Universität Bielefeld/IMW

Working Papers
Institute of Mathematical Economics

Arbeiten aus dem
Institut für Mathematische Wirtschaftsforschung

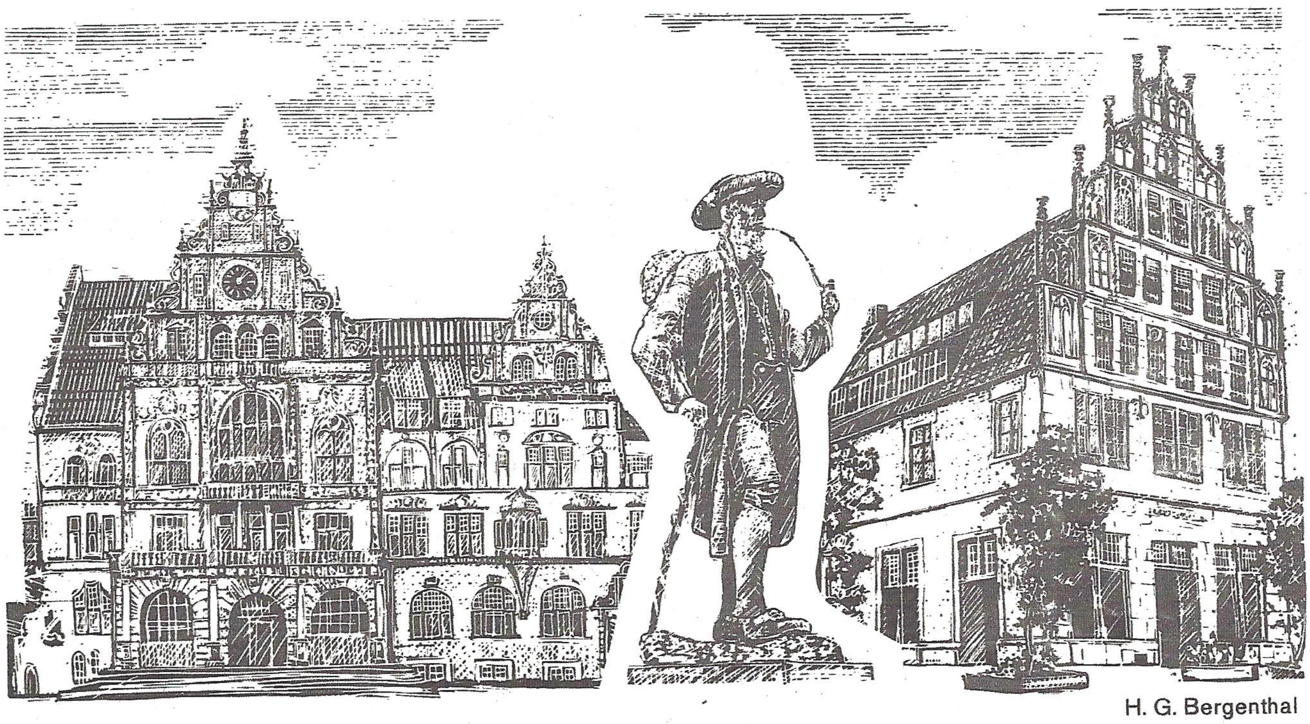
Nr. 187

A Bargaining Problem With Transferable Utility

by

Frank Weidner

August 1990



H. G. Bergenthal

Institut für Mathematische Wirtschaftsforschung
an der
Universität Bielefeld
Adresse/Address:
Universitätsstraße
4800 Bielefeld 1
Bundesrepublik Deutschland
Federal Republic of Germany

1. Introduction

In 1979 Myerson [3] defines a general framework to model a bargaining situation under uncertainty. According to the revelation principle all possible outcomes are given by the incentive compatible mechanisms.

In this paper I want to analyse a bargaining problem with transferable utility where the utility of the threatpoint is private knowledge. Myerson and Satterthwaite [5] treated a related problem where the amount of utility that can be divided is not common knowledge.

I will formulate some conditions which characterize the incentive compatible mechanisms. In the second part of the paper the assumption that the beliefs of the agents are common knowledge is dropped and I look at the safe (Myerson [4], or ex post–incentive compatible, Holmström & Myerson [1]) mechanisms. These mechanisms will be completely characterized.

In the last part of the paper I analyse the ex–post efficient mechanisms and show that they form a convex set. The boundary of this set is given too.

2. The Bargaining Problem

In this paper I want to discuss the following problem.

Two subjects, called agent 1 and agent 2, may agree to cooperate and produce a commodity together. They know the costs and the reward for the product. The agents bargain about the profit.

But there is a difficulty. Before the production begins, both agents explore other production alternatives which they can realize without their partner. (You may think of other external subjects which offer a joint production too). The agents have to decide whether to cooperate or to choose an alternative.

They want to fix a contract now which determines what should happen later at the beginning of the production. Shall the joint production start – in view of the alternatives – and how shall the profit be divided. Therefore the agents make a contract which determines what to do later, whatever the production alternatives are.

I want to model this situation in the framework presented by Myerson [3] .

At the beginning of the production the agents have decided whether to produce or not and how to share the profit. We assume that the agents are able to perform joint lotteries and that they are risk neutral. Then we may restrict ourselves to three decisions:

d_0 – they do not produce together;

d_1 – production takes place and agent 1 gets the whole profit;

d_2 – production takes place and agent 2 gets the whole profit.

Other divisions of the profit are modelled as lotteries between d_1 and d_2. These three decisions form the decision set D. d_0 is called the disagreement – (or threat–) point.

The agents are assumed to have a von-Neumann–Morgenstern utility function. We may define the utility of decision d_i, $i = 1,2$, to be 1 for agent i and 0 for the other agent.

The value of decision d_0 depends on the alternatives the agents have.

The utility level that an agent may reach with the alternatives is called the type of the agent. If the agent is of type $t \in \mathbb{R}$, then the utility of d_0 is t.

At last I assume that both agents have beliefs about the utility level the partner will reach. This belief remains unchanged during the whole process and is independent of the own type. The belief is a probability distribution on the typeset of the partner.

Now let me present the formal model.

The agents are characterized by their type sets T_1 (resp. T_2) and their beliefs p_1 (resp. p_2). We may restrict to type sets which are subsets of $[0,1]$. According to Myerson [3] we get a bargaining problem

$$\Gamma(T_1,T_2,p_1,p_2) := (D,d_0,T_1,T_2,v_1,v_2,p_1,p_2)$$

for every $T_1 \subset [0,1]$, $T_2 \subset [0,1]$, $p_1 \in \Delta(T_2)$, $p_2 \in \Delta(T_1)$.

The components are:

$D = \{d_0,d_1,d_2\}$	the decision set
T_1,T_2	the type sets
p_1,p_2	the beliefs
v_1,v_2	the utility functions on D

where for $i = 1,2$ $v_i : D \times T_1 \to \mathbb{R}$ is defined by

$$v_i(d|t) = \begin{cases} t & \text{if } d = d_0 \\ 1 & \text{if } d = d_i \\ 0 & \text{if } d = d_{-i} \end{cases} .$$

I used the shortening $-i$ for $3-i$.

All the components of the bargaining problem are common knowledge. The product $T_1 \times T_2$ of the type sets is called the state space.

The contract the agent should agree on may be type dependent. After the agents announced their types the contract determines a lottery over D.

On one hand a lottery between d_1 and d_2 is needed to model the division of the profit and on the other hand a lottery including d_0 is necessary to model the possibility of non–cooperation. Such a contract is called a mechanism.

Let $T_1 \times T_2$ be a state space. A mechanism μ (for $T_1 \times T_2$) is a function

$$\mu : D \times T_1 \times T_2 \to [0,1],$$

so that

$$\sum_{d \in D} \mu(d,\cdot,\cdot) = 1.$$

We write $\mu(d|a,b)$ instead of $\mu(d,a,b)$ for $(d,a,b) \in D \times T_1 \times T_2$.
Let $\mathcal{M}(T_1 \times T_2)$ denote the set of all mechanisms for the state space $T_1 \times T_2$.

We often write $\mathcal{M}$ instead of $\mathcal{M}(T_1 \times T_2)$ if no confusion is to be expected.

Agent i, i = 1,2, can compute the expected utility $u_i(\mu | a,b)$ of the mechanism μ in the state $(a,b) \in T_1 \times T_2$.

We define the functions

$$u_1 : \mathcal{M} \times T_1 \times T_2 \to \mathbb{R} \quad \text{and}$$

$$u_2 : \mathcal{M} \times T_1 \times T_2 \to \mathbb{R} \quad \text{by}$$

$$u_1(\mu | a,b) := \mu(d_1 | a,b) + a\, \mu(d_0 | a,b),$$

$$u_2(\mu | a,b) := \mu(d_2 | a,b) + b\, \mu(d_0 | a,b),$$

$$\forall \ \mu \in \mathcal{M},\ (a,b) \in T_1 \times T_2.$$

3. The utility of a mechanism

Let the bargaining problem $\Gamma(T_1,T_2,p_1,p_2)$ be fixed for the rest of this chapter.

After the agents have agreed on a mechanism and have explored their alternatives to determine their type, the agents have to announce their type and then the lottery takes place. The utility level an agent reached with the alternatives is not known by the other agent and (I assume that it) cannot be proven to him. Knowing this, the agent may have an incentive to announce not his true type in order to reach a lottery that is better for him than the lottery truthtelling will induce.

I also assume that the types are announced simultaneously. Therefore we have to compute the expected utilities according to the beliefs if we want to decide whether the agents have an incentive to deviate from truthtelling.

If agent 1 is type $a \in T_1$ and agent 2 is type $b \in T_2$ and agent 1 announces type $a' \in T_1$, while agent 2 announces his true type, then agent 1 gets the utility

$$u_1^*(\mu,a'\,|\,a,b) := \mu(d_1\,|\,a',b) + a\,\mu(d_0\,|\,a',b).$$

Therefore we define the function

$$u_1^* : \mathcal{M} \times T_1 \times T_1 \times T_2 \rightarrow \mathbb{R} \quad \text{by}$$

$$u_1^*(\mu,a'\,|\,a,b) := \mu(d_1\,|\,a',b) + a\mu(d_0\,|\,a',b)$$

$$\forall\, \mu \in \mathcal{M},\ a',a \in T_1,\ b \in T_2.$$

Analogously we define the function

$$u_2^*: \mathcal{M} \times T_1 \times T_2 \times T_2 \rightarrow \mathbb{R} \quad \text{by}$$

$$u_2^*(\mu,b'|a,b) := \mu(d_2|a,b') + b\,\mu(d_0|a,b')$$

$$\forall\, \mu \in \mathcal{M},\ b,b' \in T_2,\ a \in T_1.$$

In the moment agent 1 announces his type he doesn't know the type of the other agent. He may compute the expected probabilities of the decisions implemented by the mechanism if he announces $a \in T_1$ and agent 2 announces his true type.

These expected probabilities are given by

$$\bar{\mu}^1(d|a) := \sum_{b \in T_2} \mu(d|a,b)\, p_1(b) \quad \forall\, d \in D,\ a \in T_1.$$

Analogously we define

$$\bar{\mu}^2(d|b) := \sum_{a \in T_1} \mu(d|a,b)\, p_2(a) \quad \forall\, d \in D,\ b \in T_2.$$

The expected probabilities depend on the beliefs. This is not indicated by the notation.

Agent 1 computes the expected utility of the mechanism μ if his type is $a \in T_1$ and he announces $a' \in T_1$ while agent 2 always announces his true type.

This utility $U_1^*(\mu,a'|a)$ is given by the function

$$U_1^*: \mathcal{M} \times T_1 \times T_1 \rightarrow \mathbb{R}$$

defined by

$$U_1^*(\mu,a'\,|\,a) := \bar{\mu}^1(d_1\,|\,a') + a\,\bar{\mu}^1(d_0\,|\,a').$$

If agent 1 also tells the truth he will receive the expected utility given by the function $U_1 : \mathscr{M} \times T_1 \to \mathbb{R}$ which is defined by

$$U_1(\mu\,|\,a) := U_1^*(\mu,a\,|\,a) \qquad \forall\, a \in T_1.$$

In the same way we define U_2^* and U_2.

When the agents contract a mechanism they surely want the opposite to announce his true type later. A contract is called incentive compatible if no agent of any type expects (according to his beliefs) to gain from lying if the other agent always announces his true type.

Formally:

Definition: Let $\Gamma(T_1,T_2,p_1,p_2)$ be a bargaining problem.

A mechanism $\mu \in \mathscr{M}(T_1 \times T_2)$ is called incentive compatible iff

$$U_i^*(\mu,s\,|\,t) \leq U_i(\mu\,|\,t) \quad \forall\, s,t \in T_i,\ i=1,2.$$

To prevent the agents from retreating from the contract when they know the alternatives we demand that the expected reward from the contract is not lower than the profit of the alternative. Those contracts are called individually rational.

Formally:

Definition: Let $\Gamma(T_1,T_2,p_1,p_2)$ be a bargaining problem.

A mechanism $\mu \in \mathscr{M}(T_1 \times T_2)$ is called individually rational iff

$$U_i(\mu\,|\,t) \geq t \qquad \forall\, t \in T_i,\ i=1,2.$$

4. Some inequalities

In this chapter I will show some inequalities that will be useful later.

In the following let $\Gamma(T_1,T_2,p_1,p_2)$ be a fixed bargaining problem.

Let $i \in \{1,2\}$, $s,t \in T_i$ and $\mu \in \mathcal{M}(T_1 \times T_2)$.
The following three inequalities are equivalent.

1) $U_i^*(\mu,s|t) \leq U_i(\mu|t)$,

2) $U_i(\mu|t) - U_i(\mu|s) \geq (t-s)\,\bar{\mu}^i(d_0|s)$,

3) $\bar{\mu}^i(d_i|s) - \bar{\mu}^i(d_i|t) \leq t(\bar{\mu}^i(d_0|t) - \bar{\mu}^i(d_0|s))$.

To see the equivalence of 1) and 2) look at the equation

$$4)\quad U_i^*(\mu,s|t) = \bar{\mu}^i(d_i|s) + t\,\bar{\mu}^i(d_0|s)$$

$$= \bar{\mu}^i(d_i|s) + s\,\bar{\mu}^i(d_0|s) + (t-s)\,\bar{\mu}^i(d_0|s)$$

$$= U_i(\mu|s) + (t-s)\,\bar{\mu}^i(d_0|s).$$

Therefore one gets

$$U_i(\mu|t) - U_i^*(\mu,s|t) = U_i(\mu|t) - U_i(\mu|s) - (t-s)\,\bar{\mu}^i(d_0|s).$$

The left side is greater than zero iff 2) holds. This proves the equivalence of 1) and 2).

By definition

$$U_i^*(\mu,s|t) \leq U_i(\mu|t)$$

is equivalent to

$$\bar{\mu}^i(d_i|s) + t\,\bar{\mu}^i(d_0|s) \le \bar{\mu}^i(d_i|t) + t\,\bar{\mu}^i(d_0|t).$$

This is equivalent to 3).

If we change the roles of s and t we see that the following three pairs of inequalities are equivalent:

1') $U_i^*(\mu,s|t) \le U_i(\mu|t)$ and $U_i^*(\mu,t|s) \le U_i(\mu|s)$,

2') $(t-s)\,\bar{\mu}^i(d_0|s) \le U_i(\mu|t) - U_i(\mu|s) \le (t-s)\,\bar{\mu}^i(d_0|t)$,

3') $s(\bar{\mu}^i(d_0|t) - \bar{\mu}^i(d_0|s)) \le \bar{\mu}^i(d_i|s) - \bar{\mu}^i(d_i|t) \le t(\bar{\mu}^i(d_0|t) - \bar{\mu}^i(d_0|s))$.

If $s < t$ and 1') holds, then we may conclude:

$$U_i(\mu|t) \ge U_i(\mu|s),$$

$$\bar{\mu}^i(d_0|t) \ge \bar{\mu}^i(d_0|s),$$

$$\bar{\mu}^i(d_i|t) \le \bar{\mu}^i(d_i|s),$$

$$\bar{\mu}^i(d_{-i}|t) \le \bar{\mu}^i(d_i|s).$$

The first inequality follows from 2') (the left part) and the fact that $\bar{\mu}^i(d_0|\cdot) \ge 0$. Comparison of the left and the right side of 2') leads to the second inequality. If one uses the second inequality and 3') one sees that the third inequality holds.

Finally we see that

$$\bar{\mu}^i(d_{-i}|s) - \bar{\mu}^i(d_{-i}|t) = \bar{\mu}^i(d_0|t) + \bar{\mu}^i(d_i|t) - (\bar{\mu}^i(d_0|s) + \bar{\mu}^i(d_i|s))$$

$$= \bar{\mu}^i(d_0|t) - \bar{\mu}^i(d_o|s) - (\bar{\mu}^i(d_i|s) - \bar{\mu}^i(d_i|t))$$

$$\geq \bar{\mu}^i(d_0|t) - \bar{\mu}^i(d_0|s) - s(\bar{\mu}^i(d_i|s) - \bar{\mu}^i(d_i|t))$$

$$\geq 0$$

because $\sum_{k=0}^{2} \bar{\mu}^i(d_k|\cdot) = 1$, $s \leq 1$ together with the third inequality, and 3').

Now we can prove the following

<u>Lemma 1:</u> If $i \in \{1,2\}$, $t_1, t_2, t_3 \in T_i$, $t_1 < t_2 < t_3$,

$$U_i^*(\mu, t_1|t_2) \leq U_i(\mu|t_2),\ U_i^*(\mu, t_2|t_1) \leq U_i(\mu|t_1),$$

$$U_i^*(\mu, t_2|t_3) \leq U_i(\mu|t_3) \quad \text{and} \quad U_i^*(\mu, t_3|t_2) \leq U_i(\mu|t_2),$$

then

$$U_i^*(\mu, t_1|t_3) \leq U_i(\mu|t_3) \quad \text{and} \quad U_i^*(\mu, t_3|t_1) \leq U_i(\mu|t_1).$$

<u>Proof:</u> The assumptions imply

$$\bar{\mu}^i(d_0|t_1) \leq \bar{\mu}^i(d_0|t_2) \leq \bar{\mu}^i(d_0|t_3).$$

We know that

$$U_i^*(\mu, t_1|t_2) \leq U_i(\mu|t_2)$$

is equivalent to

$$\bar{\mu}^i(d_i|t_1) - \bar{\mu}^i(d_i|t_2) \leq t_2(\bar{\mu}^i(d_0|t_2) - \bar{\mu}^i(d_0|t_1)).$$

Because $t_3 > t_2$ and $\bar{\mu}^i(d_0|t_2) - \bar{\mu}^i(d_0|t_1) \geq 0$ we conclude

$$\bar{\mu}^i(d_i|t_1) - \bar{\mu}^i(d_i|t_2) \leq t_3(\bar{\mu}^i(d_0|t_2) - \bar{\mu}^i(d_0|t_1)).$$

Rearranging the terms yields

$$U_i^*(\mu,t_1|t_3) \leq U_i^*(\mu,t_2|t_3).$$

The right side does not exceed $U_i(\mu|t_3)$. This proves the first inequality in the lemma. The second may be proven in the same manner.

#

This lemma shows that "local incentive compatibility" implies "global incentive compatibility" in the following sense.

Let T_1 and T_2 be finite sets. If one wants to show that a mechanism is incentive compatible it suffices to prove that no type has an incentive to pretend to be one of his two neighbour types.

$T_1 = T_2 = [0,1]$ and $\mathscr{V}$ is an open covering of $[0,1]$, then we see that $\mu \in \mathscr{M}(T_1 \times T_2)$ is incentive compatible iff for all $V \in \mathscr{V}$, $i = 1,2$, it holds

$$U_i^*(\mu,t'|t) \leq U_i(\mu|t) \qquad \forall\, t,t' \in V.$$

This localization property will appear again when we characterize the incentive compatible mechanisms.

5. Incentive compatible mechanisms

For the rest of this paper we restrict ourselves to the case

$$T_1 = T_2 = [0,1].$$

Let the beliefs p_1, p_2 be fixed in this chapter. Let $\mu \in \mathcal{M}$ be incentive compatible and $i \in \{1,2\}$.

Because of equation 2') we get

$$|U_i(\mu|t) - U_i(\mu|s)| \leq |t-s| \max_{s,t} \{\bar{\mu}^i(d_0|t), \bar{\mu}^i(d_0|s)\} \leq |t-s|, \quad \forall\, t,s \in T_i.$$

Therefore the function $U_i(\mu|\cdot) : [0,1] \to \mathbb{R}$ is Lipschitz continuous. This implies that $U_i(\mu|\cdot)$ is differentiable almost everywhere and the derivation f_i is integrable and

$$U_i(\mu|t) - U_i(\mu|0) = \int_0^t f_i(s)\, ds \qquad \forall\, t \in [0,1].$$

(see Natanson [6], chapter IX, §§1,2.)

The function $\bar{\mu}^i(d_0|\cdot) : [0,1] \to [0,1]$ is isotone and therefore measurable and continuous almost everywhere. (see Natanson [6], chapter VIII, §1.)

Equation 2') implies $f_i = \bar{\mu}^i(d_0|\cdot)$ almost everywhere.

<u>Lemma 2:</u> Let $\mu \in \mathcal{M}$.

μ is incentive compatible iff for $i = 1,2$

1) $\bar{\mu}^i(d_0|\cdot)$ is isotone and

2) $U_i(\mu|t) - U_i(\mu|0) = \int_0^t \bar{\mu}^i(d_0|s)\, ds \qquad \forall\, t \in T_i.$

Remark: 1) Both conditions are local properties. Equation 2) states that $\bar{\mu}^i(d_0|\cdot)$ is the derivation of $U_i(\mu|\cdot)$.

2) Equation 2) is equivalent to

$$\bar{\mu}^i(d_i|t) - \bar{\mu}^i(d_i|0) = \int_0^t \bar{\mu}^i(d_0|s)\,ds - t\,\bar{\mu}^i(d_0|t) \qquad \forall\, t \in T_i.$$

<u>Proof:</u> The only if part was proven above.

Now let $\mu \in \mathcal{M}$ fulfill 1) and 2).

Let $t_1, t_2 \in T_i$, $i \in \{1,2\}$.

We have to prove that

$$U_i^*(\mu, t_2|t_1) \leq U_i(\mu|t_1).$$

First case: Let $t_1 < t_2$.

$$U_i^*(\mu, t_2|t_1) = U_i(\mu|t_2) + (t_1 - t_2)\,\bar{\mu}^i(d_0|t_2) \qquad \text{(see equ. 4) of chapter 4)}$$

$$= U_i(\mu|0) + \int_0^{t_2} \bar{\mu}^i(d_0|s)\,ds - (t_2 - t_1)\,\bar{\mu}^i(d_0|t_2)$$

$$= U_i(\mu|0) + \int_0^{t_1} \bar{\mu}^i(d_0|s)\,ds + \int_{t_1}^{t_2} \bar{\mu}^i(d_0|s)\,ds - (t_2 - t_1)\,\bar{\mu}^i(d_0|t_2)$$

$$\leq U_i(\mu|0) + \int_0^{t_1} \bar{\mu}^i(d_0|s)\,ds$$

$$= U_i(\mu|t_1)$$

because $\bar{\mu}^i(d_0|t_2) \geq \bar{\mu}^i(d_0|s) \quad \forall\, s \in [t_1, t_2]$.

Second case: Let $t_2 < t_1$.

$$U_i^*(\mu,t_2|t_1) = U_i(\mu|t_2) + (t_1 - t_2)\,\bar{\mu}^i(d_0|t_2)$$

$$= U_i(\mu|0) + \int_0^{t_2} \bar{\mu}^i(d_0|s)\,ds + (t_1 - t_2)\,\bar{\mu}^i(d_0|t_2)$$

$$= U_i(\mu|0) + \int_0^{t_1} \bar{\mu}^i(d_0|s)\,ds - \int_{t_2}^{t_1} \bar{\mu}^i(d_0|s)\,ds + (t_1 - t_2)\,\bar{\mu}^i(d_0|t_2)$$

$$\leq U_i(\mu|0) + \int_0^{t_1} \bar{\mu}^i(d_0|s)\,ds$$

$$= U_i(\mu|t_1)$$

because $\bar{\mu}^i(d_0|t_2) \leq \bar{\mu}^i(d_0|s) \quad \forall\, s \in [t_2,t_1]$.

#

Now let $\mu \in \mathcal{M}$ be incentive compatible and individually rational. Then we must have

$$U_i(\mu|1) \geq 1 \qquad \text{for } i = 1,2.$$

Obviously we may conclude

$$U_i(\mu|1) = 1 \qquad \text{for } i = 1,2.$$

By equation 2) of the lemma this is equivalent to

$$\bar{\mu}^i(d_i|0) = U_i(\mu|0) = 1 - \int_0^1 \bar{\mu}^i(d_0|t)\,dt.$$

<u>Lemma 3:</u> If $\mu \in \mathcal{M}$ is incentive compatible, then μ is individually rational iff

$$U_i(\mu|1) = 1 \qquad \text{for } i = 1,2.$$

<u>Proof:</u> Let $i \in \{1,2\}$.

Of course the condition $U_i(\mu|1) = 1$ is necessary for μ to be individually rational. If $U_i(\mu|1) = 1$ then

$$\bar{\mu}^i(d_i|1) + \bar{\mu}^i(d_0|1) = 1.$$

Therefore we get for all $t \in T_i = [0,1]$

$$\begin{aligned} U_i(\mu|t) \geq U_i^*(\mu,1|t) \quad &= \bar{\mu}^i(d_i|1) + t\,\bar{\mu}^i(d_0|1) \\ &\geq t(\bar{\mu}^i(d_i|1) + \bar{\mu}^i(d_0|1)) \\ &= t. \end{aligned}$$

#

<u>Corollary:</u> If $\mu \in \mathcal{M}$ is incentive compatible then μ is individually rational iff

$$\bar{\mu}^i(d_i|0) = 1 - \int_0^1 \bar{\mu}^i(d_0|s)\,ds \qquad \text{for } i = 1,2.$$

6. Some problems with the model

The model proceeds from the viewpoint that all the components of the bargaining problem are common knowledge. If your are careful you take all possible reservations levels into account and you have to choose $T_1=T_2=[0,1]$. Two agents may agree to this consideration if they talk about their bargaining problem. (The measurement of utility is not a critical point as long as they agree to be risk neutral because the conceptions are invariant under linear transformations of scale.) But there is no incentive for the agents to reveal their true beliefs. If the agents don't trust the opposite they cannot define the set of possible (= incentive compatible) contracts. How shall the agents find a way out. One way to deal with this difficulty may be to introduce the universal belief space of Mertens and Zamir [2]. I don't believe that two persons (except two mathematicians) will contract a mechanism depending on this big space. Perhaps it is possible to make an approximation with finite spaces. An easier way out is to deal with safe mechanisms, which are those mechanisms that are incentive compatible with regard to every belief. Then it is necessary to claim that the mechanism is ex post individually rational; that is: the mechanism gives a higher expected utility than the alternative for every type of every agent no matter what the beliefs are.

<u>Definition:</u> Let T_1, T_2 be type sets.

A mechanism $\mu \in \mathcal{M}(T_1 \times T_2)$ is called safe iff

$$u_1^*(\mu, a' \mid a,b) \leq u_1(\mu \mid a,b) \quad \text{and}$$

$$u_2^*(\mu, b' \mid a,b) \leq u_2(\mu \mid a,b)$$

for all $a, a' \in T_1$, $b, b' \in T_2$.

If the mechanism μ is safe then μ is incentive compatible with respect to every belief. On the other hand: If μ is incentive compatible with respect to every belief, then μ has to be incentive compatible if agent 1 is sure that agent 2 is type $b \in T_2$ and agent 2 is sure that agent 1 is type $a \in T_1$. This condition must be true for all $a \in T_1$, $b \in T_2$. Therefore μ is safe.

A mechanism μ is called ex post rational iff μ is individually rational with respect to every belief or equivalently.

<u>Definition:</u> A mechanism $\mu \in \mathcal{M}(T_1 \times T_2)$ is called ex post (individually) rational iff

$$u_1(\mu | a,b) \geq a \text{ and } u_2(\mu | a,b) \geq b \text{ for all } (a,b) \in T_1 \times T_2.$$

The proof of this equivalence proceeds in the same way as above.

If $\mu \in \mathcal{M}(T_1 \times T_2)$ is ex post rational, then μ implements d_0 with probability 1 whenever the agents announce two types with a sum greater than 1:
If μ is ex post rational then $\mu(d_0 | a,b) = 1$ for all $(a,b) \in T_1 \times T_2$ with $a + b > 1$.

<u>Definition:</u> A mechanism $\mu \in \mathcal{M}(T_1 \times T_2)$ is called (ex post) feasible iff μ is safe and ex post rational.

The set of all feasible mechanisms is denoted by $\mathcal{B}(T_1 \times T_2)$.

The considerations above show that it is possible to carry over the results we got for incentive compatible and individually rational mechanism to safe and ex post rational mechanisms.

We restrict ourselves to the case $T_1 = T_2 = [0,1]$.

Let $\mu \in \mathcal{M}$.

Fact I $\quad \mu$ is safe iff

1) $\mu(d_0|\cdot,b)$ and $\mu(d_0|a,\cdot)$

are isotone for all $a,b \in [0,1]$,

and

2) $u_1(\mu|a,b) - u_1(\mu|0,b) = \int_0^a \mu(d_0|t,b)\,dt \qquad \forall\, a,b \in [0,1]$,

and

3) $u_2(\mu|a,b) - u_1(\mu|a,0) = \int_0^b \mu(d_0|a,t)\,dt \qquad \forall\, a,b \in [0,1]$.

The equation 2) is equivalent to

2') $\mu(d_1|a,b) - \mu(d_1|0,b) = \int_0^a \mu(d_0|t,b)dt - a\,\mu(d_0|a,b) \quad \forall\, a,b \in [0,1]$,

and equation 3) is equivalent to

3') $\mu(d_2|a,b) - \mu(d_2|a,0) = \int_0^b \mu(d_0|a,t)dt - b\,\mu(d_0|a,b) \quad \forall\, a,b \in [0,1]$.

Fact II $\quad$ If $\mu \in \mathcal{M}$ is safe then it holds for all $a,b \in [0,1]$

a) $\mu(d_0|\cdot,b)$ and $\mu(d_0|a,\cdot)$ are isotone,

b) $\mu(d_1|\cdot,b)$, $\mu(d_2|\cdot,b)$, $\mu(d_1|a,\cdot)$ and $\mu(d_2|a,\cdot)$ are antitone,

c) $u_1(\mu|\cdot,b)$ and $u_2(\mu|a,\cdot)$ are Lipschitz continuous and isotone.

Fact III $\quad$ If $\mu \in \mathcal{M}$ is safe then μ is ex post rational iff

$u_1(\mu|1,b) = 1$ and $u_2(\mu|a,1) = 1 \quad$ for all $a,b \in [0,1]$,

These two equations are equivalent to

$\mu(d_1|0,b) = 1 - \int_0^1 \mu(d_0|t,b)\,dt \qquad \forall\, b \in [0,1]$

and

$\mu(d_2|a,0) = 1 - \int_0^1 \mu(d_0|a,t)\,dt \qquad \forall\, a \in [0,1]$.

7. Characterization of safe mechanisms

The next two theorems completely characterize the feasible mechanisms. To shorten the notation we define

$$S := \{(a,b) \in [0,1]^2 \mid a + b < 1\},$$

$$\bar{S} := \{(a,b) \in [0,1]^2 \mid a + b \leq 1\}.$$

Lemma 4: If $\mu \in \mathcal{M}$ is safe then

$$\mu(d \mid a,b) = \mu(d \mid a,0) + \mu(d \mid 0,b) - \mu(d \mid 0,0) \quad \forall\, (a,b) \in S;\ d \in D.$$

Proof: Because μ is safe fact I implies

$$\mu(d_1 \mid a,b) = \mu(d_1 \mid 0,b) + \int_0^a \mu(d_0 \mid t,b)\, dt - a\, \mu(d_0 \mid a,b),\ \forall\, (a,b) \in [0,1]^2$$

and

$$\mu(d_2 \mid a,b) = \mu(d_2 \mid a,0) + \int_0^b \mu(d_0 \mid a,t)\, dt - b\, \mu(d_0 \mid a,b),\ \forall\, (a,b) \in [0,1]^2.$$

Using the equation $\sum_{d \in D} \mu(d \mid \cdot,\cdot) = 1$ we get for all $(a,b) \in [0,1]^2$:

$$\mu(d_0 \mid a,b) = 1 - \mu(d_1 \mid a,b) - \mu(d_2 \mid a,b)$$

$$= 1 - \mu(d_1 \mid 0,b) - \mu(d_2 \mid a,0) + (a + b)\, \mu(d_0 \mid a,b)$$

$$- \int_0^a \mu(d_0 \mid t,b)\, dt - \int_0^b \mu(d_0 \mid a,t)\, dt.$$

This is equivalent to the equation

$$(1 - a - b)\, \mu(d_0 \mid a,b) + \int_0^a \mu(d_0 \mid t,b)\, dt + \int_0^b \mu(d_0 \mid a,t)\, dt$$

$$= 1 - \mu(d_1 \mid 0,b) - \mu(d_2 \mid a,0) \quad \forall\, (a,b) \in [0,1]^2.$$

Now we restrict the function $\mu(d_0|\cdot,\cdot)$ to S.

Let $g : S \rightarrow \mathbb{R}$ be defined by

$$g(a,b) = \mu(d_0|a,0) + \mu(d_0|0,b) - \mu(d_0|0,0), \quad \forall (a,b) \in S.$$

To see that this function is a solution of the integral equation we use conditions 2) and 3) of fact I to transform the right side.

$$\mu(d_1|0,b) = 1 - \mu(d_0|0,b) - \mu(d_2|0,b)$$

$$= 1 - \mu(d_0|0,b) - \mu(d_2|0,0) - \int_0^b \mu(d_0|0,t)\,dt + b\,\mu(d_0|0,b)$$

$$= 1 - \mu(d_2|0,0) - (1-b)\,\mu(d_0|0,b) - \int_0^b \mu(d_0|0,t)\,dt.$$

Analogously we get

$$\mu(d_2|a,0) = 1 - \mu(d_1|0,0) - (1-a)\,\mu(d_0|a,0) - \int_0^a \mu(d_0|t,0)\,dt.$$

Therefore we write

$$1 - \mu(d_1|0,b) - \mu(d_2|a,0)$$

$$= -1 + \mu(d_2|0,0) + \mu(d_1|0,0) + (1-b)\,\mu(d_0|0,b)$$

$$+ (1-a)\,\mu(d_0|a,0) + \int_0^a \mu(d_0|t,0)\,dt + \int_0^b \mu(d_0|0,t)\,dt$$

$$= -\mu(d_0|0,0) + (1-b)\,\mu(d_0|0,b) + (1-a)\,\mu(d_0|a,0)$$

$$+ \int_0^a \mu(d_0|t,0)\,dt + \int_0^b \mu(d_0|0,t)\,dt.$$

If we insert g into the left side of the integral equation we get for all $(a,b) \in S$:

$$(1-a-b)\, g(a,b) + \int_0^a g(t,b)\, dt + \int_0^b g(a,b)\, dt$$

$$= (1-a-b)\, (\mu(d_0|a,0) + \mu(d_0|0,b) - \mu(d_0|0,0))$$

$$+ \int_0^a \mu(d_0|t,0)\, dt + a\, \mu(d_0|0,b) - a\, \mu(d_0|0,0)$$

$$+ \int_0^b \mu(d_0|0,t)\, dt + b\, \mu(d_0|a,0) - b\, \mu(d_0|0,0)$$

$$= -\mu(d_0|0,0) + (1-b)\, \mu(d_0|0,b) + (1-a)\, \mu(d_0|a,0)$$

$$+ \int_0^a \mu(d_0|t,0)\, dt + \int_0^b \mu(d_0|0,t)\, dt.$$

This is exactly the right side of the integral equation.

Let $f : S \to \mathbb{R}$ be defined by

$$f(a,b) = \mu(d_0|a,b) - g(a,b) \qquad \forall\, (a,b) \in S.$$

$\mu(d_0|\cdot,\cdot)$ and g are solutions of the integral equation. Therefore f fulfills the equation

$$(1-a-b)\, f(a,b) + \int_0^a f(t,b)\, dt + \int_0^b f(a,t)\, dt = 0 \qquad \forall\, (a,b) \in S.$$

The following lemma will prove that f is the zero function.

Of course our f, as defined above, fulfills the assumption of the following lemma and the lemma implies that f is the zero function therefore. This shows that g and $\mu(d_0|\cdot,\cdot)$ coincide on S, which proves the lemma for $d = d_0$.

To prove the lemma for $d = d_1$ and $d = d_2$ we use equation 2') resp. 3') of fact I and the result for $d = d_0$.

If $(a,b) \in S$ then:

$$\mu(d_1 | a,b) = \int_0^a \mu(d_0 | t,b)\, dt - a\, \mu(d_0 | a,b) + \mu(d_1 | 0,b)$$

$$= \int_0^a \mu(d_0 | t,0)\, dt + a\, \mu(d_0 | 0,b) - a\, \mu(d_0 | 0,0)$$

$$- a\, \mu(d_0 | a,0) - a\, \mu(d_0 | 0,b) + a\, \mu(d_0 | 0,0) + \mu(d_1 | 0,b)$$

$$= \int_0^a \mu(d_0 | t,0)\, dt - a\, \mu(d_0 | a,0) + \mu(d_1 | 0,b)$$

$$= \mu(d_1 | a,0) - \mu(d_1 | 0,0) + \mu(d_1 | 0,b).$$

In the same manner one can prove the lemma for $d = d_2$.

#

<u>Lemma 5:</u> If the function $f : S \to \mathbb{R}$ fulfills the inequality

$$(1 - a - b)\, |f(a,b)| \leq \int_0^a |f(t,b)|\, dt + \int_0^b |f(a,t)|\, dt$$

for all $(a,b) \in S$, then $f = 0$.

<u>Proof:</u> For $r \geq 0$ define

$$S_r := \{(a,b) \in S \mid a + b \leq r\}$$

and

$$m_r := \sup \left\{ |f(a,b)| \;\middle|\; (a,b) \in S_r \right\}.$$

In a first step we will see that $m_r = 0$ for all $r < \frac{1}{2}$.

Let $0 \leq r < \frac{1}{2}$ and suppose $m_r > 0$.

If $0 < \epsilon < r$ then there exists $(a,b) \in S_r$, so that $m_r - |f(a,b)| < \epsilon$.

This implies

$$\frac{1}{2}(m_r - \epsilon) \leq (1 - a - b)(m_r - \epsilon)$$

$$\leq (1 - a - b)\,|f(a,b)|$$

$$\leq \int_0^a |f(t,b)|\,dt + \int_0^b |f(a,t)|\,dt$$

$$\leq (a + b)\,m_r$$

$$\leq r \cdot m_r.$$

Now we conclude

$$\left(\frac{1}{2} - r\right) m_r \leq \frac{1}{2}\epsilon.$$

Because ϵ was chosen arbitrary this implies that $m_r = 0$.

Let $r_0 := \sup\{r \geq 0 \mid m_r = 0\}$.

In the first step we saw that $r_0 \geq \frac{1}{2}$ and in the next step we will see that $r_0 = 1$, which is enough to close the proof.

Suppose $r_0 < 1$.

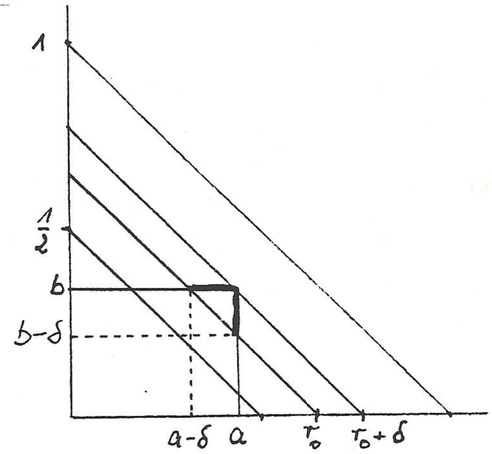
Then let $\delta := \dfrac{1 - r_0}{4}$.

By definition $|f(a,b)| = 0$ for all $(a,b) \in S$ with $a + b < r_0$.

Therefore

$$\int_0^a |f(t,b)|\,dt + \int_0^b |f(a,t)|\,dt \leq 2 \cdot \delta \cdot m_{r_0+\delta} \text{ for all } (a,b) \in S_{r_0+\delta}.$$

To see this look at the following picture.



Now let $0 < \epsilon < m_{r_0+\delta}$, then there exists $(a,b) \in S_{r_0+\delta}$ so that $|f(a,b)| \geq m_{r_0+\delta} - \epsilon$. We get with the help of the considerations above

$$(1 - r_0 - \delta)\,(m_{r_0+\delta} - \epsilon) \leq (1 - a - b)\ |f(a,b)|$$

$$\leq \int_0^a |f(t,b)|\ dt + \int_0^b |f(a,t)|\ dt$$

$$\leq 2\,\delta \cdot m_{r+\delta}.$$

This implies

$$(1 - r_0 - 3\delta)\ m_{r_0+\delta} \leq (1 - r_0 - \delta)\ \epsilon.$$

Because ϵ was chosen arbitrary and by definition of δ we conclude

$$m_{r_0+\delta} = 0.$$

This is a contradiction to the definition of r_0. Therefore $r_0 = 1$.

#

If we know that μ is a safe mechanism and we know the probabilities $\mu(d_0|\cdot,\cdot)|_{(T_1\times\{0\})\cup(\{0\}\times T_2)}$ then we can compute $\mu(d_0|\cdot,\cdot)|_S$ with the help of lemma. If in addition μ is ex post rational then we know

$$\mu(d_0|a,b)=1,\ \forall\ (a,b)\in[0,1]^2\backslash\bar{S}.$$

We can also compute $\mu(d_1|0,b)$ for all $b\in[0,1]$, and $\bar{\mu}(d_2|a,0)$ for all $a\in[0,1]$ with the help of fact III.

Because $\mu(\ \ |a,b)$ is a lottery for all $(a,b)\in[0,1]^2$ and because of lemma 4 we know $\mu(\ \ |a,b)$ for all $(a,b)\in[0,1]^2$ with the exception of those (a,b) with $a+b=1$.

For all $(a,b)\in[0,1]^2$ we have

$$\begin{aligned} & u_1(\mu|a,b)+u_2(\mu|a,b) \\ = \ & \mu(d_1|a,b)+\mu(d_2|a,b)+(a+b)\,\mu(d_0|a,b) \\ = \ & 1-(1-a-b)\,\mu(d_0|a,b), \end{aligned}$$

because $\sum_{d\in D}\mu(d|\cdot,\cdot)=1$.

Therefore $u_1(\mu|a,b)+u_2(\mu|a,b)\leq 1\ \ \forall\ (a,b)\in\bar{S}$.

If μ is ex post rational we conclude:

$$u_1(\mu|a,1-a)=a,\quad \forall\ a\in T_1,$$

$$u_2(\mu|1-b,b)=b,\quad \forall\ b\in T_2.$$

The utility that a safe and ex post rational mechanism gives to the different types of the agents is completely determined if μ is given on $(T_1 \times \{0\}) \cup (\{0\} \times T_2)$.

The following lemma sharpens this result.

Lemma 6: If the mechanism μ is safe and ex post rational then

$$\mu(d_0 | a,0) + \mu(d_0 | 0,1-a) = 1 + \mu(d_0 | 0,0)$$

$$\forall\, a \in C := \{t \in (0,1) \,|\, \mu(d_0 | \cdot ,0) \text{ is continuous in } t\}.$$

Remark:
The complement of the set C is a set with measure zero because $\mu(d_0 | \cdot ,0)$ is isotone.

Proof: We compute the utilities for the agent 1.

If $(a,b) \in S$ then

$$u_1(\mu | a,b) = \mu(d_1 | a,b) + a\, \mu(d_0 | a,b)$$

$$= \mu(d_1 | 0,b) + \int_0^a \mu(d_0 | t,b)\, dt \text{ (fact I)}$$

$$= 1 - \mu(d_2 | 0,b) - \mu(d_0 | 0,b) + \int_0^a \mu(d_0 | t,b)\, dt$$

$$= 1 - \mu(d_2 | 0,0) - \int_0^b \mu(d_0 | 0,t)\, dt + b\, \mu(d_0 | 0,b) + \int_0^a \mu(d_0 | t,0)\, dt$$

$$+ a\, \mu(d_0 | 0,b) - a\, \mu(d_0 | 0,0) - \mu(d_0 | 0,b) \quad \text{(lemma 4, fact I)}$$

$$= \mu(d_1 | 0,0) + \mu(d_0 | 0,0) - (1 - a - b)\, \mu(d_0 | 0,b) - a\, \mu(d_0 | 0,0)$$

$$+ \int_0^a \mu(d_0 | t,0)\, dt - \int_0^b \mu(d_0 | 0,t)\, dt.$$

$u_1(\mu|\cdot,b)$ is a continuous function for every $b \in [0,1]$. Therefore we get for every $a \in (0,1]$:

$$u_1(\mu|a,1-a) = \lim_{a' \nearrow a} u_1(\mu|a',1-a)$$

$$= \mu(d_1|0,0) + (1-a)\,\mu(d_0|0,0) + \int_0^a \mu(d_0|t,0)\,dt - \int_0^{1-a} \mu(d_0|0,t)\,dt$$

Because $u_1(\mu|a,1-a) = a$ we conclude

$$\int_0^{1-a} \mu(d_0|0,t)\,dt = \mu(d_1|0,0) + (1-a)\,\mu(d_0|0,0) + \int_0^a \mu(d_0|0,t)\,dt - a,$$

$$\forall\, a \in (0,1].$$

We get by differentiation

$$-\mu(d_0|0,1-a) = -\mu(d_0|0,0) + \mu(d_0|0,a) - 1 \text{ for almost every } a \in (0,1].$$

μ is isotone. Therefore this equation must hold for every $a \in C$.

#

The lemmas above determine some conditions that all safe and ex post rational mechanisms suffice. The following corollary shows that these conditions are sufficient to prove that a mechanism is feasible.

<u>Corollary:</u> A mechanism $\mu \in \mathcal{M}$ is feasible iff

1) $\mu(d_0|\cdot,b)$ and $\mu(d_0|a,\cdot)$ are isotone for all $a,b \in [0,1]$.

2) $\mu(d|a,b) = \mu(d|a,0) + \mu(d|0,b) - \mu(d|0,0), \quad \forall\,(a,b) \in S,\ d \in D;$

3) $\mu(d_1|a,0) = \mu(d_1|0,0) + \int_0^a \mu(d_0|t,0)\,dt - a\,\mu(d_0|a,0), \quad \forall a \in [0,1];$

4) $\mu(d_2|0,b) = \mu(d_2|0,0) + \int_0^b \mu(d_0|0,t)\,dt - b\,\mu(d_0|0,b), \quad \forall\, b \in [0,1];$

5) $\mu(d_0|a,0) + \mu(d_0|0,1-a) = 1 + \mu(d_0|0,0)$ for almost every $a \in [0,1]$;

6) $\mu(d_1|0,0) = 1 - \int_0^1 \mu(d_0|a,0)\, da$;

7) $u_1(\mu|a,b) = a,\ u_2(\mu|a,b) = b \qquad \forall\ (a,b) \in [0,1]^2 \setminus S.$

Remark:

Because $\sum_{d \in D} \mu(d|\cdot,\cdot) = 1$ we only need two of the equations under 2).

Of course there exists an equivalent characterization of the ex post rational mechanisms if you change the roles of agent 1 and 2 in condition 5) and 6).

Condition 7) implies $\mu(d_0|a,b) = 1$ for $(a,b) \in [0,1]^2$ with $a + b > 1$.

Proof: I want to use fact I to prove that the 7 conditions are sufficient to show that the mechanism is feasible.
The conditions 2) and 3) yield

$$(*) \qquad \mu(d_1|a,b) - \mu(d_1|0,b) = \mu(d_1|a,0) - \mu(d_1|0,0)$$

$$= \int_0^a \mu(d_0|t,0)\, dt - a\, \mu(d_0|a,0) = \int_0^a \mu(d_0|t,b)\, dt - a\, \mu(d_0|a,b)$$

$$\forall\ (a,b) \in S.$$

Therefore condition 2) of fact I is fulfilled for $(a,b) \in S$.

In the same way one can prove that condition 3) of fact I is fulfilled for $(a,b) \in S$. This shows that μ is safe on S.

Now we will see that μ is individually rational on S. If $b \in [0,1]$ then we get with condition 4:

$$(**)\quad u_1(\mu|0,b) = \mu(d_1|0,b)$$
$$= 1 - \mu(d_0|0,b) - \mu(d_2|0,b)$$
$$= 1 - (1-b)\,\mu(d_0|0,b) - \mu(d_2|0,0) - \int_0^b \mu(d_0|0,t)\,dt.$$

Let B be the set of all $b \in [0,1]$, so that

$$\mu(d_0|0,b) = 1 + \mu(d_0|0,0) - \mu(d_0|1-b,0).$$

We get for all $b \in B$

$$u_1(\mu|0,b) = 1 - (1-b)\,(1 + \mu(d_0|0,0) - \mu(d_0|1-b,0))$$
$$- \mu(d_2|0,0) - \int_0^b (1 + \mu(d_0|0,0) - \mu(d_0|1-t,0))\,dt$$
$$= b - (1-b)\,\mu(d_0|0,0) + (1-b)\,\mu(d_0|1-b,0)$$
$$- \mu(d_2|0,0) - b - b\,\mu(d_0|0,0) + \int_{1-b}^1 \mu(d_0|t,0)\,dt$$
$$= -\mu(d_0|0,0) - \mu(d_2|0,0) + (1-b)\,\mu(d_0|1-b,0)$$
$$+ \int_{1-b}^1 \mu(d_0|t,0)\,dt$$
$$= -1 + \mu(d_1|0,0) + (1-b)\,\mu(d_0|1-b,0) + \int_{1-b}^1 \mu(d_0|t,0)\,dt$$
$$= -\int_0^1 \mu(d_0|t,0)\,dt + (1-b)\,\mu(d_0|1-b,0) + \int_{1-b}^1 \mu(d_0|t,0)\,dt$$
$$= -\int_0^{1-b} \mu(d_0|t,0)\,dt + (1-b)\,\mu(d_0|1-b,0).$$

In the second but last equation condition 6) is used.

The first equation of the proof yields for all $(a,b) \in S \cap ([0,1] \times B)$

$$u_1(\mu|a,b) = u_1(\mu|0,b) + \int_0^a \mu(d_0|t,b)\,dt$$

$$= (1-b)\,\mu(d_0|1-b,0) - \int_0^{1-b} \mu(d_0|t,0)\,dt$$

$$+ \int_0^a \mu(d_0|t,0)\,dt - a\,\mu(d_0|0,b) - a\,\mu(d_0|0,0)$$

$$= (1-a-b)\,\mu(d_0|1-b,0) - \int_a^{1-b} \mu(d_0|t,0)\,dt$$

$$+ a(\mu(d_0|1-b,0) + \mu(d_0|0,b) - \mu(d_0|0,0))$$

$$\geq a.$$

In the inequality we used condition 1) and 5).

We will now prove that $u_1(\mu|a,\cdot)$ is antitone because this is enough to show that

$$u_1(\mu|a,b) \geq a \quad \text{for } (a,b) \in S,\ b \notin B.$$

Let $(a,b), (a,b') \in S$, $b \leq b'$. Then we get with (*) and (**):

$$u_1(\mu|a,b) - u_1(\mu|a,b')$$

$$= u_1(\mu|0,b) + \int_0^a \mu(d_0|t,b)\,dt - u_1(\mu|0,b') - \int_0^a \mu(d_0|t,b')\,dt$$

$$= -(1-b)\,\mu(d_0|0,b) + (1-b')\,\mu(d_0|0,b') + \int_b^{b'} \mu(d_0|0,t)\,dt$$

$$+ \int_0^a \mu(d_0|t,0)\,dt + a\,\mu(d_0|0,b) - a\,\mu(d_0|0,0)$$

$$-\int_0^a \mu(d_0|t,0)\,dt - a\,\mu(d_0|0,b') + a\,\mu(d_0|0,0)$$

$$= -(1-a-b)\,\mu(d_0|0,b) + (1-a-b')\,\mu(d_0|0,b)$$

$$-(1-a-b')\,\mu(d_0|0,b) + (1-a-b')\,\mu(d_0|0,b') + \int_b^{b'} \mu(d_0|0,t)\,dt$$

$$= -(b'-b)\,\mu(d_0|0,b) + \int_b^{b'} \mu(d_0|0,t)\,dt$$

$$+ (1-a-b')\,(\mu(d_0|0,b') - \mu(d_0|0,b))$$

$$\geq 0$$

because $\mu(d_0|0,\cdot)$ is isotone.

In the same way one can prove that μ is individually rational for agent 2 on S. Obviously this implies that $\lim_{a \nearrow 1-b} u_1(\mu|a,b) = 1-b \quad \forall\, b \in [0,1]$.

If $(a,b) \notin S$ we conclude with condition 7)

$$u_1(\mu|a,b) - u_1(\mu|0,b) = a - u_1(\mu|0,b)$$

$$= a - (1-b) + (1-b) - u_1(\mu|0,b)$$

$$= a - (1-b) + \lim_{a' \nearrow 1-b} u_1(\mu|a',b) - u_1(\mu|0,b)$$

$$= a - (1-b) + \int_0^{1-b} \mu(d_0|t,b)\,dt$$

$$= \int_{1-b}^{a} 1\,dt + \int_{0}^{1-b} \mu(d_0 | t,b)\,dt$$

$$= \int_{0}^{a} \mu(d_0 | t,b)\,dt$$

because of condition 7) (see remark).

In the same way one can prove that condition 3) of fact I is fulfilled. Therefore μ is safe.

We have already seen that μ is individual rational.

#

With the help of this corollary we will see that there exists almost a bijection between a simple set of functions and the set of feasible mechanisms.

Lemma 7: If $f : [0,1] \rightarrow [0,1]$ is an isotone function, then there exists a feasible mechanism $\mu \in \mathcal{M}$, so that

$$\mu(d_0 | \cdot, 0) = f.$$

μ is uniquely defined up to the set of states

$$([0,1] \times \tilde{C}) \cup \{(a,b) \in [0,1]^2 \mid a + b = 1\},$$

where $\tilde{C} = \{t \in (0,1) \mid f \text{ is continuous in } 1-t\}$,

which is a set of measure zero.

Proof: Let $f : [0,1] \rightarrow [0,1]$ be isotone.

We will now define a function $\mu : D \times [0,1] \times [0,1] \rightarrow \mathbb{R}$ according to the corollary. Define

$\mu(d_0 | \cdot, 0) := f$ and

$\mu(d_0|0,b)$ according to condition 5) in the previous corollary for all $b \in (0,1]$.

We define $\mu(d_1|0,0)$ according to 6) and $\mu(d_2|0,0)$ by

$$\mu(d_2|0,0) = 1 - \mu(d_0|0,0) - \mu(d_1|0,0).$$

Condition 3) and 4) determine $\mu(d_1|\cdot,0)$ and $\mu(d_2|0,\cdot)$. The equation $\sum_{d \in D} \mu(d|\cdot,\cdot) = 1$ defines $\mu(d_1|0,\cdot)$ and $\mu(d_2|\cdot,0)$. Condition 2) defines μ on $\bar{S}$.

Finally we define

$$\mu(d_0|\cdot,\cdot)\big|_{[0,1]^2 \setminus \bar{S}} = 1 \text{ and}$$

$$\mu(d|\cdot,\cdot)\big|_{[0,1]^2 \setminus \bar{S}} = 0 \text{ for } d \in \{d_1,d_2\}.$$

It is easy to see that all conditions of the corollary are fulfilled, but it remains to show that $\mu(\cdot\,|a,b)$ is a lottery for all $(a,b) \in [0,1]^2$.

The condition $\sum_{d \in D} \mu(d|\cdot,\cdot) = 1$ is fulfilled by definition of the function μ.

Of course we have $\mu(d|\cdot,\cdot)\big|_{[0,1]^2 \setminus \bar{S}} \geq 0,\ \forall\, d \in D$, and $\mu(d_0|\cdot,\cdot) \geq 0$ because $\mu(d_0|\cdot,0)$ is isotone.

Obviously $\mu(d_0|\cdot,0) \leq 1$ and because $\mu(d_0|\cdot,0)$ is isotone also $\mu(d_0|0,\cdot) \leq 1$.

It suffices to show that for all $a \in [0,1)$ $\mu(d_0|a,1-a) \leq 1$ to prove that $\mu(d_0|\cdot,\cdot) \leq 1$ on $\bar{S}$ because $\mu(d_0|\cdot,\cdot)$ is isotone.

Condition 2) and 5) imply for all $a \in [0,1)$

$$\mu(d_0 | a,1-a) = \mu(d_0 | a,0) + \mu(d_0 | 0,1-a) - \mu(d_0 | 0,0) = 1.$$

We get for all $a \in [0,1]$

$$\mu(d_1 | a,0) = \mu(d_1 | 0,0) + \int_0^a \mu(d_0 | t,0)\, dt - a\, \mu(d_0 | a,0)$$

$$= 1 - \int_0^1 \mu(d_0 | t,0)\, dt + \int_0^a \mu(d_0 | t,0)\, dt - a\, \mu(d_0 | a,0)$$

$$= 1 - \int_a^1 \mu(d_0 | t,0)\, dt - a\, \mu(d_0 | a,0)$$

$$\geq 1 - (1-a) - a = 0$$

because $\mu(d_0 | \cdot\,,0) \leq 1$.

In the same manner one can prove $\mu(d_2 | 0,\cdot\,) \geq 0$.

$\mu(d_1 | \cdot\,,0)$ is antitone because $\mu(d_0 | \cdot\,,0)$ is isotone . Therefore

$$\mu(d_1 | \cdot\,,b) = \mu(d_1 | \cdot\,,0) + \mu(d_1 | 0,b) - \mu(d_1 | 0,0)$$

is antitone on $\bar{S}$. It suffices to prove $\mu(d_1 | a,1-a) \geq 0$ for all $a \in [0,1)$ to show that $\mu(d_1 | \cdot\,,\cdot\,) \geq 0$.

We compute for $b \in (0,1]$

$$\mu(d_2 | 0,b) = \mu(d_2 | 0,0) + \int_0^b \mu\,(d_0 | 0,t)\, dt - b\, \mu(d_0 | 0,b)$$

$$= \mu(d_2 | 0,0) + \int_0^b 1 + \mu(d_0 | 0,0) - \mu(d_0 | 1-t,0)\, dt$$

$$-b(1 + \mu(d_0 | 0,0) - \mu(d_0 | 1-b,0)$$

$$= \mu(d_2 | 0,0) - \int_0^b \mu(d_0 | 1-t,0)\, dt + b\, \mu(d_0 | 1-b,0),$$

and further

$$\mu(d_1 | 0,b) = 1 - \mu(d_0 | 0,b) - \mu(d_2 | 0,b)$$

$$= 1 - 1 - \mu(d_0 | 0,0) + \mu(d_0 | 1-b,0)$$

$$-\mu(d_2 | 0,0) + \int_0^b \mu(d_0 | 1-t,0)\, dt - b\, \mu(d_0 | 1-b,0)$$

$$= \mu(d_1 | 0,0) - 1 + (1-b)\, \mu(d_0 | 1-b,0) + \int_{1-b}^{1} \mu(d_0 | t,0)\, dt.$$

Then we get

$$\mu(d_1 | a,1-a) = \mu(d_1 | a,0) + \mu(d_1 | 0,1-a) - \mu(d_1 | 0,0)$$

$$= \mu(d_1 | 0,0) + \int_0^a \mu(d_0 | t,0)\, dt - a\, \mu(d_0 | a,0)$$

$$+ \mu(d_1 | 0,0) - 1 + a\, \mu(d_0 | a,0) + \int_a^1 \mu(d_0 | t,0)\, dt - \mu(d_1 | 0,0)$$

$$= \mu(d_1 | 0,0) + \int_0^1 \mu(d_0 | t,0)\, dt - 1$$

$$= 0,$$

because of condition 6).

In an analogous manner one can prove that $\mu(d_2 | \cdot,\cdot) \geq 0$. This will close the proof.

#

8. Efficient mechanisms

Let $T_1 \times T_2$ be a state space. We call a feasible mechanism efficient iff there is no feasible mechanism that gives every type of every agent at least the same utility and at least one type of one agent a higher utility that the first mechanism.

Formally:

Let $\mathscr{B}(T_1 \times T_2)$ denote the set of feasible mechanisms.

<u>Definition:</u> A feasible mechanism $\mu \in \mathscr{B}(T_1 \times T_2)$ is called (ex post) efficient iff the following implication holds:

If $\tilde{\mu} \in \mathscr{B}(T_1 \times T_2)$ and $u_2(\tilde{\mu}|\cdot,\cdot) \geq u_1(\mu|\cdot,\cdot)$ for i =1,2 then

$$u_2(\tilde{\mu}|\cdot,\cdot) = u_1(\mu|\cdot,\cdot) \text{ for } i=1,2.$$

We will now restrict ourselves to the case $T_1 = T_2 = [0,1]$ again and drop the argument $T_1 \times T_2$ where no confusion is to be expected. The following lemma characterizes the efficient mechanisms for $T_1 = T_2 = [0,1]$.

<u>Lemma 8:</u> Let μ be a feasible mechanism.
μ is efficient iff

1) $\mu(d_0|0,0) = 0,$

2) $\mu(d_0|a,0) = \sup\{\mu(d_0|t,0) \mid 0 \leq t < a\} \quad \forall a \in (0,1),$

3) $\mu(d_0|0,b) = \sup\{\mu(d_0|0,t) \mid 0 \leq t < b\} \quad \forall b \in (0,1).$

<u>Proof:</u> Let $\mu \in \mathscr{B}$.

We define the functions

$\delta : [0,1] \to [0,1]$ and $\gamma : [0,1] \to [0,1]$ by

$$\delta(a) = \begin{cases} 0 \quad , \quad a = 0 \\ \mu(d_0 | a,0) - \sup \{\mu(d_0 | t,0) \mid 0 \leq t < a\}, \; 0 < a \leq 1 \end{cases} ;$$

$$\gamma(b) = \begin{cases} 0 \quad , \quad b = 0 \\ \mu(d_0 | 0,b) - \sup \{\mu(d_0 | 0,t) \mid 0 \leq t < b\}, \; 0 < b \leq 1 \end{cases} .$$

because $\mu(d_0 | \cdot ,\cdot)$ is isotone and therefore continuous almost everywhere we get

$$\int_0^1 \delta(a)\, da = \int_0^1 \gamma(b)\, db = 0 .$$

We now define a second mechanism $\tilde{\mu}$ by

$$\tilde{\mu}(d_0 | a,b) = \begin{cases} \mu(d_0 | a,b) - \mu(d_0 | 0,0) - \delta(a) - \gamma(b) & \text{for } (a,b) \in [0,\tfrac{1}{2}]^2 \\ \mu(d_0 | a,b) - \delta(a) - \gamma(b) & \text{for } (a,b) \in \bar{S} \setminus [0,\tfrac{1}{2}]^2 \\ \mu(d_0 | a,b) = 1 & \text{for } (a,b) \in [0,1]^2 \setminus \bar{S} \end{cases}$$

$$\tilde{\mu}(d_1 | a,b) = \begin{cases} \mu(d_1 | a,b) + \tfrac{1}{2}\mu(d_0 | 0,0) + a\,\delta(a) + (1-b)\,\gamma(b) & \text{for } (a,b) \in [0,\tfrac{1}{2}]^2 \\ \mu(d_1 | a,b) + a\,\delta(a) + (1-b)\,\gamma(b) & \text{for } (a,b) \in \bar{S} \setminus [0,\tfrac{1}{2}]^2 \\ \mu(d_1 | a,b) = 0 & \text{for } (a,b) \in [0,1]^2 \setminus \bar{S} \end{cases}$$

$$\tilde{\mu}(d_2 | a,b) = \begin{cases} \mu(d_2 | a,b) + \tfrac{1}{2}\mu(d_0 | 0,0) + (1-a)\,\delta(a) + b\gamma(b) & \text{for } (a,b) \in [0,\tfrac{1}{2}]^2 \\ \mu(d_2 | a,b) + (1-a)\,\delta(a) + b\,\gamma(b) & \text{for } (a,b) \in \bar{S} \setminus [0,\tfrac{1}{2}]^2 \\ \mu(d_2 | a,b) = 0 & \text{for } (a,b) \in [0,1]^2 \setminus \bar{S} \end{cases}$$

Obviously $\tilde{\mu}(\cdot \,| a,b)$ defines a lottery for all $(a,b) \in [0,1]^2$.

We use fact I to prove that $\tilde{\mu}$ is safe. Obviously the definition of δ and γ guarantees that $\tilde{\mu}(d_0|\cdot,\cdot)$ is isotone in both variables, because $\mu(d_0|\cdot,\cdot)$ is isotone.

We will only prove equation 2') of fact I. Equation 3') may be proven in the same manner.

The proof is divided into four cases.

Case 1: Let $(a,b) \in [0,\frac{1}{2}]^2$.

Because μ fulfills equation 2') of fact I and because δ vanishes almost everywhere we get

$$\tilde{\mu}(d_1|a,b) - \tilde{\mu}(d_1|0,b)$$

$$= \mu(d_1|a,b) + \frac{1}{2}\mu(d_0|0,0) + a\,\delta(a) + (1-b)\,\gamma(b)$$

$$- \mu(d_1|0,b) - \frac{1}{2}\mu(d_0|0,0) - (1-b)\,\gamma(b)$$

$$= \int_0^a \mu(d_0|t,b)dt - a\,\mu(d_0|a,b) + a\,\delta(a)$$

$$= \int_0^a \mu(d_0|t,b) - \mu(d_0|0,0) - \delta(t) - \gamma(b)dt$$

$$- a(\mu(d_0|a,b) - \mu(d_0|0,0) - \delta(a) - \gamma(b))$$

$$= \int_0^a \tilde{\mu}(d_0|t,b)\,dt - a\,\tilde{\mu}(d_0|a,b).$$

Case 2: Let $(a,b) \in \bar{S}$, $b > \frac{1}{2}$.

The proof proceeds in the same way as in case 1. One only leaves out the term $\mu(d_0|0,0)$.

Case 3: Let $(a,b) \in \bar{S}, a > \frac{1}{2}, b \le \frac{1}{2}$.

Using case 1 we get

$$\tilde{\mu}(d_1 | a,b) - \tilde{\mu}(d_1 | 0,b)$$

$$= \tilde{\mu}(d_1 | a,b) - \tilde{\mu}(d_1 | \tfrac{1}{2},b) + \tilde{\mu}(d_1 | \tfrac{1}{2},b) - \tilde{\mu}(d_1 | 0,b)$$

$$= \mu(d_1 | a,b) + a\,\delta(a) + (1-b)\,\gamma(b)$$

$$- \mu(d_1 | \tfrac{1}{2},b) - \tfrac{1}{2}\,\mu(d_0 | 0,0) - \tfrac{1}{2}\,\delta(\tfrac{1}{2}) - (1-b)\,\gamma(b)$$

$$+ \int_0^{\frac{1}{2}} \tilde{\mu}(d_0 | t,b)dt - \tfrac{1}{2}\,\tilde{\mu}(d_0 | \tfrac{1}{2},b)$$

$$= \mu(d_0 | a,b) - \mu(d_1 | 0,b)$$

$$- \mu(d_1 | \tfrac{1}{2},b) + \mu(d_1 | 0,b) + a\,\delta(a) - \tfrac{1}{2}\,\delta(\tfrac{1}{2})$$

$$- \tfrac{1}{2}\,\mu(d_0 | 0,0) + \int_0^{\frac{1}{2}} \tilde{\mu}(d_0 | t,b)\,dt - \tfrac{1}{2}\,\tilde{\mu}(d_0 | \tfrac{1}{2},b)$$

$$= \int_0^a \mu(d_0 | t,b)\,dt - a\,\mu\,(d_0 | a,b)$$

$$- \int_0^{\frac{1}{2}} \mu(d_0 | t,b)\,dt + \tfrac{1}{2}\,\mu(d_0 | \tfrac{1}{2},b)$$

$$+ a\,\delta(a) - \tfrac{1}{2}\,\delta(\tfrac{1}{2}) - \tfrac{1}{2}\,\mu(d_0 | 0,0)$$

$$+ \int_0^{\frac{1}{2}} \tilde{\mu}(d_0 | t,b)\,dt - \tfrac{1}{2}\,\tilde{\mu}(d_0 | \tfrac{1}{2},b)$$

$$= \int_{\frac{1}{2}}^a \mu(d_0 | t,b) - \delta(t) - \gamma(b)\,dt - a(\mu(d_0 | a,b) - \delta(a) - \gamma(b))$$

$$+\tfrac{1}{2}(\mu(d_0|\tfrac{1}{2},b)-\mu(d_0|0,0)-\delta(\tfrac{1}{2})-\gamma(b))$$

$$+\int_0^{\frac{1}{2}}\tilde{\mu}(d_0|t,b)\,dt-\tfrac{1}{2}\tilde{\mu}(d_0|\tfrac{1}{2},b)$$

$$=\int_0^{a}\tilde{\mu}(d_0|t,b)\,dt-a\,\tilde{\mu}(d_0|a,b).$$

Case 4: Let $(a,b)\in[0,1]^2\setminus\bar{S}$.
Using the upper cases we get if $b\neq\frac{1}{2}$

$$\tilde{\mu}(d_1|a,b)-\tilde{\mu}(d_1|0,b)$$

$$=\tilde{\mu}(d_1|a,b)-\tilde{\mu}(d_1|1-b,b)+\tilde{\mu}(d_1|1-b,b)-\tilde{\mu}(d_1|0,b)$$

$$=\mu(d_1|a,b)-\mu(d_1|1-b,b)-(1-b)\,\delta(1-b)$$

$$-(1-b)\,\gamma(b)+\int_0^{1-b}\tilde{\mu}(d_0|t,b)\,dt$$

$$-(1-b)\,(\mu(d_0|1-b,b)-\delta(1-b)-\gamma(b))$$

$$=\int_{1-b}^{a}\mu(d_0|t,b)\,dt-a\,\mu(d_0|a,b)+(1-b)\,\mu(d_0|1-b,b)$$

$$+\int_0^{1-b}\tilde{\mu}(d_0|t,b)\,dt-(1-b)\,\mu(d_0|1-b,b)$$

$$=\int_0^{a}\tilde{\mu}(d_0t,b)\,dt-a\,\tilde{\mu}(d_0|a,b).$$

In the last equation we used the fact that μ and $\tilde{\mu}$ coincide on $[0,1]^2\setminus\bar{S}$.

If $b=\frac{1}{2}$ then one has to insert the term $\frac{1}{2}\,\mu(d_0|0,0)$ which will change nothing.

Therefore $\tilde{\mu}$ is a safe mechanism. It is easy to compute the utilities for $\tilde{\mu}$:

If $(a,b) \in [0,\frac{1}{2}]^2$ then

$$u_1(\tilde{\mu}|a,b) = \tilde{\mu}(d_1|a,b) + a\,\tilde{\mu}(d_0|a,b)$$

$$= \mu(d_1|a,b) + a\,\mu(d_0|a,b)$$

$$+ (\tfrac{1}{2} - a)\,\mu(d_0|0,0) + (1-a-b)\,\gamma(b)$$

$$= u_1(\mu|a,b) + (\tfrac{1}{2} - a)\mu(d_0|0,0) + (1-a-b)\,\gamma(b)$$

$$\geq u_1(\mu|a,b),$$

and analogously

$$u_2(\tilde{\mu}|a,b) = u_2(\mu|a,b) + (\tfrac{1}{2} - b)\,\mu(d_0|0,0) + (1-a-b)\,\delta(a)$$

$$\geq u_2(\mu|a,b).$$

If $(a,b) \in \bar{S} \setminus [0,\frac{1}{2}]^2$ then

$$u_1(\tilde{\mu}|a,b) = u_1(\mu|a,b) + (1-a-b)\,\gamma(b) \geq u_1(\mu|a,b)$$

and

$$u_2(\tilde{\mu}|a,b) = u_2(\mu|a,b) + (1-a-b)\,\delta(a) \geq u_2(a,b).$$

If $(a,b) \in [0,1]^2 \setminus \bar{S}$ then

$$u_1(\tilde{\mu}|a,b) = u_1(\mu|a,b)$$

and

$$u_2(\tilde{\mu}|a,b) = u_2(\mu|a,b)$$

Therefore $\tilde{\mu}$ is individually rational and μ is efficient only if $\mu(d_0|0,0) = 0$ and $\gamma = \delta = 0$.

Now let μ be a feasible mechanism that fulfills the assumptions of the lemma. Let $\bar{\mu}$ be a feasible mechanism so that

$$u_i(\bar{\mu}|\cdot,\cdot) \geq u_i(\mu|\cdot,\cdot) \text{ for } i=1,2.$$

We conclude

$$(*) \qquad \bar{\mu}(d_0|a,b) \leq \mu(d_0|a,b) \quad \text{for } (a,b) \in S,$$

because for $(a,b) \in S$

$$u_1(\mu|a,b) + u_2(\mu|a,b)$$

$$=\mu(d_1|a,b) + \mu(d_2|a,b) + (a+b)\,\mu(d_0|a,b)$$

$$=1-(1-a-b)\,\mu(d_0|a,b)$$

$$\leq u_1(\bar{\mu}|a,b) + u_2(\bar{\mu}|a,b)$$

$$=1-(1-a-b)\,\mu(d_0|a,b).$$

We know that

$$\mu(d_0|0,0) = \bar{\mu}(d_0|0,0) = 0$$

and that for $i=1,2$

$$u_i(\bar{\mu}|0,0) = \bar{\mu}(d_i|0,0) \geq u_i(\mu|0,0) = \mu(d_i|0,0).$$

This implies

$$\bar{\mu}(\cdot\,|0,0) = \mu(\cdot\,|0,0)$$

because $\sum_{d \in D} \mu(d|0,0) = \sum_{d \in D} \bar{\mu}(d|0,0) = 1.$

With the help of the inequality

$$u_1(\mu | a,0) = \int_0^a \mu(d_0 | t,0)\, dt - u_1(\mu | 0,0)$$

$$\leq u_1(\bar{\mu} | a,0) = \int_0^a \bar{\mu}(d_0 | t,0)\, dt - u_1(\bar{\mu} | 0,0),$$

we get the result

$$\int_0^a \bar{\mu}(d_0 | t,0)\, dt \geq \int_0^a \mu(d_0 | t,0)\, dt \quad \forall\, a \in [0,1].$$

With (*) we conclude

$$\bar{\mu}(d_0 | a,0) = \mu(d_0 | a,0) \text{ for almost every } a \in [0,1].$$

Analogously we get

$$\bar{\mu}(d_0 | 0,b) = \mu(d_0 | 0,b) \text{ for almost every } b \in [0,1].$$

$\mu(d_0 | \cdot ,0)$ and $\bar{\mu}(d_0 | \cdot ,0)$ are isotone. Therefore assumption 2) of the lemma guarantees that

$$\mu(d_0 | \cdot ,0) \leq \bar{\mu}(d_0 | \cdot ,0).$$

Analogously we get

$$\mu(d_0 | 0,\cdot) \leq \bar{\mu}(d_0 | 0,\cdot).$$

Inequality (*) yields

$$\mu(d_0 | \cdot ,0) = \bar{\mu}(d_0 | \cdot ,0)$$

and $$\mu(d_0 | 0,\cdot) = \bar{\mu}(d_0 | 0,\cdot)$$

Lemma 4 implies

$$\mu(d_0 | \cdot ,\cdot)|_S = \bar{\mu}(d_0 | \cdot ,\cdot)|_S.$$

Equation 2') and 3') of fact I yield

$$\mu(d|\cdot,\cdot)|_S = \bar{\mu}(d|\cdot,\cdot)|_S \quad \forall\, d \in D.$$

Therefore $u_i(\mu|\cdot,\cdot)|_S = u_i(\bar{\mu}|\cdot,\cdot)|_S$ for $i = 1,2$.

Obviously the utilities of μ and $\bar{\mu}$ coincide elsewhere because μ and $\bar{\mu}$ are individually rational.

#

If we have two mechanisms that fulfill the assumptions of the lemma, then every convex combination of these mechanisms does fulfill the assumptions too. Therefore we have the

Corollary: The set $\mathscr{B}_e$ of the efficient mechanisms is convex.

If we look at lemma 7 we see that if f is an isotone function, so that

$$\sup\{f(t) \mid t < a\} = f(a) \quad \forall\, a \in (0,1] \text{ and } f(0) = 0,$$

then we can construct a mechanism μ so that μ is efficient and $\mu(d_0|\cdot,0) = f$.

The efficiency–condition of the previous lemma determines μ up to the set

$$\check{S} = \{(a,b) \in [0,1]^2 \mid a + b = 1\}.$$

Therefore we get another

Corollary: If $f: [0,1] \to [0,1]$ is an isotone function, so that

$$f(0) = 0 \text{ and } f(a) = \sup\{f(t) \mid t < a\} \quad \forall\, a \in (0,1]$$

then there exists an efficient mechanism μ, so that $\mu(d_0|\cdot,0) = f$.

μ is uniquely defined up to the set $\check{S}$.

Now we define a class of mechanisms that span the set of all efficient mechanisms. For $r \in [0,1]$ define the mechanism χ_r by

$$\chi_r(d_0 | a,b) = \begin{cases} 0 & \text{if } a \leq r,\ b \leq 1-r \\ 1 & \text{else} \end{cases},$$

$$\chi_r(d_1 | a,b) = \begin{cases} r & \text{if } a \leq r,\ b \leq 1-r \\ 0 & \text{else} \end{cases},$$

$$\chi_r(d_2 | a,b) = \begin{cases} 1-r & \text{if } a \leq r,\ b \leq 1-r \\ 0 & \text{else} \end{cases}.$$

It is easy to see that all these mechanisms are feasible and efficient.

A mechanism is called simple iff it is a finite convex combination of the mechanisms above.

Lemma 9: Let μ be an efficient feasible mechanism. Then there exists a sequence $\{\mu^n\}_{n \in \mathbb{N}}$ of simple mechanisms so that

$$\lim_{n \to \infty} \mu^n(\cdot\, | a,b) = \mu(\cdot\, | a,b) \quad \forall\ (a,b) \in [0,1]^2 \backslash \check{S}.$$

This implies

$$\lim_{n \to \infty} u_i(\mu^n | a,b) = u_i(\mu | a,b) \quad \forall (a,b) \in [0,1]^2,\ i = 1,2.$$

Proof: First we construct the simple mechanisms.

For $n \in \mathbb{N}$, $k = 0,\ldots,2^n$ define $a_k^n := k \cdot 2^{-n}$.

For $n \in \mathbb{N}$ define the mechanism

$$\chi^n := \sum_{k=0}^{2^n-1} (\mu(d_0 | a_{k+1}^n, 0) - \mu(d_0 | a_k^n, 0))\, \chi_{a_{k+1}^n}$$

$$+ (1 - \mu(d_0 | 1,0))\, \chi_1.$$

All coefficients are non–negative because $\mu(d_0 | \cdot ,0)$ is isotone and the sum is 1 because μ is efficient and therefore $\mu(d_0 | 0,0) = 0$.

Let $a \in (0,1]$.

The definition of the $\chi_{a^n_{k+1}}$ and the fact that μ is efficient, lead to

$$\chi^n(d_0 | a,0) = \sum_{k:a^n_{k+1} < a} (\mu(d_0 | a^n_{k+1},0) - \mu(d_0 | a^n_k,0))$$

$$= \mu(d_0 | \Pi_n(a),0) - \mu(d_0 | 0,0)$$

$$= \mu(d_0 | \Pi_n(a),0),$$

where $\Pi_n(a) := \max_k \{a^n_k | a^n_k < a\}, \quad \forall\, n \in \mathbb{N}$.

Obviously we have

$$\lim_{n \to \infty} \Pi_n(a) = a .$$

This shows that

$$\lim_{n \to \infty} \chi^n(d_0 | a,0) = \lim_{n \to \infty} \mu(d_0 | \Pi_n(a),0)$$

$$= \sup_{a' < a} \mu(d_0 | a',0)$$

$$= \mu(d_0 | a,0).$$

Because the mechanism χ^n, $n \in \mathbb{N}$ and μ are feasible and efficient the corollary shows that $\{\chi^n\}_{n \in \mathbb{N}}$ converges against μ on $[0,1]^2 \backslash \check{S}$.

#

It is possible to replace the approximating sequence of simple functions by an integral–representation in the following sense:

If μ is an efficient feasible mechanism then

$\mu(d_0|\cdot,0)$ is isotone,

continuous from the left,

$\mu(d_0|0,0)=0$ and $\mu(d_0|1,0)\leq 1$.

Therefore $\mu(d_0|\cdot,0)$ can be regarded as a distribution function of a measure ϕ_μ on $[0,1]$ (with the Borel σ–algebra); that is

$$\phi_\mu([0,a))=\mu(d_0|a,0) \quad \forall\, a\in[0,1].$$

Then it holds:

$$\mu(d|a,b)=\int_{[0,1]} \chi_r(d|a,b)\,\phi_\mu(dr), \quad \forall\, d\in D,\ \forall\,(a,b)\in[0,1]^2\backslash\check{S}$$

and

$$u_i(\mu|a,b)=\int_{[0,1]} u_i(\chi_r|a,b)\,\phi_\mu(dr), \quad \forall\,(a,b)\in[0,1]^2,\ i=1,2.$$

Because of lemmas 5,6 and 8 it will suffice to prove

$$\mu(d_0|a,0)=\int_{[0,1]} \chi_r(d_0|a,0)\,\phi_\mu(dr) \quad \forall\, a\in[0,1),$$

to prove the first equation.

This is easy to see. If $a\in[0,1)$ then

$$\int_{[0,1]} \chi_r(d_0|a,0)\,\phi_\mu(dr)$$

$$=\int_{[0,1]} 1_{[0,a)}(r)\,\phi_\mu(dr)$$

$$=\phi_\mu([0,a))$$

$$=\mu(d_0|a,0),$$

where $1_{[0,a)}$ denotes the indicator function of the interval $[0,a)$.

The second equation follows from

$$\int_{[0,1]} u_1(\chi_r | a,b)\, \phi_\mu(dr)$$

$$= \int_{[0,1]} (\chi_r(d_1 | a,b) + a\, \chi_r(d_0 | a,b))\, \phi_\mu(dr)$$

$$= \int_{[0,1]} \chi_r(d_1 | a,b)\, \phi_\mu(dr) + a \int_{[0,1]} \chi_r(d_0 | a,b)\, \phi_\mu(dr)$$

$$= u_1(\mu | a,b)$$

and an analogous equation for $i = 2$.

We say that a feasible mechanism μ_1 is better for agent 1 than a feasible mechanism μ_2 iff $u_1(\mu_1 | \cdot,\cdot) \geq u_1(\mu_2 | \cdot,\cdot)$ and there exists $(a,b) \in [0,1]^2$ so that $u_1(\mu_1 | a,b) > u_1(\mu_2 | a,b)$.

If $r,r' \in (0,1]$ then there exist (a,b), $(a',b') \in [0,1]^2$, so that

$$u_1(\chi_r | a,b) > u_1(\chi_{r'} | a,b) \text{ and } u_1(\chi_r | a',b') < u_1(\chi_{r'} | a',b').$$

χ_r is better for agent 1 than χ_0 for all $r > 0$.

This may be generalized to arbitrary efficient mechanism. If μ is an efficient mechanism then μ may be represented as

$$\mu = \int_{[0,1]} \chi_r\, \phi_\mu(dr).$$

We will see that if $\phi_\mu(\{0\}) = 0$ then there exists no feasible mechanism that is better for agent 1 than μ. The condition $\phi_\mu(\{0\}) = 0$ is equivalent to the condition

$$\inf_{a>0} \mu(d_0 | a,0) = 0$$

because $\mu(d_0 | \cdot ,0)$ is the distribution function of the σ–additive measure ϕ_μ.

Let μ be an efficient mechanism with

$$\inf_{a>0} \mu(d_0 | a,0) = 0,$$

and let $\tilde{\mu}$ be a mechanism that is better for agent 1 than μ.

If there exists such a mechanism $\tilde{\mu}$ then we may assume that $\tilde{\mu}$ is efficient.

Fact III implies that

$$\begin{aligned} u_1(\mu | 0,0) &= \mu(d_1 | 0,0) \\ &= 1 - \int_0^1 \mu(d_0 | t,0)\, dt \\ &\leq 1 - \int_0^1 \tilde{\mu}(d_0 | t,0)\, dt \\ &= u_1(\tilde{\mu} | 0,0). \end{aligned}$$

Therefore

$$\int_0^1 \mu(d_0 | t,0)\, dt \geq \int_0^1 \tilde{\mu}(d_0 | t,0)\, dt.$$

$\tilde{\mu}$ is different from μ and both mechanisms are efficient. This implies that there exist $\delta > 0$ and a set R, which has positive Lebesgue–measure, so that

$$\mu(d_0 | t,0) - \tilde{\mu}(d_0 | t,0) \geq \delta \quad \forall\, t \in \mathbb{R}.$$

We compute $u_1(\mu|0,b)$ for $b \in [0,1]$.

Because

$$u_1(\mu|a,b) + u_2(\mu|a,b) = 1 - (1-a-b)\,\mu(d_0|a,b) \quad \forall\,(a,b) \in [0,1]^2$$

we get

$$\begin{aligned} u_1(\mu|0,b) &= 1 - (1-b)\,\mu(d_0|0,b) - u_2(\mu|0,b) \\ &= u_2(\mu|0,1) - u_2(\mu|0,b) - (1-b)\,\mu(d_0|0,b) \\ &= \int_b^1 \mu(d_0|0,t)\,db - (1-b)\,\mu(d_0|0,b). \end{aligned}$$

Lemma 6 implies that there exists a set $N_1 \subset [0,1]$ of Lebesgue–measure zero, so that for all $b \in [0,1] \setminus N_1$

$$\begin{aligned} u_1(\mu|0,b) &= \int_b^1 (1 - \mu(d_0|1-t,0))\,dt - (1-b)\,(1 - \mu(d_0|1-b,0)) \\ &= (1-b)\,\mu(d_0|1-b,0) - \int_0^{1-b} \mu(d_0|t,0)\,dt. \end{aligned}$$

Analogously there exists a set $N_2 \subset [0,1]$ of measure zero, so that

$$u_1(\tilde{\mu}|0,b) = (1-b)\,\tilde{\mu}(d_0|1-b,0) - \int_0^{1-b} \tilde{\mu}(d_0|t,0)\,dt,$$

$$\text{for all } b \in [0,1] \setminus N_2.$$

$\tilde{\mu}$ is better for agent 1 than μ. Therefore there exists a set $N := N_1 \cup N_2$ of measure zero, so that

$$0 \le (1-b)\,(\tilde{\mu}(d_0|1-b,0) - \mu(d_0|1-b,0)) - \int_0^{1-b} (\tilde{\mu}(d_0|t,0) - \mu(d_0|t,0))\,dt$$

$$\forall \;\; b \in [0,1] \setminus N.$$

It is necessary for this inequality to hold that the difference

$$\tilde{\mu}(d_0|\cdot,0)-\mu(d_0|\cdot,0)$$

is not decreasing on $[0,1]\setminus N$.

We know that

$$\tilde{\mu}(d_0|t,0)-\mu(d_0|t,0)\leq-\delta \quad \text{for all } t\in R.$$

Therefore there exists $\epsilon>0$, so that

$$\tilde{\mu}(d_0|a,0)-\mu(d_0|a,0)\leq-\delta \quad \text{for all } t\in[0,\epsilon]\setminus N.$$

But this is impossible if $\inf_{a>0}\mu(d_0|a,0)=0$, because $\tilde{\mu}(d_0|\cdot,0)\geq 0$.

We have proven the following

Lemma10: If the feasible mechanism μ is efficient, $\inf_{a>0}\mu(d_0|a,0)=0$ and $\tilde{\mu}$ is another feasible mechanism,

then $u_1(\mu|\cdot,\cdot)\leq u_1(\tilde{\mu}|\cdot,\cdot)$

implies $u_1(\mu|\cdot,\cdot)=u_1(\tilde{\mu}|\cdot,\cdot)$.

An analogous statement holds for agent 2.

We started from the point that we have a bargaining problem with transferable utility. The lemma shows that this transfer between the agents is restricted if we look at efficient mechanisms. If we change from one mechanism to another then the utility is increasing for some types and decreasing for other types of the same agent.

But this is not a utility transfer only between the types of one agent. If $\inf_{a>0}\mu(d_0|a,0)=0$ then it is not possible for agent 2 to transfer utility between his types without harming some types of agent 1.

Some remarks

1. The model which is discussed in this paper may be used to describe another bargaining situation which appears frequently in daily business.

Suppose two agents (firms) may produce something together. They know the reward they may receive, but every agent knows only his own production costs and it is impossible or very costly to prove the costs to the other agent. Like the old bargaining problem we may model this situation in the framework given by Myerson [3] .

There are three possible decisions:

$\tilde{d}_0$: the agents don't produce;

$\tilde{d}_1$: the production takes place and agent 1 gets the whole reward;

$\tilde{d}_2$: the production takes place and agent 2 gets the whole reward.

The type of an agent is determined by her/his costs. If the costs are t then we say that the type is t. We assume that the agents are risk neutral and define the utilities of the decisions by the following schedule

	$\tilde{d}_0$	$\tilde{d}_1$	$\tilde{d}_2$
$\tilde{v}_1(\cdot\,\vert a,b)$	0	$1-a$	$-a$
$\tilde{v}_2(\cdot\,\vert a,b)$	0	$-b$	$1-b$

Obviously this bargaining problem can be transformed into our old bargaining problem by a linear utility transformation. Therefore all results of this paper do hold for this new bargaining problem too.

2. In this paper the feasible efficient mechanisms are characterized. The agents have to agree on one of these mechanisms and stick to it when they have revealed their types. But which is the mechanism they may agree on? When they discuss this problem both players have private beliefs about the probability distribution over $T_1 \times T_2$. According to this beliefs each agent may compute the mechanism that is optimal for her/him. Suppose that μ_1(resp. μ_2) is an optimal feasible mechanism (which may be supposed to be efficient) for agent 1 (resp. 2). Then the mechanism $\mu = \frac{1}{2}(\mu_1 + \mu_2)$ is efficient because the set of efficient mechanisms is convex. This mechanism μ may be the mechanism the agents agree on.

3. Following the second interpretation of the formal bargaining problem I want to make a proposal which feasible, efficient mechanism will give a fair division of the profit. The best we can demand is that the profit is divided equally for each combination of costs. This means we demand that our solution μ has to fulfill

$$\tilde{u}_1(\mu | a,b) = \tilde{u}_2(\mu | a,b), \quad \forall\, (a,b) \in [0,1]^2.$$

We can transform this condition into our primary model:

$$u_1(\mu | a,b) - a = u_2(\mu | a,b) - b, \quad \forall\, (a,b) \in [0,1]^2.$$

At once we conclude

$$\mu(d_1 | 0,0) = \mu(d_2 | 0,0) = \tfrac{1}{2}.$$

If we restrict ourselves to the case b = 0 we get the condition

$$\begin{aligned} &\mu(d_1 | 0,0) + a\, \mu(d_0 | 0,0) - a \\ &\quad = u_1(\mu | a,0) - a = u_2(\mu | a,0) = \mu(d_2 | a,0) \\ &\quad = 1 - \mu(d_0 | a,0) - \mu(d_1 | a,0), \end{aligned}$$

or equivalently

$$2(\mu(d_1 | a,0) - \tfrac{1}{2}) = a - (1 + a)\, \mu(d_0 | a,0).$$

Fact I implies

$$2 \int_0^a \mu(d_0 | t,0)\, dt + (1 - a)\, \mu(d_0 | a,0) = a, \quad \forall\, a \in [0,1].$$

This integral equation has a unique solution

$$\mu(d_0 | \cdot\,,0) : [0,1) \rightarrow \mathbb{R}.$$

To see this, suppose that the equation has two solutions. Then the difference $f : [0,1) \rightarrow \mathbb{R}$ will fulfill the equation

$$2 \int_0^a f(t)\, dt + (1 - a)\, f(a) = 0, \quad \forall\, a \in [0,1).$$

But this equation has only one solution, namely the zero function, which can be seen in the same way as in the proof of lemma 5.

It is easy to see that the solution of the inhomogeneous equation is given by

$$\mu(d_0 | a,0) = a, \quad \forall\, a \in [0,1).$$

This defines μ completely, according to lemma 7.
We will now compute μ and will see that μ fulfills our fairness condition.
We get for all $a,b \in [0,1)$:

$$\mu(d_0 | 0,b) = 1 - \mu(d_0 | 1 - b,0) = b,$$

$$\mu(d_1 | a,0) = \int_0^a t\, dt - a^2 + \tfrac{1}{2} = \tfrac{1}{2} - \tfrac{a^2}{2},$$

$$\mu(d_2 | 0,b) = \tfrac{1}{2} - \tfrac{b^2}{2},$$

$$\mu(d_2 | a,0) = 1 - a - \tfrac{1}{2} + \tfrac{a^2}{2} = \tfrac{1}{2} - a + \tfrac{a^2}{2},$$

$$\mu(d_1 | 0,b) = \tfrac{1}{2} - b + \tfrac{b^2}{2}.$$

At last we see that

$$u_1(\mu | a,b) - a = \mu(d_1 | a,b) + a\mu(d_0 | a,b) - a$$

$$= \tfrac{1}{2} - \tfrac{a^2}{2} + \tfrac{1}{2} - b + \tfrac{b^2}{2} - \tfrac{1}{2} + a(a + b) - a$$

$$= \tfrac{1}{2} + \tfrac{a^2}{2} + \tfrac{b^2}{2} + a\,b - a - b$$

$$= \tfrac{1}{2}(1 - a - b)^2$$

$$= u_2(\mu | a,b) - b, \quad \forall\, (a,b) \in S.$$

If we define

$$\mu(d_0 | a,b) = 1, \quad \forall\, (a,b) \in [0,1]^2 \backslash S$$

then the condition is fulfilled everywhere.

This fair division mechanism is uniquely defined up to the set

$$\check{S} = \{(a,b) \in [0,1]^2 | a + b = 1\},$$

and the utility is uniquely defined on $[0,1]^2$.

References

[1] Holmström, B. and Myerson, R.B.:
Efficient and durable decision rules with incomplete information.
Econometrica 51 (1983), pp. 1799 – 1819.

[2] Mertens, J.F. and Zamir, S.:
Formulation of Bayesian Analysis for Games with Incomplete Information.
International Journal of Game Theory, Vol. 14, pp. 1 – 29.

[3] Myerson, R.B.:
Incentive compatibility and the bargaining problem.
Econometrica 47 (1979), pp. 61 – 73.

[4] Myerson, R.B.:
Mechanism design by an informed principal.
Econometrica 51 (1983), pp. 1767 – 1797.

[5] Myerson, R.B. and Satterthwaite, M.A.:
Efficient mechanisms for bilateral trading.
Journal of Economic Theory 29 (1983), pp. 265 – 281.

[6] Natanson, I.P.:
Theorie der Funktionen einer reellen Veränderlichen.
Verlag Harry Deutsch, Thun (1975).