

# A Dutch Book for Group Decision-Making?

**Forthcoming in** Benedikt Löwe, Eric Pacuit, Jan-Willem Romeijn (eds.) *Foundations of the Formal Sciences VI--Reasoning about Probabilities and Probabilistic Reasoning*.

Luc Bovens (London School of Economics)

L.Bovens@LSE.ac.uk

and

Wlodek Rabinowicz (University of Lund)

Wlodek.Rabinowicz@fil.lu.se

**Abstract.** Inspired by Todd Ebert’s problem of the hats, we construct a betting situation which seems to show that a Dutch Book can be made against a group of three rational players with common priors who are making self-interested or group-interested decisions and who have full trust in the other players’ rationality. But appearances are misleading – no such Dutch Book can be made. The moral of the story is that it is incorrect to identify degrees of beliefs with *posted* betting-rates. We signal some connections of our work to the tragedy of the commons and to strategic voting.

## 1. The Problem of the Hats and a Dutch Book for a Group of Rational Players

The problem of the hats is a mathematical puzzle introduced in 1998 by Todd Ebert.<sup>1</sup> We randomly distribute white and black hats to a group of  $n$  rational players. We do this in the dark, with each player having an independent fifty-fifty chance of receiving a hat of one colour or the other. When the lights are turned on, each player can see the colour of the hats of the other players but not of his own hat.

---

<sup>1</sup> <http://www.msri.org/people/members/sara/articles/hat.html>

They are asked to simultaneously name the colour of their own hats or to pass. If at least one person is correct and nobody is in error, a prize will be awarded to the group. They can engage in pre-play communication (before the hats are distributed) to design a strategy. What is the optimal strategy? For  $n = 3$ , the solution is simple. If all and only the players who see two hats of the same colour call out the opposite colour to the colour of their own hat, whereas the others pass, then the group wins in all cases in which the hats are not of the same colour—that is, in  $\frac{3}{4}$  of the cases. It is an open problem, however, whether there exists a general solution, for an arbitrary  $n$ , to this optimisation problem.

We do not intend to tackle the problem of the hats, but take it as a starting point for setting up a Dutch Book, or what seems like a Dutch book, against a group of rational players with common priors and full trust in each other's rationality.

Again distribute the hats amongst a group of three players. This time we disallow pre-play communication. Clearly, the chance that

(A) not all hats are of the same colour

is  $\frac{3}{4}$ . The light is switched on and all players can see the colour of the hats of the other persons, but not the colour of their own hats. Then no matter what combination of hats was assigned, at least one player will see two hats of the same colour. For her the chance that not all hats are of the same colour strictly depends on the colour of her own hat and hence equals  $\frac{1}{2}$ .

On Lewis's principal principle, a rational player will let her degrees of belief be determined by these chances. So before the light is switched on, all players will assign degree of belief of  $\frac{3}{4}$  to (A) and after the light is turned on, at least one player will assign degree of belief of  $\frac{1}{2}$  to (A). Suppose that before the light is turned on a bookie offers to sell a single bet on (A) with stakes \$4 at a price of \$3 and

subsequently offers to buy a single bet on (A) with stakes \$4 at a price of \$2 after the light is turned on. Suppose, finally, that all of the above is common knowledge among the players.

If, following Ramsey, the degree of belief equals the betting rate (i.e. the price-stake ratio) at which the player is both willing to buy and willing to sell a bet on a given proposition, then any of the players would be willing to buy the first bet and at least one player would be willing to sell the second bet. Whether all hats are of the same colour or not, the bookie can make a Dutch book—she has a guaranteed profit of \$1.

	(A) is true: Not all hats are of same colour	(A) is false: All hats are of same colour
Bet 1	Player buys bet for \$3 Bookie pays out \$4	Player buys bet for \$3 Bet is lost
Bet 2	Bookie buys bet for \$2 Player pays out \$4	Bookie buys bet for \$2 Bet is lost
Payoffs	Bookie gains \$1	Bookie gains \$1

So, seemingly, the bookie has succeeded in making a Dutch book against a group of rational players. But the fact that a Dutch book can be made is a mark of some form of irrationality. There are two possibilities. Either each player is trying to increase her own payoff. Then the Dutch Book would not be too worrisome. After all, prisoner's dilemmas have a similar structure—when each player acts to her own advantage, the group payoff is suboptimal. In this case individual rationality would just not be in line with group rationality. Or, alternatively, the players are supposed to act in the interest of the group

as a whole, i.e. to maximize the group's total payoff, rather than their own winnings. In this case the Dutch book would be worrisome—it would be indicative of an internal breakdown in group rationality.

In order not to keep the reader in suspense, let us say straightaway that the rational course of action is not to sell the bookie the second bet and hence that no Dutch book can be made—neither when the players are trying to increase their own payoffs nor when they are trying to increase the group payoff. But nonetheless, our solution to the paradox will prove rewarding. It forces us to think carefully about how one should state Ramsey's claim that degrees of belief are the betting rates at which a player is willing to buy or sell bets on a given proposition. We will conclude with suggestions how the reasoning in this hats puzzle is relevant to the tragedy of the commons and to strategic voting.

## **2. The Dutch Book Disarmed for Self-Interested Decision-Makers**

Let us focus on the second bet – the one that the bookie offers to buy. Suppose the players are trying to maximise their own expected payoffs. Then we need to determine the probability  $p_i$  with which a player  $i$  should step forward and offer to sell the bet if she sees two hats of the same colour. (Obviously, she shouldn't step forward to sell the bet if she sees two hats of different colours. For in that case she knows that the bet would be won by the bookie.) Since there is no pre-play communication and the players are symmetrically placed, we are looking for a symmetrical Nash equilibrium  $\langle p_a, p_b, p_c \rangle = \langle p, p, p \rangle$  for players Alice, Bob and Carol.

Suppose Alice sees two hats of the same colour. We calculate the expected utility for Alice of stepping forward given that she sees two hats of the same colour,  $E[U_a(\langle 1, p, p \rangle)]$ , by first conditioning on the random variable  $S$  with values  $S = 0$  when the hats are of different colours and  $S = 1$  when all the hats are of the same colour. Obviously,  $P(S = 1) = 1 - P(S = 0)$ .

$$(1) \quad E[U_a(<1, p, p>)] = E[U_a(<1, p, p>|S = 0)]P(S = 0) + E[U_a(<1, p, p>|S = 1)](1 - P(S = 0))$$

Since Alice has seen two hats of the same colour,  $P(S = 0) = 1/2$ . If the hats are of different colours, Alice is the only one who will step forward and so her payoff  $E[U_a(<1, p, p>|S = 0)] = -\$2$ . (Remember that the bookie will win the second bet if  $S = 0$ .) To calculate  $E[U_a(<1, p, p>|S = 1)]$ , we condition on the random variable  $N$  with values  $N = i$  when  $i$  other persons beside Alice decide to step forward to sell the bet, for  $i$  ranging over 0, 1 and 2. Hence,

$$(2) \quad E[U_a(<1, p, p>|S = 1)] = \text{Sum}_{i=0,1,2} E[U_a(<1, p, p>|S = 1, N = i)]P(N = i|S = 1)$$

The assumption is that the bookie chooses randomly among the players who offer to sell the bet he wants to buy. Thus, if  $i$  other players beside Alice step forward to sell the bet, each of them, Alice included, has the chance of  $1/(i+1)$  of being the seller. The values of the components of the sum in the right-hand side of Eq. (2) can therefore be read off from the following matrix (remember that the player who gets to sell the bet will win \$2 if  $S = 1$ ):

$i$	$E[U_a(<1, p, p> S = 1, N = i)]$	$P(N = i S = 1)$
0	$2 (= 2 \times 1/1)$	$(1-p)^2$
1	$1 (= 2 \times 1/2)$	$2p(1-p)$
2	$2/3 (= 2 \times 1/3)$	$p^2$

Hence,

$$(3) \quad E[U_a(<1, p, p>)] = (-2)(1/2) + (2(1-p)^2 + 2p(1-p) + (2/3)p^2)(1/2) \\ = 1/3(p-3)p$$

We now examine symmetric Nash equilibria in pure strategies. In view of Eq. (3), unilateral deviation from  $\langle 1, 1, 1 \rangle$  to  $\langle 0, 1, 1 \rangle$  increases Alice's payoff from  $E[U_a(\langle 1, 1, 1 \rangle)] = 1/3(1-3)1 = -2/3$  to  $E[U_a(\langle 0, 1, 1 \rangle)] = 0$ , but unilateral deviation from  $\langle 0, 0, 0 \rangle$  to  $\langle 1, 0, 0 \rangle$  leaves Alice's payoff constant at  $E[U_a(\langle 0, 0, 0 \rangle)] = E[U_a(\langle 1, 0, 0 \rangle)] = 1/3(0-3)0 = 0$ . Hence  $\langle 0, 0, 0 \rangle$  is the only symmetric equilibrium in pure strategies. To examine whether there is an equilibrium in mixed strategy, we note that, if  $0 < p < 1$ , then  $\langle p, p, p \rangle$  can be a Nash equilibrium only if the pure strategies that  $p$  is the mixture of have the same expected utility:

$$(4) \quad E[U_a(\langle 1, p, p \rangle)] = E[U_a(\langle 0, p, p \rangle)] \\ 1/3(p-3)p = 0$$

However, Eq. (4) has no solution under the constraint  $0 < p < 1$ . Hence, there exists only one symmetric equilibrium, viz.  $\langle 0, 0, 0 \rangle$ .

### 3. The Dutch Book Disarmed for Group-Interested Decision-Makers

Now suppose that each group member instead is out to maximize the payoff of the group. So what should a player do who sees two hats of the same colour?

Consider

$$(5) \quad E[U_g(\langle 0, p, p \rangle)] = E[U_g(\langle 0, p, p \rangle) | S = 0]P(S = 0) + E[U_g(\langle 0, p, p \rangle) | S = 1](1 - P(S = 0))$$

Once again,  $P(S = 0) = 1/2$ . (Note that it is Alice's probability that is in question, since it is her expectation of the group utility that we are after. Thus, it would be more appropriate to write  $E_a[U_g(\langle 0, p, p \rangle)]$  instead of  $E[U_g(\langle 0, p, p \rangle)]$ , but we omit the extra index to keep the notation simpler.)

Furthermore,  $E[U_g(\langle 0, p, p \rangle) | S=0] = 0$ , since, if  $S = 0$ , Alice is the only player who sees two hats of the same colour. However, if  $S = 1$ , i.e., if all hats are of the same colour, then the two other players will also see two hats of the same colour and step forward with probability  $p$ . We condition on the random variable  $N = i$ , for  $i$  being the number of the other players who step forward. So,

$$(6) \quad E[U_g(\langle 0, p, p \rangle) | S = 1] = \text{Sum}_{i=0,1,2} E[U_g(\langle 0, p, p \rangle) | S = 1, N = i] P(N = i | S = 1)$$

$i$	$E[U_g(\langle 0, p, p \rangle)   S = 1, N = i]$	$P(N = i   S = 1)$
0	0	$(1-p)^2$
1	2	$2p(1-p)$
2	2	$p^2$

Consequently,

$$(7) \quad E[U_g(\langle 0, p, p \rangle)] = (0)1/2 + (0(1-p)^2 + 2(2p(1-p)) + 2p^2)1/2 \\ = -(p-2)p$$

Now, consider

$$(8) \quad E[U_g(\langle 1, p, p \rangle)] = E[U_g(\langle 1, p, p \rangle) | S = 0] P(S = 0) + E[U_g(\langle 1, p, p \rangle) | S = 1] (1 - P(S = 0))$$

Again  $P(S = 0) = 1/2$ .  $E[U_g(\langle 1, p, p \rangle) | S = 0] = -2$  and  $E[U_g(\langle 1, p, p \rangle) | S = 1] = 2$ . So  $E[U_g(\langle 1, p, p \rangle)] = 1/2(2) + 1/2(-2) = 0$  for all values of  $p$ .

$\langle 1, 1, 1 \rangle$  is not a Nash equilibrium, since unilateral deviation to  $\langle 0, 1, 1 \rangle$  increases the group's payoff from  $E[U_g(\langle 1, 1, 1 \rangle)] = 0$  to  $E[U_g(\langle 0, 1, 1 \rangle)] = -(1-2)1 = 1$ .  $\langle 0, 0, 0 \rangle$  is a Nash equilibrium, since

unilateral deviation leaves the group's payoff at  $E[U_g(\langle 0, 0, 0 \rangle)] = E[U_g(\langle 1, 0, 0 \rangle)] = 0$ . We then investigate whether there are equilibria in mixed strategies. To do so, we solve the following equation for  $p$  under the constraint  $0 < p < 1$ :

$$(9) \quad E[U_g(\langle 0, p, p \rangle)] = E[U_g(\langle 1, p, p \rangle)] \\ - (p - 2)p = 0$$

Eq. (9) has no solution under the constraint  $0 < p < 1$ . Hence, there exists only one symmetric equilibrium, viz.  $\langle 0, 0, 0 \rangle$ .

#### 4. An Intuitive Account

So why is it that a person whose degree of belief for some proposition is  $\frac{1}{2}$  should not be posting her betting rates accordingly? Why should she refrain from expressing a willingness to sell a bet for \$2 that pays \$4? Since we are looking for a symmetric solution, we need to consider what would happen if every player did declare herself willing to sell a bet at odds that correspond to her betting rate. This would mean that the strategy of each player would be to step forward and offer to accept the second bet if she sees two hats of the same colour. The profile of the player's strategies would thus be  $\langle 1, 1, 1 \rangle$ . We will now provide intuitive arguments to the effect that  $\langle 1, 1, 1 \rangle$  cannot be the rational solution neither in the self-interested nor in the group-interested case.

Let us first consider the case of self-interested decision-making. There are two states—one that is favourable and one that is unfavourable for selling the bet. In the favourable state, all the hats are of the same colour, the bookie loses the bet and the player gains \$2. In the unfavourable state, the hats are of different colours, the bookie wins the bet and the player loses \$2. In the favourable state, three players will step forward and the bookie takes a single bet by randomising over the willing players. So there is

only a  $1/3$  chance that a player who steps forward will actually sell the bet. However, in the unfavourable state, only one player steps forward and she is sure to get the bet. Even though my degree of belief that the hats are of different colours is  $1/2$ , it is irrational to express my willingness to sell a bet at the matching betting rate if this willingness translates into a greater opportunity to sell the bet when I am bound to lose and a lesser opportunity to sell the bet when I am bound to win.

Consider the following analogy. Suppose that my degree of belief in the proposition that it will snow tomorrow at noon is  $1/2$ . Suppose that you are going to the bookmaker in town tomorrow morning and are willing to place a bet for me on that proposition that costs \$2 and pays \$4. I might refrain from letting you do so on grounds of the following reasoning. In the favourable state, viz. when it snows at noon, this is likely to be preceded by a cold night and there is only a small chance that your car will start in the morning. Thus you will probably be unable to drive to town and place the bet. In the unfavourable state, viz. when it does not snow at noon tomorrow, there is a very good chance that you will make it into town. Then clearly it would be irrational for me to give you the assignment to place the bet on my behalf.

In the case of group-interest decision-making, it is also irrational to express my willingness to sell a bet at the matching betting rate. Suppose that we would all do precisely that. Then I would reason as follows when seeing two hats of the same colour. In the favourable state, with all hats of the same colour, two other players will step forward and nothing is lost by my not stepping forward. In the unfavourable state, with the hats being of different colours, I am the only one who would step forward, so I can save the group from a loss by not stepping forward. Hence unilateral deviation from  $\langle 1, 1, 1 \rangle$  to the pure strategy of not stepping forward improves the group payoff. We can conclude that simply stepping forward – i.e. stepping forward with probability 1 – when seeing two hats of the same colour cannot be the rational strategy.

To understand why  $\langle p, p, p \rangle$  is not a Nash Equilibrium for any  $p > 0$ , it is instructive to construct some graphs. First let us look at self-interested decision-making. We have plotted  $E[U_a(\langle 0, p, p \rangle)]$ ,  $E[U_a(\langle 1, p, p \rangle)]$  and  $E[U_a(\langle p, p, p \rangle)] = pE[U_a(\langle 1, p, p \rangle)] + (1-p)E[U_a(\langle 0, p, p \rangle)]$  in figure 1. Note that if the group members display even the slightest inclination to bet, then the expected value of each person's payoff is lower than if they would have refrained from betting. This is easy to understand. For the  $\langle 1, 1, 1 \rangle$  strategy, we said that it was not a good idea to express one's willingness to bet if this translates in placing the bet for sure in the unfavourable situation but only having a one in three chance of placing the bet in the favourable situation. Neither is it a good idea to be inclined to express one's willingness to bet with some positive chance  $p$ , if it is the case that if one were to act on this inclination, then the expression of one's willingness would translate in placing the bet for sure in the unfavourable situation and a less than maximal chance of placing the bet in the favourable situation.

Let us now turn to the group-interested decision-making. We have plotted  $E[U_g(\langle 0, p, p \rangle)]$ ,  $E[U_g(\langle 1, p, p \rangle)]$  and  $E[U_g(\langle p, p, p \rangle)]$  in figure 2. Note that for any value of  $p > 0$ , unilateral deviation to  $\langle 0, p, p \rangle$  increases the payoff function. In this case, the issue is not about loss avoidance. If I am determined to step forward upon seeing two hats of the same colour, then the expected payoff for the group will be zero. But the point is that I can do better on the group's behalf. If the others express their willingness to step forward when seeing two hats of the same colour or if they are inclined to do so with positive chance  $p$ , then I can exploit this by refraining from stepping forward. This would only marginally decrease the chance for a win for the group in the favourable situation when all hats are of the same colour, but it would guarantee the absence of a loss for the group in the unfavourable situation when the hats are of different colours. Since the win and the loss are equal in size and the probability of the favourable state is the same as that of the unfavourable state, the expected payoff to the group is positive if I decide to stay put.

## 5. Discussion

The general lesson is this: Willingness to bet is one thing, but a binding declaration of such willingness is another. Betting rates at which we are equally willing to buy or to sell a bet on a proposition are one thing, but posting these betting rates is a different matter. Degrees of belief might match our betting rates. (There are counter-examples to this claim as well, but nothing we say here addresses this issue.) But they certainly need not match our posted commitments to bet. The former can be expected to differ from the latter if declarations of willingness to bet do not automatically translate into betting opportunities.

In future work, we intend to explore two applications of this problem.

First, suppose the bookie ‘sweetens the pie’ by offering to buy the second bet at stakes  $\$4 - \epsilon$  for a small  $\epsilon$ . In this case the Nash equilibrium is  $\langle p, p, p \rangle$  for some  $p$  such that  $1 > p > 0$  and that equilibrium is no longer the same for self-interested and group-interested decision-making: The value of  $p$  is greater for the former than for the latter. The situation has the structure of a tragedy of the commons. Self-interested fishermen tend to put too many boats on the sea and over-fish so that they can barely make a living, whereas the group-interest demands that we restrict the number of boats on the sea. Self-interested decision-makers will pick the value of  $p$  so that their expected payoff equals 0, whereas the group-interest demands lower values of  $p$  yielding a higher expected utility for the group.

Second, the structure of the decision-problem in the second bet of the ‘sweetened’ hats problem, with group-interested players, is similar to the decision-problem faced by juries. In the hats problem, there is a *group choice* of either taking a bet or not. There is an *individual choice* of either stepping forward to take the bet or not. The *procedure* is that group choice of taking the bet is taken if and only if at least one person steps forward. The *situation* is that all hats either are single-coloured or not. Each person receives a *private signal* as to whether or not the hats might be single-coloured, but this private signal is

not fully reliable. (It is not fully reliable, if both hats you observe are of the same colour.) In a jury vote, there is a *group choice* of either acquitting or convicting the defendant. There is an *individual choice* of voting innocent or guilty. The *procedure* is that the group choice of acquitting is taken if and only if at least one person votes innocent. The *situation* is that the defendant either is guilty or innocent. Each person receives a *private signal* as to whether the defendant is guilty or innocent, but this private signal is not fully reliable. In the hats problem, the question is whether it is rational to step forward to take the bet when I receive a private signal suggesting that all hats are of the same colour—which is a private signal that the bet is favourable from the group’s point of view. In the jury problem, the question is whether it is rational to step forward to vote innocent when I receive a private signal that the defendant is innocent—which is a private signal that the acquittal is favourable from the jury’s point of view. There are differences in detail between the two cases, but both of them can be modelled as instances of the same formal decision-problem.

The core idea of strategic voting (Banks, 1999; Feddersen and Pesendorfer, 1998 and 1999) is that jury members in a unanimous jury will not vote according to their private signal. Suppose they were to do so. As a jury member, my innocent vote only matters if it is pivotal, i.e. if all others have voted guilty. But, in this case, there is overwhelming evidence that the suspect is guilty since all others must have received private signals of guilt. So even if I receive an innocent vote, I should vote guilty. But if all reason like this, then the suspect will be convicted, even if everyone receives a signal of innocence. Still, even if one takes into consideration that the others’ votes might not express their private signals, voting innocent upon receiving a signal of innocence is too rash. Just as in the ‘sweetened’ version of the hats problem, with group-interested players, strategic considerations force jury members to adopt a randomised strategy of voting innocent with probability higher than zero but lower than one, when they receive an innocent signal.

Our methodology to determine a randomised strategy  $\langle p, p, p \rangle$  for group-interested decision-making in the ‘sweetened’ hats problem is (with some qualifications) the same as the methodology for determining a randomised strategy  $\langle p, \dots, p \rangle$  for voting innocent upon receiving an innocent signal in jury voting. And depending on the values of the parameters, there are indeed situations in which the rational choice in jury voting is to randomise with a surprisingly low chance of voting innocent upon receiving an innocent signal, just like the rational solution in the hats problem is to randomise with a surprisingly low chance of stepping forward to take the bet upon seeing two hats of the same colour.

## References

Banks, J.S. (1999) ‘Committee Proposals and Restrictive Rules’ *Proceedings of the National Academy of Sciences of the United States of America*, 96, 14, 8295-8300.

Feddersen, T. and Pesendorfer, W. (1998) ‘Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting’ *The American Political Science Review*, 92, 1, pp. 23-35.

Feddersen, T. and Pesendorfer, W. (1999) ‘Elections, Information Aggregation, and Strategic Voting’ *Proceedings of the National Academy of Sciences of the United States of America*, 96, pp. 10572-74.

Robinson, S. ‘Why Mathematicians Now Care about their Hat Color’ *New York Times-Science*, April 10, 2001.

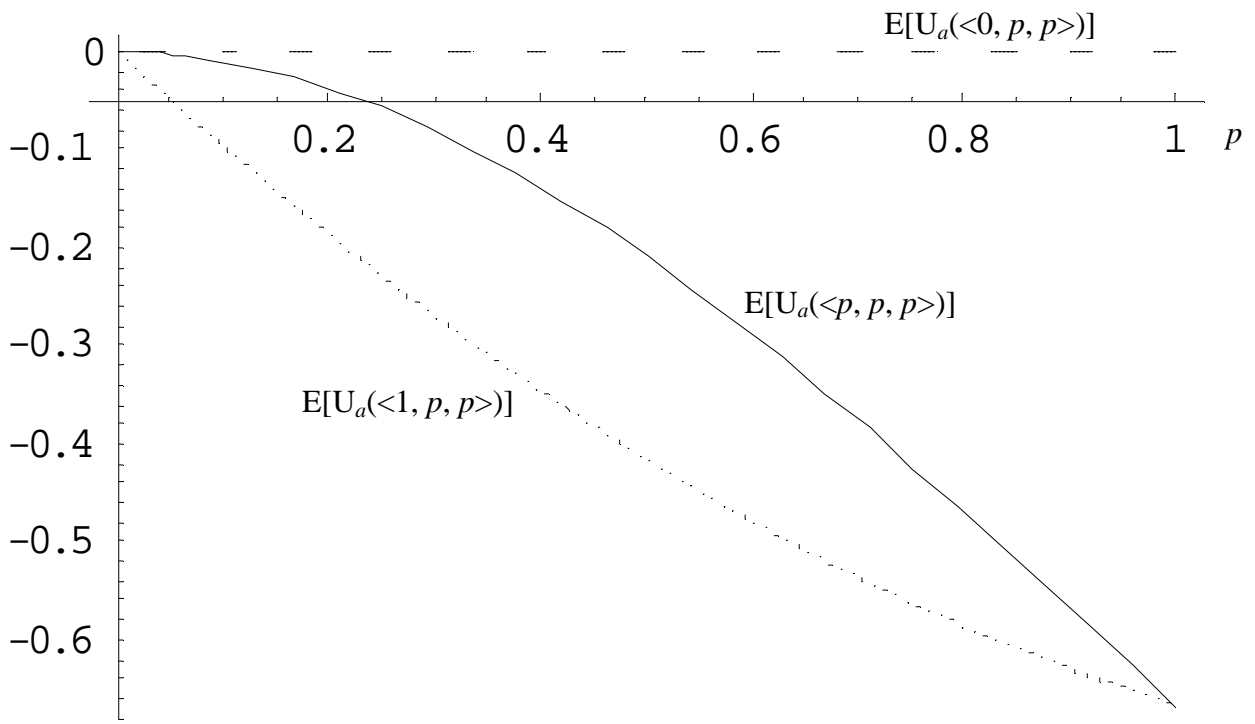


Figure 1. Self-Interested Decision-Making

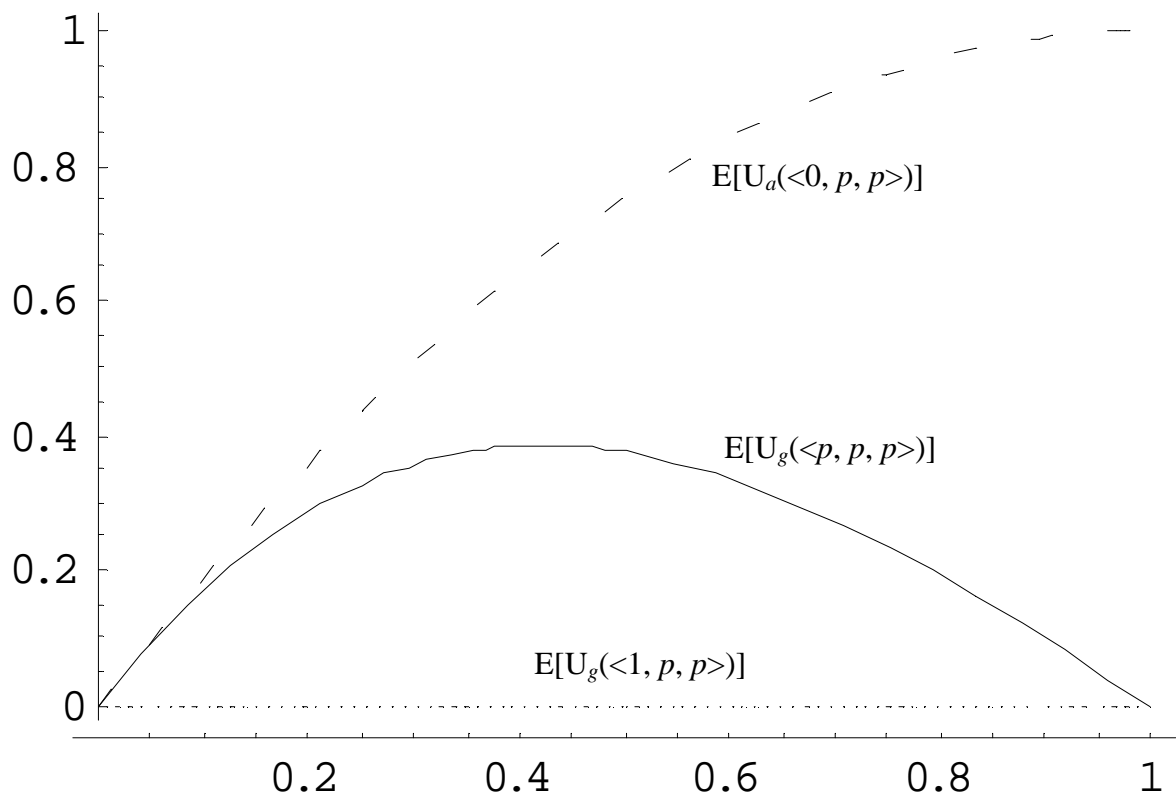


Figure 2. Group-Interested Decision-Making