, focusing on the case of equal welfare weights: $\lambda = 0.5$ and uniform type distribution G. For this case, we also provide some analytical results that back up some of our numerical findings. In the next section we provide numerical results for other cases. For our numerical work we discretize the type space and postulate 80 equally spaced types.[5] 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. For the computations reported in this section we used the implementation of the interior point

[4]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 [1] has found such conditions. However, generalizing his results to our context seems hard.

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

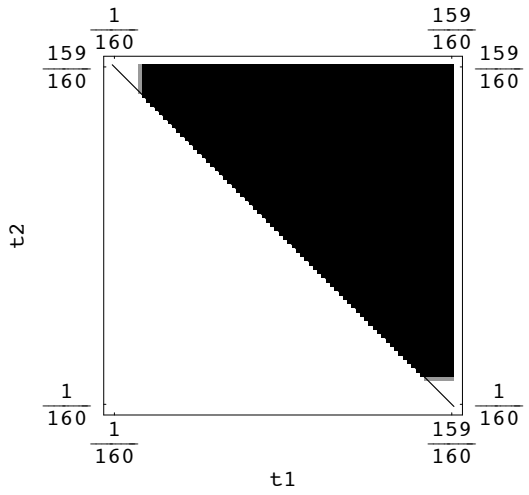

Figure 1: Probability of the compromise under the second-best rule for uniform type distribution and equal welfare weights

algorithm for linear programming that is available in MATHEMATICA 6.0 for LINUX x86. Our computations take account only of "local" incentive constraints: No type can gain from pretending to be a neighboring type. As in other standard models, local incentive constraints imply global incentive constraints. This is, for example, the logic behind Lemma 2.

Figure 1 shows the probability of the compromise B under the second-best rule. The figure shows a grid representing the possible 80×80 type pairs. Each grid point is associated with a square whose color represents the probability with which the compromise is chosen by the second-best rule. If the square is white, the probability of B is 0. If the square is black, the probability of B is 1. If the color is grey, the probability is between 0 and 1. A darker shade of grey implies a larger probability of B.[6]

A surprising aspect of Figure 1 is how similar the second-best and the first-best rules are. First-best decision rules assign probability 1 to the compromise B if the types are above the diagonal connecting the points $(1, 0)$ and $(0, 1)$, and they assign probability zero to B if the types are below this diagonal. The second-best rule is identical to this rule except that the area in which the compromise is implemented is cut off in the extreme corners of the unit square. Our calculations suggest that the compromise is implemented with very small probability only if the type of one of the agents is 13/160, and it is not implemented at all

[6]Gridlines have been suppressed in Figure 1, as well as in all other figures below.

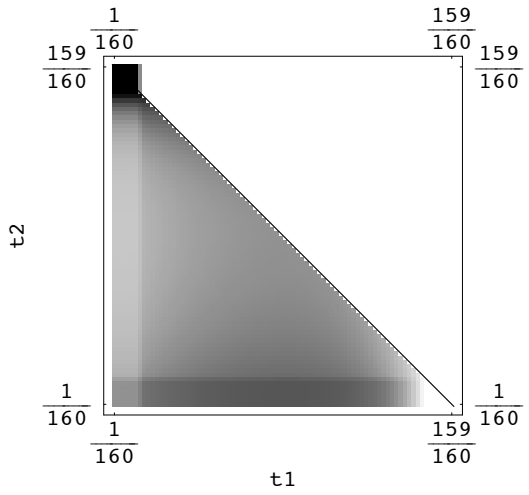

Figure 2: The probability of alternative A under the second-best rule for uniform type distribution and equal welfare weights

if the type of one agent is less than or equal to 11/160. A consequence is that the ex ante welfare loss under the second-best rule, relative to the first-best, is very small, approximately 0.015%.

It is instructive to compare Figure 1 to the second-best mechanism in the public goods problem that corresponds to the compromise problem. In this case, the second-best rule can easily be analytically determined. It implements production of the public good if and only if the sum of types is above 1.25. Geometrically, instead of cutting off corners as in Figure 1, the diagonal is shifted to the North East in the second-best public good rule. The associated relative welfare loss is approximately 2.23%. Thus, numerically, it appears that the interim individual rationality constraint in the public goods problem is a more restrictive constraint than the ex post liquidity constraint in the compromise problem.

In Figure 2 we report the probabilities of alternative A under the second-best rule. The method that we use for the graphical representation of these probabilities is the same as in Figure 1. In the first-best, the allocation of probabilities to alternatives A and C below the diagonal is not relevant for welfare. In the second-best rule, this probability is chosen by the optimization routine to provide at the interim stage incentives for agents 1 and 2 to report their true valuations of the compromise. The probability assigned to A by the second-best rule for type pairs below the diagonal does not seem to follow any particular pattern.

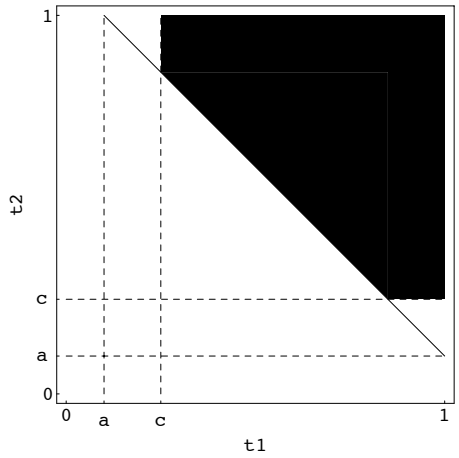

Figure 3: Cropped Triangle Rules

We now complement the numerical results of Figures 1 and 2 by some analytical insights. To make the problem analytically tractable we consider a subclass of decision rules with only two parameters. We choose this subclass so that it includes a rule that is very close to the one shown in Figure 1, and so that it includes the rule that would be second-best in the public goods setting that corresponds to our model. The subclass of decision rules to which we restrict attention is described in Definition 6.

Definition 6 *A decision rule f is called a* cropped triangle rule *if the probability of the compromise is of the form:*

$$f_B(t_1, t_2) = \begin{cases} 1 & \text{if } t_1 \geq c,\ t_2 \geq c \text{ and } t_1 + t_2 \geq 1 + a, \\ 0 & \text{otherwise,} \end{cases}$$

where $a \in [0, 1]$ *and* $c \in [a, \frac{1+a}{2}]$.

The function f_B for a typical two-parameter rule is illustrated in Figure 3, where $f_B(t_1, t_2) = 1$ in the shaded area.

In Appendix A, contained in a supplementary document archived in the "Supplementary Materials" section of the J. Econ. Theory web site, we show that among all incentive compatible *cropped triangle rules* those that maximize expected welfare with equal welfare weights have parameters $a = 0$ and $c = c^*$ where c^* is the unique c that solves $-1+12c-6c^2-4c^3 = 0$ in the interval $[0, 1]$ ($c^* \approx 0.0874$). Note how similar this solution is to the rule of Figure 1

where corners are cut at approximately $13/160 \approx 0.0813$. For the optimal *cropped triangle rule* the welfare loss relative to first-best is $2(c^*)^3/7$, which is approximately 0.0191%.

In Appendix B, contained in the supplementary material available from the J. Econ. Theory web site, we also show that the class of *cropped triangle rules* includes a rule that can be analytically shown to be second-best in the public goods problem not just among all incentive compatible *cropped triangle rules*, but among *all* rules, and even if one neglects that f_A and f_B need to be between zero and one. This rule has $a = 1.25$ and $c = 0$. We show in Appendix B that this rule can be implemented with probabilities f_A and f_C that are between zero and one. In other words, in the public goods problem that corresponds to our problem, if one determines second-best taking into account only interim individual rationality, but not the ex post liquidity constraints, then one obtains an optimal solution that also satisfies the ex post liquidity constraints. In this sense, the interim individual rationality constraints are more restrictive than the ex post liquidity constraints. This explains why the welfare loss is larger in the public goods problem than in the compromise problem.

7. Second-Best Rules: The General Case

In this section we explore the robustness of our insights into second-best decision rules that we obtained in Section 6 for uniformly distributed types and equal welfare weights. The first step in our robustness check is to consider changes in the type distribution G while maintaining the assumption of equal welfare weights. We shall focus on the case that the types follow a discretized Beta-distribution. We vary separately each of the two shape-parameters of the Beta-distribution from 0.5 to 5 in increments of 0.5. We thus obtain 100 different type distributions. The uniform distribution corresponds to the case that both shape parameters are equal to 1.

We first attempt to give some insight into how the second-best decision rule varies with the type distribution when welfare weights are equal. To this end we have computed numerically the second-best decision rule for all 100 type distributions. We report in Figure 4 for each pair of types the average deviation of the second-best decision rule from the first-best rule, where the average is taken across our 100 type distributions. For grid points marked by white

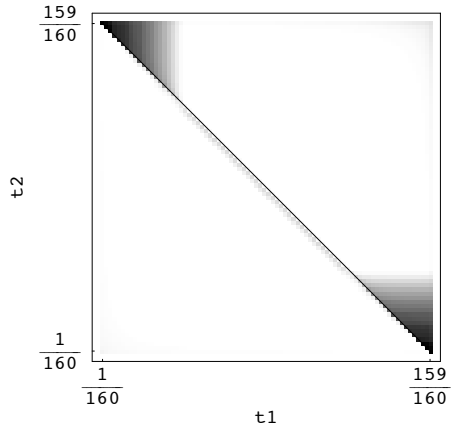

Figure 4: Average deviations from first-best for 100 type distributions

squares the average deviation from first-best is zero. For grid points marked in black the average deviation is 1. If the average deviation is between zero and 1, we have indicated the value by choosing an appropriate level of grey, where darker grey implies a larger deviation.

Figure 4 suggests that the observation made in the case of the uniform distribution that deviations from first-best occur only in the extreme corners of the unit square seems to hold regardless of the type distribution. The different shades of grey in Figure 4 indicate that the 100 second-best rules differ from each other only with regard to the threshold at which the extreme corners of the first-best decision rule have been cropped. The magnitude of the threshold appears to be related to the type distribution.

Next, we describe how the magnitude of the welfare loss under the second-best decision rule varies with the type distribution. For each of our 100 second-best rules we have calculated the associated ex ante welfare loss, relative to the first-best. Figure 5 displays a histogram of these welfare losses. The main point to notice is that in all cases the ex ante welfare loss is less than 0.025%. Thus, the observation made in the case of the uniform distribution that the ex ante welfare loss is very small seems to hold quite generally.

To provide a standard of comparison for evaluating these very small welfare losses we have computed numerically the second-best public good rule, and the relative welfare loss associated with it, for all 100 type distributions.[7] We find that in all 100 cases the relative

[7]To reduce computing time, the comparison of the second-best in the compromise and the public good settings is based on a discretization of 20 types.

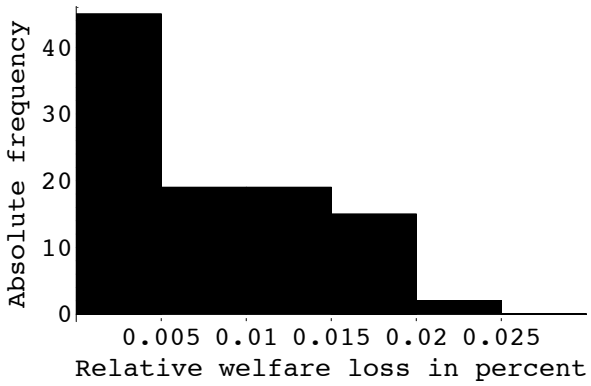

Figure 5: Frequency distribution of ex ante welfare losses for 100 type distributions

welfare loss associated with the second-best public goods rule is strictly larger than the relative welfare loss under the second-best in the compromise setting. This suggests that also for distributions other than the uniform distribution the participation constraints are more restrictive than the ex post liquidity constraints. To explore this observation further, we have calculated numerically, for all 100 type distributions, the second-best decision rule when both participation constraints *and* ex post liquidity constraints are imposed. Our computations yield the same level of ex ante social welfare in the second-best of the public good setting and in the second-best of the setting with participation constraints *and* liquidity constraints. This indicates that, just as with uniformly distributed types, the participation constraints are more restrictive than the ex post liquidity constraints in the sense that the latter constraints will not be binding if participation constraints are imposed.

The second part of our robustness check is to study, for the case of uniformly distributed types, the effect of a change in agent 1's welfare weight λ. We have computed the second-best decision rule and the relative welfare loss associated with it for 52 different values of λ equally spaced between 0.5 and approximately 0.99. We first study the effect of a change in agent 1's welfare weight on the shape of the second-best decision rule. We display in Figures 6 and 7 our results for two examples of asymmetric welfare weights: $\lambda \approx 0.505$ and $\lambda = 0.6$.

Figures 6 and 7 illustrate the second-best decision rules by displaying the probability of alternative A on the left hand side and the probability of the compromise B on the right hand side. Both figures show, in addition to the diagonal that connects points $(0, 1)$ and $(1, 0)$, a steeper line that represents all type-pairs for which the λ-weighted sum of the agents' types

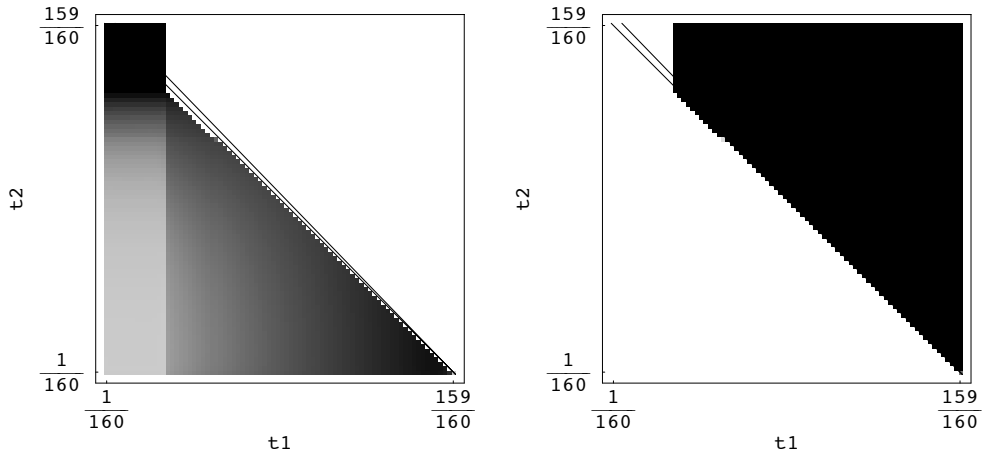

Figure 6: Probabilities of A and B, resp., under the second-best rule for uniform type distribution and $\lambda \approx 0.505$

equals agent 1's welfare weight λ. For type-pairs above this line the first-best rule selects the compromise B with probability 1. Below this line it selects alternative A.

The second-best probabilities of the compromise in Figures 6 and 7 involve two types of distortions relative to the first-best: First, a distortion of the slope of the line above which the first-best decision rule implements the compromise B. This type of distortion is absent in the case of equal welfare weights. Second, the area in which the second-best rule selects the compromise appears to be cropped at the top left corner of the unit square. Figure 7 suggests furthermore that the area in which the second-best rule selects the compromise is also cropped in the bottom right corner of the unit square once agent 1's welfare weight becomes sufficiently large. These second distortions are reminiscent of the distortions that we found in the case of equal welfare weights. The second-best probabilities of alternative A in Figures 6 and 7 show that under asymmetric welfare weights the second-best rule differs from first-best also in that agent 1's favorite alternative A is *not* always chosen when the compromise is not selected, but instead alternative C is sometimes chosen. This is necessary to provide incentives to agent 1 to reveal his type.

We finally describe how the magnitude of the relative welfare loss under the second-best decision rule varies with agent 1's welfare weight. The relation is non-monotone. Welfare loss increases for λ close to, but above 0.5, and then decreases. The maximum welfare loss is approximately 5%. Two opposing forces cause the non-monotonicity. To provide

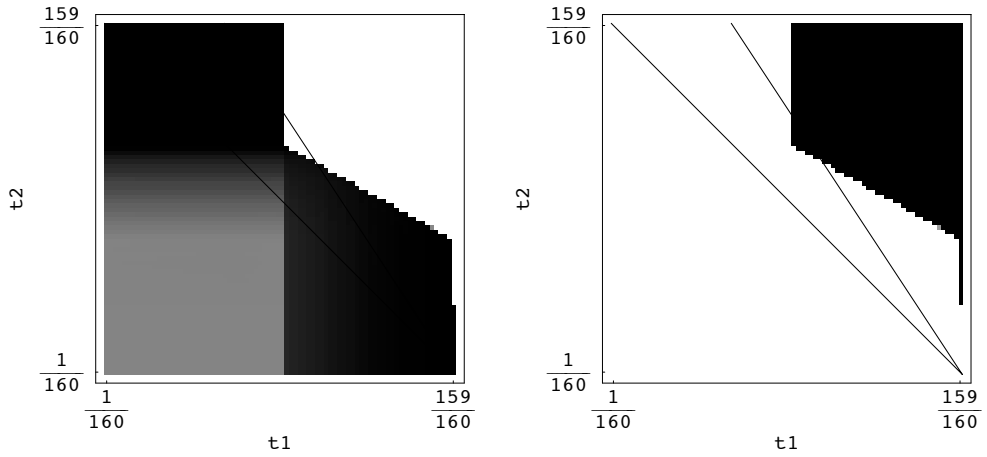

Figure 7: Probabilities of A and B, resp., under the second-best rule for uniform type distribution and $\lambda = 0.6$

incentives the second-best decision rule must choose alternative C with positive probability. The relative importance of this deviation from first-best increases as λ gets larger. On the other hand, for large λ, first-best and second-best implement the compromise less frequently, and therefore the need to provide incentives decreases, and with it the deviation of second-best from first-best.

8. Interim Efficient Rules

So far we have evaluated decision rules using welfare weights that do not depend on types. In this section we consider briefly the case that welfare weights are allowed to depend on types. There are thus two functions $\lambda_i : [0,1] \to \mathbb{R}_+$ for $i = 1, 2$, and the welfare associated with a decision rule f is:

$$\int_0^1 \lambda_1(t_1) U_1(t_1) g(t_1) dt_1 + \int_0^1 \lambda_2(t_2) U_2(t_2) g(t_2) dt_2.$$

We revisit with this specification of welfare weights the question addressed in Section 5 whether first-best rules are incentive compatible. We give two examples. Suppose that weights are given by: $\lambda_i(t_i) = t_i$ for $i = 1, 2$. The first-best rule is then uniquely determined except for a set of pairs (t_1, t_2) of measure zero. The rule is shown in Figure 8. Clearly, this rule is not incentive compatible for any distribution G of types. As each agent's type

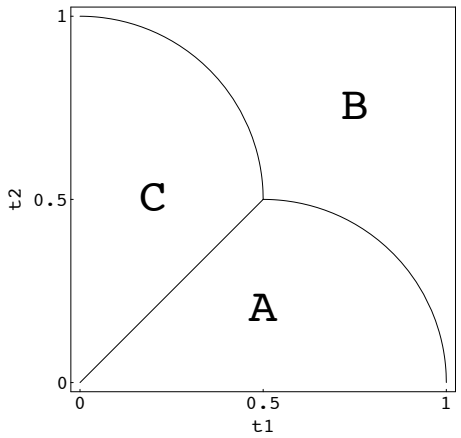

Figure 8: First best rule when $\lambda_i(t_i) = t_i$ for $i = 1, 2$.

increases from 0 to 0.5, both the interim probability of B and the interim probability of the agent's most preferred alternative increase. Therefore, each agent has an incentive to report their type as $t_i = 0.5$ even if the true type is $t_i < 0.5$.

In Figure 8, as agents report a higher type, the first-best mechanism not only infers that agents have a higher valuation of the compromise, but it also attaches higher weights to agents. It might be argued that this effect generates the incentive to distort preferences for agents with low types. We next consider an example in which agents' weights are decreasing in their types: $\lambda_i(t_i) = 1 - t_i$ for $i = 1, 2$. The first-best rule for these weights is shown in Figure 9. Once again it is obvious that this rule is not incentive compatible for any distribution G of types. As each agent's type increases from 0 to 0.5, the probability of the most preferred alternative of that agent decreases, while the probability of the compromise is zero. Therefore, each agent will have an incentive to report their type as $t_i = 0$ even if the true type is $t_i \in (0, 0.5]$. The decreasing weight creates an incentive to understate one's type. The two examples that we have given suggest that for most λ_i functions the first-best will not be incentive compatible. For given λ_i functions, the first-best choice will be uniquely determined for almost all (t_1, t_2), and there is no reason why in general the first-best rule should provide adequate incentives to truthfully reveal one's type.[8]

[8]We have not proved a formal version of the intuition developed in the text, nor have we analyzed second-best rules in the case of type-dependent weights. It seems that a further investigation of these issues would not add much further insight beyond what we have obtained so far.

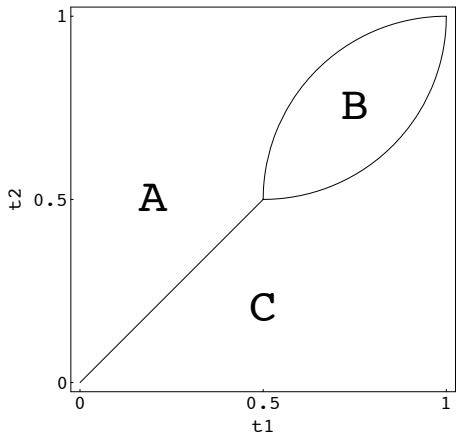

Figure 9: First best rule when $\lambda_i(t_i) = 1 - t_i$ for $i = 1, 2$.

9. Conclusion

For a simple compromise problem with non-transferrable utility we have shown the impossibility of implementing the first-best, and we have determined some characteristics of second-best decision rules. 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

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

[2] T. Börgers, Undominated Strategies and Coordination in Normalform Games, Soc. Choice Welfare 8 (1991), 65-78.

[3] P. Cramton, R. Gibbons, P. Klemperer, Dissolving a Partnership Efficiently, Econometrica 55 (1987), 615-632.

[4] J. Crémer, R. P. McLean, Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist When Demands Are Interdependent, Econometrica 53 (1985), 345-61.

[5] C. d'Aspremont, L.-A. Gérard-Varet, Incentives and Incomplete Information, J. Public Econ. 11 (1979), 25-45.

[6] B. Dutta, H. Peters, A. Sen, Strategy-Proof Cardinal Decision Schemes, Soc. Choice Welfare 28 (2007), 163-179.

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

[8] W. Güth, M. Hellwig, The Private Supply of a Public Good, J. Econ. (Z. Nationalökon.), Supplement 5 (1986), 121-159.

[9] M. Hellwig, Public-Good Provision with Many Participants, Rev. Econ. Stud. 70 (2003), 589-614.

[10] B. Holmström, R. Myerson, Efficient and Durable Decision Rules With Incomplete Information, Econometrica 51 (1983), 1799-1819.

[11] R. Hortala-Vallve, Inefficiencies on Linking Decisions, University of Oxford, Working Paper, 2007.

[12] M. Jackson, H. Sonnenschein, Overcoming Incentive Constraints by Linking Decisions, Econometrica 75 (2007), 241-257.

[13] V. Krishna, Auction Theory, Academic Press, San Diego, 2002.

[14] G. Mailath, A. Postlewaite, Asymmetric Information Bargaining Problems with Many Agents, Rev. Econ. Stud. 57 (1990), 351-368.

[15] P. Milgrom, Putting Auction Theory to Work, Cambridge University Press, Cambridge, 2004.

[16] R. Myerson, M. Satterthwaite, Efficient Mechanisms for Bilateral Trading, J. Econ. Theory 29 (1983), 265-281.

[17] R. Rob, Pollution Claim Settlements Under Private Information, J. Econ. Theory 47 (1989), 307-333.

[18] M. A. Satterthwaite, Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions, J. Econ. Theory 10 (1975), 187-217.

Online Supplement to: "Efficient Compromising"

by Tilman Börgers[1] and Peter Postl[2]

published in the Journal of Economic Theory

In this online appendix we prove the analytical results mentioned in Section 6 of the main paper. The structure of this appendix is as follows: In Appendix A we characterize the optimal cropped triangle rule. In Appendix B we show that the second-best in the public goods problem can be implemented using an incentive compatible cropped triangle rule.

Appendix A: Welfare Maximization Among Incentive Compatible Cropped Triangle Rules

The claim that we prove in this appendix concerns the maximization of expected welfare with weight $\lambda = 0.5$ among all incentive compatible cropped triangle rules. To maximize expected welfare among all these rules we shall restrict attention to cropped triangle rules that satisfy the following symmetry condition: $f_A(t,t') = f_C(t',t)$ and $f_B(t,t') = f_B(t',t)$ for all $(t,t') \in [0,1]^2$. To see that this is without loss of generality consider any incentive compatible decision rule (f_A, f_B, f_C), and define a new, symmetric decision rule by swapping the roles of players 1 and 2 and alternatives A and C with probability 0.5. Thus, with probability 0.5 the original rule is applied, and with probability 0.5 player 1 finds himself in the role of player 2 in the original rule. But since the original rule was incentive compatible for players 1 and 2, so is the new rule. Moreover, with equal welfare weights, expected welfare remains unchanged. Note that the argument is not restricted to cropped triangle rules but is general. For cropped triangle rules, moreover, the function f_B is symmetric by construction. The above argument only establishes that for cropped triangle rules it is without loss of generality to assume that also the functions f_A and f_C satisfy the symmetry condition.

[1]Department of Economics, University of Michigan, 611 Tappan Street, 337 Lorch Hall, Ann Arbor, MI 48109-1220, U.S.A., tborgers@umich.edu.

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

The structure of our argument will be as follows. First, we establish a necessary condition that all functions f_B that are part of a symmetric, incentive compatible cropped triangle rule have to satisfy. Then we determine the welfare maximizing choice of f_B where f_B is as in Figure 3 of the main paper and satisfies the necessary conditions. Thus, we solve a relaxed optimization problem. Finally, we show that the solution satisfies the constraints of the original optimization problem, that is, that we can construct functions f_A and f_C that make the rule f_B incentive compatible.

If a function f_B is part of a symmetric, incentive compatible cropped triangle rule, then there must be interim probability functions p_i mapping types into interim probabilities of preferred outcomes that make the rule f_B incentive compatible. For given f_B the functions p_i are determined by Lemma 3, part (ii) (see main paper), once we have fixed the boundary values $p_i(1)$. By the symmetry assumption, $p_i(1)$ will be the same for both i, and we denote it for simplicity by p. Our necessary condition will be that it must be possible to find some $p \in [0,1]$ such that, if we substitute this p for $p_i(1)$ in Lemma 3, part (ii), we obtain functions $p_i(t_i)$ that satisfy together with f_B the "ex ante adding up" constraint, i.e. the ex ante expected probability of the compromise and the ex ante expected values of $p_1(t_1)$ and $p_2(t_2)$ add up to one.

To work out the necessary condition in detail, we first note that for cropped triangle rules, the interim probability of the compromise is:

$$q_i(t_i) = \begin{cases} 0 & \text{if } 0 \leq t_i \leq c, \\ t_i - a & \text{if } c \leq t_i \leq 1 + a - c, \\ 1 - c & \text{if } 1 + a - c \leq t_i \leq 1. \end{cases}$$

The ex ante probability of the compromise, that is the ex ante expected value of q_i (for i either 1 or 2) can most easily be calculated as the size of the shaded area in Figure 3 of the main paper, which is:

$$\frac{1}{2}(1-a)^2 - (c-a)^2$$

where the size of the shaded area in Figure 3 of the main paper was determined as the size

of a rectangular triangle with two sides of length $1-a$ minus the size of the two smaller triangles that are "cropped" in Figure 3 of the main paper, and that are rectangular with two sides of length $c-a$.

Next, we infer, using Lemma 3 of the main paper, the interim probabilities of the most preferred alternatives p_i. Obviously, if $1+a-c \leq t_i \leq 1$, incentive compatibility requires:

$$\begin{aligned} p_i(t_i) &= p_i(1) \\ &= p. \end{aligned}$$

If $c \leq t_i \leq 1+a-c$ we have

$$\begin{aligned} p_i(t_i) &= p_i(1) + q_i(1) - q_i(t_i)t_i - \int_{t_i}^{1} q_i(s_i)ds_i \\ &= p + (1-c) - (t_i - a)t_i \\ &\quad - \int_{t_i}^{1+a-c} (s_i - a)ds_i - (1-(1+a-c))(1-c) \end{aligned}$$

where the integral was calculated in two parts, and the second part equals the size of a rectangle with sides of length $1-(1+a-c)$ and $1-c$. We continue the calculation as follows:

$$\begin{aligned} &= p + (1-c) - (t_i - a)t_i \\ &\quad - \left[\frac{1}{2}(s_i)^2 - as_i\right]_{t_i}^{1+a-c} - (c-a)(1-c) \\ &= p + (1-c) - (t_i - a)t_i \\ &\quad -\frac{1}{2}(1+a-c)^2 + a(1+a-c) \\ &\quad +\frac{1}{2}(t_i)^2 - at_i - (c-a)(1-c) \\ &= p + \frac{1}{2}(1+a-c)^2 - \frac{1}{2}(t_i)^2 \end{aligned}$$

For $0 \leq t_i \leq c$ we have:

$$\begin{aligned} p_i(t_i) &= p_i(1) + q_i(1) - q_i(t_i)t_i - \int_{t_i}^{1} q_i(s_i)ds_i \\ &= p + (1-c) - \frac{1}{2}(1-a)^2 + (c-a)^2 \end{aligned}$$

where the integral was calculated as the size of a rectangular triangle with two sides of length $1-a$ minus the size of the two smaller triangles that are "cropped" in Figure 3 of the main paper.

Now we are in a position to determine the ex ante expected value of p_i:

$$\begin{aligned} & p + c\left[(1-c) - \frac{1}{2}(1-a)^2 + (c-a)^2\right] \\ & + \int_{c}^{1+a-c} \frac{1}{2}(1+a-c)^2 - \frac{1}{2}\left(t_i\right)^2 dt_i \\ = \; & p + c\left[(1-c) - \frac{1}{2}(1-a)^2 + (c-a)^2\right] \\ & + (1+a-2c)\frac{1}{2}(1+a-c)^2 - \frac{1}{6}\left[(t_i)^3\right]_{c}^{1+a-c} \\ = \; & p + c\left[(1-c) - \frac{1}{2}(1-a)^2 + (c-a)^2\right] \\ & + (1+a-2c)\frac{1}{2}(1+a-c)^2 - \frac{1}{6}(1+a-c)^3 + \frac{1}{6}c^3 \\ = \; & \frac{1}{3} + p - c + c^2 + \frac{1}{3}c^3 + a + a^2 + \frac{1}{3}a^3 - 2ac - a^2c \end{aligned}$$

where the last step was verified by MATHEMATICA.

The necessary condition with which we shall work is now that twice this value, plus the ex ante expected value of q_i must equal 1:

$$\begin{aligned} \frac{2}{3} + 2p - 2c + 2c^2 + \frac{2}{3}c^3 + 2a + 2a^2 + \frac{2}{3}a^3 - 4ac - 2a^2c & \\ + \frac{1}{2}(1-a)^2 - (c-a)^2 &= 1 \Leftrightarrow \\ -\frac{1}{12} - \frac{1}{2}a - \frac{3}{4}a^2 - \frac{1}{3}a^3 + c + ac + a^2c - \frac{1}{2}c^2 - \frac{1}{3}c^3 &= p \end{aligned}$$

where the last step was again verified by MATHEMATICA. The constraint that we shall work with when maximizing expected welfare is now that there must be some $p \in [0,1]$ such that

the above equation holds. This is the same as the requirement that the left hand side of the above equation is contained in $[0,1]$. In the following, we denote this expression by $E(a,c)$.

We seek to determine the welfare-maximizing choice of a and c subject to the condition $E(a,c) \in [0,1]$. We proceed in two steps. We first ask which choices, if any, of $c \in \left[a, \frac{1+a}{2}\right]$ are optimal for given $a \in [0,1]$. Then we ask which choice of a is best.

Note that for given a welfare is maximized by choosing c as small as possible. The smallest admissible value of c is $c = a$. If for this choice of c we have $E(a,c) \in [0,1]$, then it is the optimal choice.

$$\begin{aligned} E(a,a) &\in [0,1] \Leftrightarrow \\ -\frac{1}{12} - \frac{1}{2}a - \frac{3}{4}a^2 - \frac{1}{3}a^3 + a + a^2 + a^3 - \frac{1}{2}a^2 - \frac{1}{3}a^3 &\in [0,1] \Leftrightarrow \\ -\frac{1}{12} + \frac{1}{2}a - \frac{1}{4}a^2 + \frac{1}{3}a^3 &\in [0,1]. \end{aligned}$$

Numerically, we can determine using MATHEMATICA that this is the case if and only if

$$a \geq 0.178846.$$

For smaller values of a MATHEMATICA shows that we have: $E(a,a) < 0$. On the other hand, $E(a, \frac{1+a}{2}) > 0$ where $c = \frac{1+a}{2}$ is the largest admissible value of c. The proof is as follows:

$$\begin{aligned} E(a, \frac{1+a}{2}) &= -\frac{1}{12} - \frac{1}{2}a - \frac{3}{4}a^2 - \frac{1}{3}a^3 + \frac{1+a}{2} + a\frac{1+a}{2} \\ &\quad + a^2\frac{1+a}{2} - \frac{1}{2}\left(\frac{1+a}{2}\right)^2 - \frac{1}{3}\left(\frac{1+a}{2}\right)^3 \\ &= \frac{1}{8}\left(2 + a + 12a^2 + 13a^3\right) > 0 \end{aligned}$$

where the simplification in the last step was obtained using MATHEMATICA. By the continuity of $E(a,c)$ in c we can now conclude that there is a smallest $c \in [a, \frac{1+a}{2}]$ such that $E(a,c) = 0$. This c is the optimal choice, given a.

To determine the optimal choice of a, we first prove that the optimal choice of c increases in a. This is obvious for $a \geq 0.178846$. For smaller values of a it follows from the fact that the value of $E(a, c)$ decreases in a. To show this we calculate:

$$\begin{aligned} \frac{\partial E}{\partial a} &= -\frac{1}{2} - \frac{3}{2}a - a^2 + c + 2ac \\ &\leq -\frac{1}{2} - \frac{3}{2}a - a^2 + \frac{1+a}{2} + 2a\frac{1+a}{2} \\ &= 0. \end{aligned}$$

As the optimal c is increasing in a, it follows immediately that expected welfare is decreasing in a, assuming that for each a the optimal c is chosen. Therefore, the optimal choice of a is $a = 0$. The corresponding choice of c is the smallest c for which $E(0, c) = 0$. This equation is equivalent to:

$$-\frac{1}{12} + c - \frac{1}{2}c^2 - \frac{1}{3}c^3 = 0.$$

MATHEMATICA shows that there is a unique solution c^* in $[0, 1]$ of this equation, and that it is: $c^* \approx 0.087373$.

We have now solved the relaxed maximization problem, and we complete the argument by constructing functions f_A and f_C that make the optimal f_B incentive compatible. We define f_A as follows:

$$f_A(t_1, t_2) = \begin{cases} \frac{1}{2} & \text{if } t_1 \leq c^* \text{ and } t_2 \leq c^*, \\ 1 - \frac{(1-c^*-t_2)(t_2-c^*)}{2c^*} & \text{if } t_1 \leq c^* \text{ and } c^* \leq t_2 \leq 1 - c^* \\ 1 & \text{if } t_1 \leq c^* \text{ and } t_2 > 1 - c^* \\ \frac{(1-c^*-t_1)(t_1-c^*)}{2c^*} & \text{if } c^* < t_1 \leq 1 - c^* \text{ and } t_2 \leq c^*, \\ \frac{1}{2} & \text{if } c^* < t_1 \leq 1 - c^* \text{ and } c^* < t_2 \leq 1 - t_1, \\ 0 & \text{otherwise.} \end{cases}$$

The function f_C is defined symmetrically, and we omit the formal definition. The construction of f_A is shown in Figure 1. In Figure 1 we refer to a function h. We define for every $t \in [0, 1]$:

$$h(t) \equiv \frac{(1 - c^* - t)(t - c^*)}{2c^*}.$$

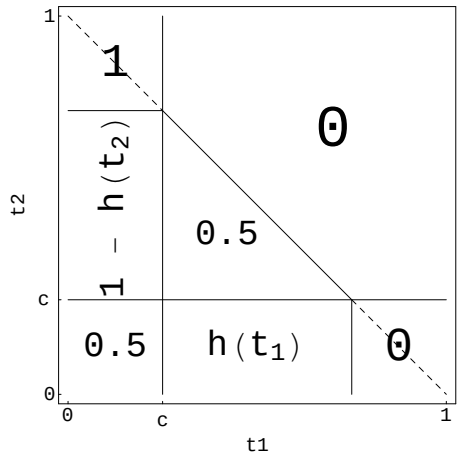

Figure 1: The probability of alternative A under the optimal cropped triangle rule

To check that what we have defined are actually probabilities we need to verify that

$$h(t) \in [0,1] \text{ for all } t \in [c^*, 1-c^*].$$

It is obvious that $h(t)$ is non-negative for all relevant t. Plotting it in MATHEMATICA one can verify that it is never more than 1. We also need to check that the probabilities that we have defined add up to 1 for every type vector. This is obvious.

It remains to verify that these probabilities give rise to the interim probabilities $p_i(t_i)$ that make the decision rule incentive compatible. Clearly, the implied interim probabilities of the compromise $q_i(t_i)$ are monotonically increasing in type t_i. It remains to verify that the interim probabilities of the preferred alternatives are those required by part (ii) of Lemma 3 in the main paper. We have:

$$\begin{aligned}
p_i(t_i) &= p = 0 \text{ when } 1-c^* \leq t_i \leq 1. \\
p_i(t_i) &= c^* h(t_i) + \frac{1}{2}(1-c^*-t_i) \\
&= \frac{(1-c^*-t_i)(t_i-c^*)}{2} + \frac{1}{2}(1-c^*-t_i) \\
&= \frac{(1-c^*-t_i)(1-c^*+t_i)}{2} \\
&= \frac{1}{2}(1-c^*)^2 - \frac{1}{2}(t_i)^2 \text{ when } c^* \leq t_i \leq 1-c^*.
\end{aligned}$$

In these first two cases we thus obtain the expressions that are required by incentive com-

patibility and that were derived above. For the remaining, third, case: $0 \leq t_i \leq c^*$, no calculation is needed. The conclusion can be derived from two observations. First, the total probability of the preferred alternative that is assigned by our rule in this case is ex ante the same as required by incentive compatibility. This is because our mechanism clearly has the property that ex ante probabilities add up to 1. Indeed, it also has this property ex post. Thus, the probability assigned to the preferred alternative ex ante if $0 \leq t_i \leq c^*$ is 1 minus the probability assigned to the preferred alternative if $t > c^*$. Moreover, the probabilities assigned to the preferred alternative and that are required by incentive compatibility add up to 1. This is indeed the constraint under which we determined the optimal mechanism. Because for $t_i > c^*$ our mechanism assigns the correct probabilities to the preferred alternative, the same must be true ex ante if $0 \leq t_i \leq c^*$. The second observation is that incentive compatibility requires the probability assigned to the preferred alternative to be constant for $0 \leq t_i \leq c^*$. Our mechanism has this property. Therefore, it must assign exactly the probabilities required by incentive compatibility to the preferred alternative for $0 \leq t_i \leq c^*$.

Appendix B: The Second Best Public Goods Rule as a Cropped Triangle Rule

In this appendix we prove the claim in Section 6 of the main paper that the function f_B that corresponds to the second-best in the public goods problem with equal welfare weights and uniform type distribution can be implemented as an incentive compatible cropped triangle rule. Recall that the second-best public goods decision rule corresponds to a function f_B of the type described in Figure 3 of the main paper with parameters $a = c = 0.25$. Clearly, this rule implies that q_i is increasing in t_i for $i = 1, 2$. By Lemma 3 we therefore have an incentive compatible decision rule if and only if the interim probabilities of the preferred alternatives satisfy:

$$p_i(t_i) = \begin{cases} p_i(1) \quad \frac{3}{4} - t_i\left(t_i - \frac{1}{2}\right) - \int_{t_i}^{1} \left(s_i - \frac{1}{4}\right) ds_i & \\ = p_i(1) + \frac{1}{2} - \frac{1}{2}\left(t_i\right)^2 & \text{if } t_i \geq 0.25 \\ p_i(1) + \frac{1}{2} - \frac{1}{2} \cdot \left(\frac{1}{4}\right)^2 & \\ = p_i(1) + \frac{15}{32} & \text{if } t_i < 0.25 \end{cases}$$

We can achieve incentive compatibility by defining f_A by:

$$f_A(t_1,t_2)=\begin{cases}0.5 & \text{if } t_1,t_2<0.25\\ \frac{1}{8}-2\left(t_1\right)^2+2t_1 & \text{if } t_1\geq 0.25, t_2<0.25\\ \frac{7}{8}+2\left(t_2\right)^2-2t_2 & \text{if } t_1<0.25, t_2\geq 0.25\\ 0.5 & \text{if } t_1,t_2\geq 0.25, t_1+t_2<1.25\\ 0 & \text{if } t_1+t_2\geq 1.25\end{cases}$$

and defining f_C analogously. It is trivial to verify that $f_A(t_1,t_2)\in[0,1]$, $f_B(t_1,t_2)\in[0,1]$ and $f_A(t_1,t_2)+f_B(t_1,t_2)+f_C(t_1,t_2)=1$ for all $(t_1,t_2)\in[0,1]^2$. It remains to check the incentive compatibility constraint. Note first that

$$p_i(1)=\frac{1}{32}$$

for $i=1,2$. Therefore, for $t_i\geq 0.25$, we need to check that:

$$p_i(t_i)=\frac{17}{32}-0.5\left(t_i\right)^2.$$

We calculate:

$$\begin{aligned}p_i(t_i) &= \frac{1}{4}\left(\frac{1}{8}-2\left(t_i\right)^2+2t_i\right)+\left(1.25-t_i-0.25\right)\frac{1}{2}\\ &= \frac{17}{32}-\frac{1}{2}\left(t_i\right)^2.\end{aligned}$$

For $t_i<0.25$ we need to check:

$$p_i(t_i)=\frac{1}{2}.$$

We calculate:

$$\begin{aligned}p_i(t_i) &= \frac{1}{4}\cdot\frac{1}{2}+\int_{\frac{1}{4}}^{1}\frac{7}{8}+2\left(t_j\right)2-2t_j dt_j\\ &= \frac{1}{2}.\end{aligned}$$