The Impact of the Degree of Responsibility and Mutual Decision Making on Choices under Risk*

Gilbert G. Eijkelenboom[†] **Alexander Vostroknutov**[‡ §]

July 2017

Abstract

We use a within subjects design to study how responsibility for the payoffs of different number of others influences the choices under risk and how choosing together with another person changes these decisions. After controlling for the regression to the mean, we do not find an effect of responsibility for one other person. We, however, do find that the number of others matters: risk averse subjects choose riskier options when the number of others grows. Mutual responsibility makes choices for others converge to the mean risk preference in the population. It diffuses blame and, depending on individual risk preferences, encourages subjects to make riskier or more cautious choices for others than when choosing alone.

JEL classifications: C91, C92
Keywords: responsibility, blame, risk preferences, choice in groups.

*We would like to thank Matthew Embrey, Martin Strobel, the participants of the Maastricht BEELab meeting and two anonymous reviewers for great comments. The funds for this study were provided by the Marie Curie Reintegration Grant N277033 (FP7-PEOPLE-2010-RG). All mistakes are ours.

[†]School of Business and Economics, Maastricht University, P.O. Box 616, Maastricht 6200 MD, The Netherlands.

[‡]Center for Mind/Brain Sciences, University of Trento, Via delle Regole 101, 38123 Mattarello (TN), Italy. e-mail: a.vostroknutov@unitn.it

[§]Corresponding author.

1 Introduction

In daily life people have to make many important decisions that involve risk. Buying a house, moving to take a new job opportunity, and investing family savings are clear examples. In recessionary periods, like those experienced in Europe recently, these decisions are often exposed to elevated risk. Typically, it is not only the decision maker who is affected by the resulting outcomes, but other people as well. In the examples above all members of the household feel the consequences. Moreover, sometimes decisions like these are made not by sole individuals, but by groups of people. For example, in a household husband and wife might make the choices together. Similarly, in firms managers act on behalf of shareholders. Both *responsibility*, a situation in which the choice of one person affects others to the same degree, and *mutual decision making*, a situation in which several people communicate in order to make choices for themselves and a group of others, are involved. Given the impact of these decisions, it is important to investigate the influence of responsibility and group decision making on choices under risk.

In this study we put forward a hypothesis that changes in behavior under responsibility are the result of *blame avoidance*: a change in behavior that is triggered by the desire to preclude others from forming an unfavorable belief about the decision maker when her choice affects them. In case of decision making under risk and responsibility, we conjecture that blame avoidance should result in choices which are consistent with what the passive affected person(s) would choose, so that the amount of potential blame is minimized when the outcome of the risky choice is realized. In addition, blame avoidance can make decision makers sensitive to the *number* of people their choice affects and whether the choice was made alone or *together* with someone else. Intuitively, in the former case there should be "more" blame when more people are involved and in the latter case blame should diminish since several people are responsible for single decision.

To find out if choices under responsibility are in line with the blame avoidance hypothesis, we conduct an experiment in which we investigate the influence of the *degree* of responsibility and *mutual decision making* on choices under risk. By the degree of responsibility we mean the number of people that are affected by the decision made by a single individual. In particular, we look at the choices between lotteries when the realized outcome is paid to the decision maker and various number of others. By mutual decision making we mean the choice between lotteries made by *two* individuals, who should agree on a choice which would then impact themselves and a group of passive others. With this design we aim at testing the blame avoidance hypothesis as a potential explanation for the responsibility effects. We ask the following questions: How do choices under risk compare when individuals choose just for themselves and when they choose for themselves and others? Does the number of others, who are influenced by the decision, matter? How is mutual decision making different from individual decision making? Blame avoidance hypothesis gives explicit answers to these questions, and our goal is to check if the answers we find are consistent with it.

We are not the first to ask questions about responsibility and decisions under risk. A large literature exists where similar problems are tackled. In general, many studies with between subjects designs find that responsibility makes choices under risk more cautious on average (see Section 2). In a recent article, Bolton, Ockenfels, and Stauf (2015) find that risk loving subjects (as measured by the risky choices when only the subject herself is affected) become more cautious when they are responsible for the payoff of one other subject, but they do not find any shifts towards riskier choices under responsibility among risk averse subjects. The authors attribute this behavior to blame avoidance and possibly other effects.

If blame avoidance is the driving force behind behavioral changes under responsibility, we should observe not only the cautious shift in behavior, as Bolton, Ockenfels, and Stauf (2015) have found, but also the *opposite* effect ("cautious shift" and "risky shift," as literature puts it, see Section 2). Blame avoidance would predict both shifts, as subjects should try to choose as average other would have chosen in their place. We try to find out under which conditions cautious and risky shifts are present. In addition, we test a hypothesis that the increase in the number of people, that are affected by the decision under risk, makes the effect even stronger: the blame is stronger the more people are involved. Finally, we study mutual decision making, where two individuals should come to an agreement on how to choose for themselves and a group of passive others. We hypothesise that in this case both cautious and risky shifts should become equally strong due to the blame diffusion effect between the two decision makers: if an unfavorable outcome is realized the other people affected will not be able to blame a particular individual, but only the pair of the decision makers together. Overall, we postulate that all these hypotheses stem from the desire to avoid blame in the sense of Gurdal, Miller, and Rustichini (2013).

In order to be able to detect both cautious and risky shifts we use a novel completely randomized within subjects design, which allows us to see how heterogeneity in subjects' risk preferences leads to differential behavior in responsibility situations. The subjects make choices in four stages. Each stage is a different Holt and Laury task (Holt and Laury, 2002), which involves 10 decisions between two lotteries. The subjects are told that the payoff that they get from each of the four HL tasks is also delivered to 0, 1, or 3 others.[1] Moreover, in one of the four stages the

[1]This corresponds to the IR and SR+ stages of the Bolton, Ockenfels, and Stauf (2015) experiment. It should be mentioned, however, that we use *different* HL tasks and randomize the order of the stages to avoid fixed sequence effects.

subjects have to make a decision in a HL task together with another subject in the session. Thus, each subject makes choices in 1) *individual choice condition*, where only he is affected by his choice; 2) *responsibility for one other condition*, where the payoff is delivered to the subject who makes the choice and one other subject in the session; 3) *responsibility for three others condition*, where the payoff is delivered to the subject who makes the choice and three others; 4) *mutual responsibility condition*, where the subjects have to agree on their choices in pairs and the payoff is delivered to them and three others. This within subjects design allows us to directly observe the changes in choices under risk depending on the four conditions and on the individual heterogeneity, which is absolutely necessary to detect cautious and risky shifts.

We report the following results. After controlling for the *regression to the mean* that is inherent in our design, we find no effect of responsibility for one other person on the choices of our subjects as compared to the condition where they choose only for themselves.[2] When we compare the choices with responsibility for one and three others we do find evidence of the risky shift: risk averse subjects choose riskier options when choosing for three others than when choosing for one other. This change is significant beyond the effect of the regression to the mean. Our most important finding, though, is that in the mutual responsibility condition, where pairs of subjects choose for themselves and three others, we observe *symmetric* cautious and risky shifts as compared to the individual choices for three others. Subjects who are risk loving, when individually responsible, choose more cautiously when mutually responsible and vice versa for the individually responsible risk averse subjects. This result is highly significant even after taking the regression to the mean into account. Moreover, the choices in the mutual responsibility condition seem to be much more concentrated around the average risk preferences in the population than the choices in the other three conditions. Overall, we find *some* evidence for the blame avoidance hypothesis: there is an effect of the number of others influenced by the risky choice, and under mutual responsibility both cautious and risky shifts are very prominent and symmetric.

2 Related Literature

In this section we overview the literature related to choices with responsibility and choices in groups (mutual responsibility). There exists a wide range of literature on the decision making under risk, with the emphasis on the differences between individual and group risk preferences (when choice is made mutually by the group members). Stoner (1961) laid the foundation. In his study he concluded that people, as members of a group, took on significantly more risk, compared to the situation in which they chose individually. This change was referred to as a "risky shift." The same effect was found in various other studies, for example, in Gardner and

[2] Note that other studies which use within subjects designs and address similar questions do not control for the regression to the mean.

Steinberg (2005).

More recent investigations report somewhat contrasting evidence. For example, Shupp and Williams (2008), who used lotteries to explore differences in choices under risk between individuals and groups, found that groups were significantly more risk averse than individuals when choosing between lotteries that involved high levels of risk. Masclet, Colombier, Denant-Boemont, and Lohéac (2009) and Baker II, Laury, and Williams (2008) also found a "cautious shift": they observed that groups made safer choices than individuals.

The findings that people in a group are *less* risk averse can be explained with the fact that risk is shared by the group members (Wallach, Kogan, and Bem, 1964). When choosing alone, all the results of risky decisions are attributed to a single decision maker. Consequently, in case of non-desirable outcomes, this person is the only one who is seen "responsible" for the outcome. Thus, guilt and shame aversion (aversion to having negative feelings for being responsible for bad outcome not expected by others) force the decision makers to choose more cautiously.[3] This reasoning can also be applied in reverse: being solely responsible for successful events results in a higher utility level, compared to achieving the same with a group (Eliaz, Ray, and Razin, 2006). As a result, utility levels of the decision makers have a smaller spread when decisions are made in groups rather than on an individual basis. This implies that individuals who make decisions together in groups are exposed to less risk, which could explain riskier choices.

The papers cited above use various methods to test their hypotheses regarding group versus individual choices. However, none of the papers uses within subjects design and thus, only the averages between conditions are being compared. This does not allow to look at deeper determinants of the changes in risky choices as our design permits. In particular, within subjects design allows us to see how individual risk preferences determine the shifts in choice under responsibility. Between subjects designs do not allow for this possibility since only averages between conditions can be compared and the differential effects of risk preferences (cautious shift for risk loving subjects and risky shift for risk averse subjects) cannot be detected. In addition, within subjects we can directly compare the influence of the number of others on the choices and the influence of the mutual decision making, which is, again, not possible with between subjects design.[4] As an additional advantage, in our design mutual decision making involves passive group members for whom the choice is made, which adds an additional dimension never studied before. Thus, with our design, we aim at clarifying the mixed results reported in the studies

[3]Many other studies found cautious shift (e.g., Charness and Jackson, 2009; Reynolds, Joseph, and Sherwood, 2009; Pahlke, Strasser, and Vieider, 2012a; Ertac and Gurdal, 2012). It should be mentioned as well that in some studies the concept of loss aversion (Kahneman and Tversky, 1984) is used to explain similar effects, see, for example, Andersson, Holm, Tyran, and Wengström (2014); Pahlke, Strasser, and Vieider (2012b); Rohde and Rohde (2011); Vieider, Villegas-Palacio, Martinsson, and Mejía (2015).

[4]Within subjects designs have their drawbacks. For example, various demand effects can occur as well as wealth and portfolio effects. In our design we tried to minimize these effects as much as possible by 1) randomizing the order of conditions and the lotteries in each condition and 2) not providing information on payoffs between conditions.

above.

Some studies concentrate on the dynamics of group decision making. According to Pruitt (1971), leadership theory predicts that the choice shifts towards the most confident group member. Gardner and Steinberg (2005) studied risk preferences of groups of adolescents and found that peer pressure plays an important role in their behavior. In our design subjects in the mutual choice condition can freely chat with each other in order to reach mutual decision for the group. Given that we have subjects also choose for the group individually, we can see the effect mutual responsibility has on choice. In particular, we can detect if blame is "diffused" when the choice is made by several people.

All studies above compared group decision making to individual choice for the group. It is also interesting to compare an individual choosing just for himself and choosing for himself and a group (Davis, 1992; Bolton, Ockenfels, and Stauf, 2015). In this case risk is not shared by several individuals. The decision maker is the only one who decides and is, therefore, the only person responsible. Thus, the mechanism that worked for the group decision making cannot play any role here. A possible explanation for being more risk averse when deciding for others, might be that people blame the decision maker for bad outcomes that the decision maker is only partially responsible for (in case of lotteries). Gurdal, Miller, and Rustichini (2013) experimentally show that people blame others for events they are not (technically) responsible for. Thus, people who make decisions with responsibility for others may become more risk averse because they fear punishment in the form of blame.[5]

Most studies found that people who are responsible for others choose more cautiously as compared to the situation where they choose just for themselves. We add to this literature in two different ways. First, we do find both cautious and risky shifts in our data. With our within subject design we can establish how individual choice characteristics lead to either cautious or risky shift. Second, we analyze the responsibility choices for the groups of others of different sizes. This allows us to see how blame is perceived by the decision makers.

Our concept of decision making under risk with mutual responsibility combines two strands of literature discussed above: decision making under risk with responsibility and decision making in a group. Firstly, people who decide are responsible for the outcome delivered to them and passive others, but this responsibility is smaller than in the scenario where they bore sole responsibility. Secondly, if a group of people is mutually responsible for the outcomes obtained, they have to collectively choose among the options. This involves group dynamics, similar to what is observed in groups that have to agree upon making a decision. Our prime interest is how mutual responsibility for others changes the choices for the group and how group dynamics is involved. While many studies investigate the connection between decision making under

[5]We are aware of a study where the decision makers choose in a riskier manner when others are involved (Chakravarty, Harrison, Haruvy, and Rutström, 2011). However, in this study the decision maker is not affected by his own choice for others.

risk and responsibility, the topic of mutual responsibility for others has been untouched to our knowledge.

3 Experimental Design

The experiment consist of four stages and a questionnaire. In each of the four stages subjects make choices in a Holt-Laury task, or HL (Holt and Laury, 2002). Namely, in each HL task subjects make 10 choices between two lotteries, A and B. All A (B) lotteries have the same monetary outcomes, but the probabilities of these outcomes vary (Figure 8 in Appendix E illustrates). Subjects make choices in four *different* HL tasks. This means that the monetary outcomes in all HL tasks are different but the probabilities of the outcomes stay fixed across all 10 choices in all tasks (see Table 8 in Appendix A for details). The payoffs were chosen to satisfy two conditions: 1) the expected payoffs from the lotteries across the four HL tasks are approximately the same; 2) assuming the expected utility function $u(x) = x^{1-r}$ the optimal switch from choosing lottery A to lottery B happens at the same place for all four HL tasks for given intervals of r.[6] The main idea of our experiment was to have a complete within subject design with four different conditions but the same kind of task. We could not use one HL task for all four conditions simply because this could create several types of biases in subjects' choices if they were exposed to the same task four times in a row. For example, subjects could ignore the differences in conditions and choose the same switching point, or there could be an experimenter effect: subjects could deliberate on how exactly they are supposed to change their switching point depending on the condition. Thus, the payoffs in the four HL tasks were chosen so that the switching points could be interpreted in the same way independent of the task. To avoid order effects each subject was exposed to a *different* (randomized) sequence of HL tasks. Table 9 in Appendix A shows the exact ordering for each subject.

There are two treatments: Main and Control. In the Control treatment subjects choose *for themselves* in four HL tasks (in randomized order). No information about payoffs was communicated to subjects between HL tasks (see Appendix E for instructions). In the end of the experiment the software chose one of the 10 choices at random in each HL task, "played" the lottery, and assigned each subject the payoff equal to the sum of the four payoffs realized from the four chosen lotteries. It should be noted that we could not avoid paying subjects for each of the four tasks since in the Main treatment others must receive the same payoff as the subject. Therefore, it was necessary to pay subjects for each HL task to keep the Control treatment as close to the Main treatment as possible. Subjects were informed about this procedure and were

[6] Holt and Laury (2002), when choosing the payoffs for the lotteries, followed the following procedure. They assumed the constant relative risk aversion expected utility function $u(x) = x^{1-r}$ and chose the payoffs so that the approximately risk neutral utility maximizer with $-0.15 \leq r \leq 0.15$ switches to lottery B after four choices of lottery A. In addition, they wanted the switch to B after six choices of A to happen for r in the interval $[0.41, 0.68]$, which corresponds to the utility function $\sqrt{x}$. We chose the payoffs in the four HL tasks in the same way.

told that they will see the payoffs they obtained from each task in the end of the experiment.

Paying for all four lotteries might raise concerns about the possibility of the portfolio effect. Harrison and Rutström (2008, pp. 118-119) directly address this issue in a setting where subjects choose in several HL tasks one after the other (all draws are uncorrelated). They are paid for all HL tasks and the goal of the experiment is to elicit subjects' risk preferences. They point out that, in this specific case, which is exactly like our treatments, portfolio effect is actually helpful in the sense that more information about the risk preferences of the subjects can be obtained if we think of them as choosing portfolios instead of making several individual decisions. For example, when a subject makes six safe choices in one HL task and five in the other, this means that her CRRA risk coefficient lies somewhere in the intersection of the intervals of coefficients that are consistent with making these choices. In addition, Harrison and Rutström (2008) report the results of an experiment where they evaluate risk preferences in this setting but vary the number of HL tasks that are paid (1, 2 or 3). They find no difference in the estimations of risk coefficients. They conclude that in the sequential HL tasks setting portfolio effect does not influence the choices. The same should be true in our experiment.

In the Main treatment we expose subjects to different conditions which vary in the degree of responsibility and mutual decision making. In each condition subjects make choices in different HL tasks, as having fixed HL task in each condition could potentially create some undetectable biases. Thus, it was crucial for our design to have four HL tasks in which a subject with some value of r in her utility function would make the same number of safe choices. This would allow us to make inferences about conditions even though subjects made choices in different HL tasks. The purpose of the Control treatment is to ensure that the four HL tasks that we have chosen are suitable for this role.

The Main treatment is the same as the Control treatment but with one difference. The subjects are told that the payoff they obtain in each HL task (as described above) will also be paid to other subject(s). So, if the software generated some payoff x for the subject, the same amount x would be paid to some other individual(s) depending on the condition. There are four conditions which correspond to the four HL tasks.

Condition 1 (Individual Choice). Subjects choose just for themselves as in the Control treatment.

Condition 2 (Responsibility for One Other Person). Subjects choose for themselves and *one other subject in the session*. So, the payoff generated by the software goes to the subject who made the choice and someone else in the session. The identity of the other was kept anonymous.

Condition 3 (Responsibility for Three Other People). Subjects choose for themselves and *three other subjects in the session*. So, the payoff generated by the software goes to the subject who made the choice and three other people in the session. The identities of the others were kept anonymous.

Condition 4 (Mutual Responsibility). Subjects are matched in pairs. Each pair of subjects (the negotiators) sees the same HL task and should reach an agreement on how to choose in each of the 10 cases. The subjects can communicate via online chat using natural language. In order to make a mutual choice both subjects should make identical choices for all 10 cases and press the "Validate" button. After both subjects pressed the button, the software checked if the choices of both subjects were the same and only then allowed them to proceed. After the mutual choice was made, the payoff that was generated from these choices went to both negotiators and *three others* in the session. This design allows us to make comparisons between the Condition 3, where subjects choose *alone* for themselves and three others and *mutual decision making*.

The conditions were *not* presented to the subjects in a fixed order. In each time period all four conditions were played by some subjects in the session. The order was randomized. Table 11 in Appendix A indicates the exact sequence of conditions for each subject and Table 10 shows how subjects were matched. Thus, in our design, subjects were presented with random sequence of HL tasks and random sequence of conditions.

The questionnaire elicited standard demographic data and two questions related to risk aversion. The subjects answered to these questions on a Likert scale. The first question was "*How risk averse/loving are you?*" on the scale from 1 = "very risk loving" to 7 = "very risk averse." The second question was "*How risk averse/loving are you in comparison with other people?*" on the scale from 1 = "most risk loving" to 7 = "most risk averse."

The experiments were conducted at Maastricht University, School of Business and Economics (BEELAB), in April 2014 and June 2017. 24 subjects participated in one session of the Control treatment and 120 subjects in five sessions of the Main treatment. No other data were collected in any form: there were no pilots or discarded sessions. The experiment was programmed in z-Tree (Fischbacher, 2007).

4 Results

4.1 HL Tasks in the Control Treatment

Before we get to the main results we would like to ensure that in the four HL tasks the subjects, when deciding only for themselves, choose (mostly) the same switching point. If this is true, then, in the Main treatment, we can interpret the movements of the switching point between HL tasks as the result of the influence of responsibility. We use the data from the Control treatment where subjects were not responsible for anyone and just made choices in four HL tasks choosing for themselves. We look at the *number of safe choices* of an individual, which is defined as the number of safe options (lotteries A) chosen before switching to the risky option (lotteries B). Thus, high number of safe choices indicates high degree of risk aversion. For the HL tasks 1, 2, 3, and 4 we define the variables hl1, hl2, hl3, and hl4, which are equal to the number of safe choices.[7]

HL task	N obs	Mean	SE	95% Conf Int	
hl1	22	5.89	(0.19)	[5.50	6.28]
hl2	21	5.32	(0.24)	[4.81	5.82]
hl3	22	5.47	(0.21)	[5.04	5.91]
hl4	21	5.42	(0.26)	[4.89	5.96]

Table 1: Summary of the number of safe choices in the four HL tasks in the Control treatment.

Table 1 shows the average number of safe choices in the four HL tasks. The number of observations is not equal to 24 since some subjects chose to switch from option A to option B more than once. These observations were discarded. Figure 4 in Appendix B presents these data graphically, and Figure 6 (Appendix B) shows the histograms of the number of safe choices in the four HL tasks. All four tasks produce very close average numbers of safe choices. The mean number of safe choices in HL tasks 2-4 are almost exactly the same. The mean number of safe choices of the first HL is 2 standard errors away from the other three means. This might look problematic, but the difference is rather small: 5.89 for the first HL vs. 5.40 for the average of HL tasks 2-4, less than 10%. Signed-rank tests give significant difference between HL task 1 and tasks 2 and 3 ($p < 0.045$), but are insignificant for all other pair-wise comparisons ($p > 0.100$). It seems like HL task 1 is different from the other three HL tasks. However, the difference is small and this does not bias our analysis reported below for the following reason. We analyze the within subject *differences* between the number of safe choices in different conditions and all subjects choose in different HL tasks in each condition. Thus, the possible difference of the first HL task should cancel out as there is approximately the same number of subjects choosing in it in all conditions. Thus, we conclude that we are able to use the choices in the mixture of the four HL tasks in the analysis of the influence of responsibility on lottery choices.

4.2 Descriptive Statistics

Next we provide some summary statistics for the Main treatment. Each subject made choices in four conditions: 1) choice in HL task for yourself only (Condition 1); 2) choice for yourself and one other person (Condition 2); 3) choice for yourself and three others (Condition 3); 4) mutual choice with a partner for three others (Condition 4). In what follows we will denote the number of safe choices in the four conditions as t1, t2, t3, and t4.[8]

Table 2 shows the average numbers of safe choices in the four conditions. Notice that the number of observations is less than 120 in all four conditions. We have excluded the observations where subjects made more than one switch between options A and B (such behavior

[7] All variables used in the regressions and other analyses are explained in Appendix C.

[8] Notice that now we look at the number of safe choices in different *conditions*, not in different HL tasks. Subjects in each condition choose in randomized HL tasks.

Condition	N obs	Mean	SE	95% Conf Int	
t1	110	5.52	(0.12)	[5.27	5.77]
t2	112	5.67	(0.13)	[5.40	5.93]
t3	114	5.31	(0.14)	[5.03	5.60]
t4	116	5.29	(0.10)	[5.09	5.50]

Table 2: Summary of the number of safe choices in the four conditions of the Main treatment.

cannot be rationalized by expected utility maximization). Figure 5 in Appendix B shows the averages graphically, and Figure 7 (Appendix B) shows the histograms. One can easily see that the averages are very close. Signed-rank tests show no significant differences between conditions ($p > 0.116$) except between conditions 2 and 4 ($p = 0.0135$). This does not necessarily mean that there is no influence of responsibility on choices. It is possible that the subjects change their choices in different ways depending on their risk preferences (cautious and risky shifts).[9]

Next we define and summarize the *differences* in the number of safe choices between conditions, which are our main variables of interest. We use the following notation. Let tij denote the difference between the number of safe choices between conditions i and j, or, in other words, ti - tj. Thus, t21, for example, is the difference between the number of safe choices in Condition 2 and Condition 1 (t2 – t1).

	From HL ≤ 5				From HL ≥ 6				All Data			
	Mean	SE	95% Conf Int		Mean	SE	95% Conf Int		Mean	SE	95% Conf Int	
t21	0.50	(0.18)	[0.14	0.86]	−0.24	(0.16)	[−0.56	0.09]	0.12	(0.12)	[−0.13	0.37]
t32	0.23	(0.19)	[−0.15	0.60]	−0.82	(0.21)	[−1.25	−0.39]	−0.35	(0.15)	[−0.66	−0.05]
t43	0.68	(0.18)	[0.31	1.05]	−0.77	(0.18)	[−1.13	−0.42]	−0.02	(0.14)	[−0.31	0.27]

Table 3: Summary of the differences in the number of safe choices between four conditions in the Main treatment. "From HL $\leq$ 5" means that the mean was taken over subjects who chose the number of safe choices less or equal to 5 in Condition 1, 2, 3 (for t21, t32, and t43 respectively).

Table 3 shows the averages of the differences in the number of safe choices between Conditions 2-1, 3-2, and 4-3 for subgroups of subjects who chose the number of safe choices less or equal (greater or equal) than 5 (6) in Conditions 1, 2, and 3 respectively (e.g., for t32 we divide subjects into two groups according to their choices in Condition 2). We can see that for t21 risk

[9] In the only study which has the design similar to ours (Bolton, Ockenfels, and Stauf, 2015) the authors find significant overall cautious shift between individual choice condition and choice for one other person. The reason behind it is that in their experiments risk averse subjects on average do not make riskier choices when choosing for the other than when they choose for themselves only. In Bolton, Ockenfels, and Stauf (2015) risky shift is not present, so averages show overall cautious shift. This can possibly stem from Bolton, Ockenfels, and Stauf (2015) using the same HL task for individual and social choices and having individual choice always preceding social one. In the second HL task, when subjects choose for one other, they perfectly understand which choices are riskier than their own. This might preclude subjects from choosing in a riskier way. In our design subjects choose in different HL tasks and social choice precedes individual one in around half of the cases. This procedure creates no interdependencies between the two conditions and is, thus, more reliable.

loving subjects (with the number of safe choices in individual Condition 1 less than or equal to 5) choose on average more safely when they are responsible for one other person. The average of t32 for risk averse subjects (HL $\geq$ 6 in Condition 2) is negative and significant which shows that, when facing three instead of one other, risk loving subjects tend to decrease the number of safe choices on average. Thus, it looks like responsibility for one or three people influences only risk averse subjects. Finally, notice that communication in Condition 4 makes risk averse subjects choose in riskier way on average and risk loving subjects choose more cautiously.

	***N* obs**	**t-test**	**One-Sample Signed-Rank**
From HL ≤ 5			
H_1 : t21 > 0	52	$p = 0.0037$	$p = 0.0142$
H_1 : t32 > 0	48	$p = 0.1127$	$p = 0.1235$
H_1 : t43 > 0	57	$p = 0.0002$	$p = 0.0008$
From HL ≥ 6			
H_1 : t21 < 0	55	$p = 0.0752$	$p = 0.2163$
H_1 : t32 < 0	60	$p = 0.0002$	$p = 0.0010$
H_1 : t43 < 0	53	$p < 0.0001$	$p = 0.0001$

Table 4: T-tests and one-sample signed-rank tests of null hypotheses that the differences in the number of safe choices are zero. All p-values are one-tailed.

To support these conclusions we test the hypotheses that the average differences are equal to zero. Table 4 shows p-values for t-tests and one-sample signed-rank tests. We can draw similar conclusions: t32 is different from zero for risk averse subjects; and t43 is different from zero for both risk averse and risk loving subjects (controlling for multiple hypothesis testing with Bonferroni correction we use the per test threshold p-value of 0.025).

4.3 Responsibility

In this section we analyze the effects of the four conditions on the choices of our subjects. Our goal is to elicit the influence of responsibility for other people on the number of safe choices made. We look at the differences in thresholds between Conditions 2-1, 3-2, and 4-3. The graph on the left of Figure 1 shows the linear regression of t21, the difference in the number of safe choices between Conditions 2 (choice for one other) and 1 (individual choice), on the threshold choice in Condition 1 (t1). The dummy for the threshold being higher or equal to 6 and its interaction with t1 are included.[10] One can see that there is a significant trend towards choosing thresholds close to 5 or 6 safe choices for both risk loving and risk averse subjects. This could be interpreted as the effect of responsibility for one other person on the choices of the subjects. However, this interpretation is incorrect. In the within subjects design it is inherent that subjects

[10]The regression is reported in Column 2 of Table 13 in Appendix D.

who chose extreme thresholds in, say, individual Condition 1 would tend to choose less extreme thresholds in Condition 2, just because there is only one direction in which they can deviate from their individual extreme choice. Therefore, in order to establish the effect of responsibility on our subjects we need to control for the regression to the mean.

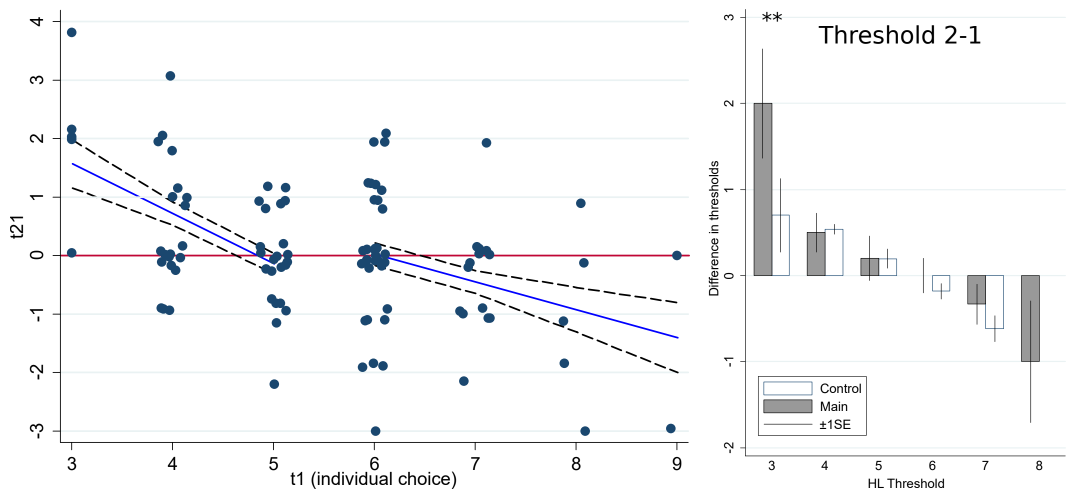

Figure 1: **Left graph.** Regression of t21 on t1, dummy for t1 being higher or equal to 6, and their interaction. The data points are jittered. **Right graph.** Average differences of the thresholds between Conditions 2 and 1, and the differences in the Control treatment. Stars stand for the significance of the rank-sum tests. ** – $p < 0.05$.

The graph on the right of Figure 1 shows the average difference in thresholds for different values of t1 in Condition 1 for the Main treatment and the same quantity for all available thresholds in the Control treatment.[11] Regression to the mean is evident in the choices of the subjects in the Control treatment: subjects who chose small number of safe choices (say, 3) tend to increase the number of safe choices and vice versa for the subjects who chose high number of safe choices. This relationship is significant as demonstrated by a simple regression in Column 1 of Table 5 below.

It might be the case that in the Main treatment the significant trend towards choosing less extreme thresholds for one other person than for yourself is simply regression to the mean and nothing else. To establish an effect of responsibility we need to show that the threshold differences between Conditions 2 and 1 are significantly higher than those in the Control treatment. Column 2 in Table 5 reports the regression. The independent variables are the thresholds in Condition 1 and in the Control treatment (hl); the dummy for the Control treatment (control); and their interaction. The dependent variable is the difference in thresholds between Conditions 2 and 1 in the Main treatment, and the deviation from the mean threshold in the Control treatment.[12] We see that the interaction control·hl is not significant. This means that the differ-

[11]In the Control treatment the differences were calculated in the following way. For each subject the average threshold was calculated from the four choices she made. Then the differences between each choice and that average were taken.

[12]Column 3 in Table 13 in Appendix D shows similar regression with addition of the dummy geq6 which indicates

	Control	Main t21	Main t32	Main t43	Main t31	Main t41	Main t42
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
control		–0.410	–0.850	–1.860***	0.185	–1.541***	–1.496**
		(0.633)	(0.657)	(0.508)	(0.745)	(0.577)	(0.582)
hl	–0.355***	–0.404***	–0.555***	–0.717***	–0.353***	–0.670***	–0.673***
	(0.061)	(0.094)	(0.106)	(0.070)	(0.123)	(0.088)	(0.089)
control·hl		0.048	0.200	0.361***	–0.003	0.315***	0.318***
		(0.112)	(0.122)	(0.093)	(0.137)	(0.107)	(0.108)
cons	1.933***	2.343***	2.783***	3.793***	1.748***	3.474***	3.429***
	(0.341)	(0.534)	(0.562)	(0.376)	(0.663)	(0.466)	(0.472)
N	86	193	194	196	193	194	195
R^2	.36	.20	.26	.51	.17	.41	.47

Table 5: OLS regressions of differences in thresholds between conditions in the Main treatment (Columns 2-7) and the Control treatment (Column 1). Errors are robust. * – $p < 0.1$; ** – $p < 0.05$; *** – $p < 0.01$.

ences in thresholds between Conditions 2 and 1, at least from the perspective of the regression analysis, are not different from the Control treatment. In other words, we do not detect any significant effect of responsibility between individual Condition 1 and Condition 2 where one other person is involved. Nonetheless, we do find a significant difference between the differences in thresholds for risk loving subjects with individual threshold of 3. The rank-sum test is significant ($p = 0.0486$, stars on the right graph of Figure 1). This means that risk loving subjects tend to become more risk averse when choosing for one other (beyond the regression to the mean), though it should be mentioned that this conclusion is based on very few data points.

Next we turn to the analysis of the effect of the number of others on the number of safe choices. The left graph on Figure 2 shows the regression of the difference in thresholds between Conditions 3 and 2 (t32) on the threshold in Condition 2.[13] Again we see the significant effect: subjects tend to choose thresholds closer to the values 5 and 6. Right graph on Figure 2 shows the comparison of the average differences between the Main and Control treatments. We do not observe any significant difference in differences between the Control treatment and the difference in Condition 3-2 except for the threshold of 7 where risk averse subjects seem to choose significantly riskier options for three others than for one other as compared to the regression to the mean (rank-sum test $p = 0.0507$, star on the right graph of Figure 2). The regression in Column 3 of Table 5 shows no significant difference in the coefficient on hl between treatments (the coefficient on control·hl is insignificant), which means that we cannot conclude that there is an effect of the number of others on responsibility.

It should be noted, however, that with the non-parametric analysis the picture is somewhat different. Table 6 reports the comparisons of the distributions of differences in thresholds between conditions in the Main treatment and the Control treatment. The mean difference for

the thresholds higher or equal to 6. The results are unchanged.

[13]The regression is reported in Column 4 of Table 13 in Appendix D.

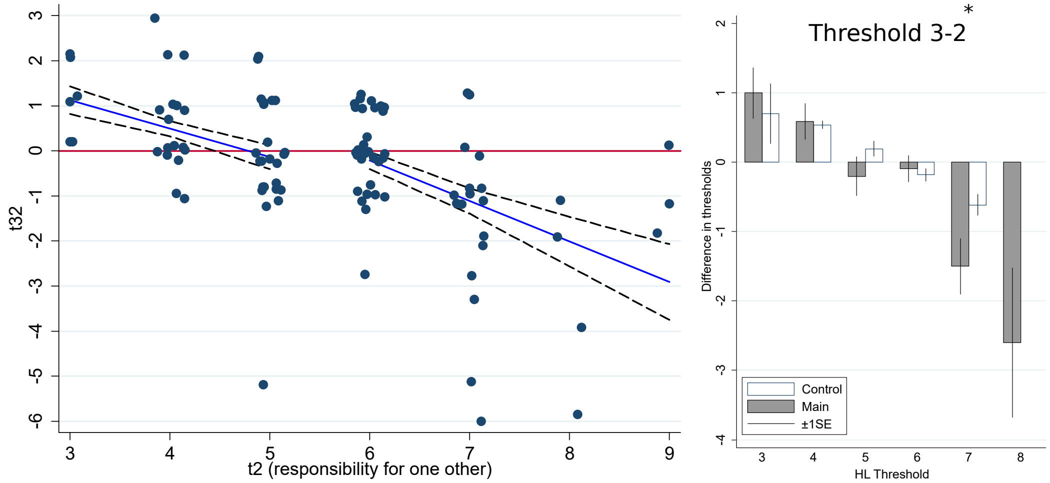

Figure 2: **Left graph.** Regression of t32 on t2, dummy for t2 being higher or equal to 6, and their interaction. The data points are jittered. **Right graph.** Average differences of the thresholds between Conditions 3 and 2, and the differences in the Control treatment. Stars stand for the significance of the rank-sum tests. * – $p < 0.10$.

thresholds higher or equal to 6 in the Control treatment is equal to -0.329, and the mean difference between Conditions 3 and 2 (for thresholds higher or equal to 6) is -0.635.[14] Subjects in the Main treatment decrease the number of safe choices by higher amount than in the Control treatment. Kolmogorov-Smirnov test shows significant difference of distributions ($p = 0.020$). This means that, from the perspective of the non-parametric analysis, we can conclude that there is a significant effect of responsibility for three rather than one person on the risk averse subjects who choose riskier options as the number of others goes up.

Overall, with the regression analysis we cannot reject the hypothesis that the difference in thresholds between Conditions 3 and 2 comes from the regression to the mean. The non-parametric analysis suggests, though, that risk averse subjects are influenced by the number of others they are responsible for as the distributions of differences in thresholds are different between the Main and Control treatments.

Lastly, we turn to the analysis of mutual responsibility. Left graph on Figure 3 shows the regression of the difference of thresholds between Conditions 4 and 3 on the threshold in Condition 3.[15] The right graph of Figure 3 shows significant difference in differences between the Main and Control treatments for thresholds 3 and 7 (rank-sum tests $p = 0.0398$ and $p = 0.0148$ correspondingly). This suggests a strong effect of communication on the choice of threshold in Condition 4. This conclusion is supported by the regression analysis: the regression in Column 4 in Table 5 shows significant difference in the coefficient on hl between the treatments (the coefficient on control·hl is 0.361***). This directly shows that communication has a strong effect on

[14]The means reported in Table 6 and Table 3 are different because in Table 3 all data points are used, whereas in Table 6 the data with thresholds 8 and 9 are discarded in order to make the support of the distributions of the thresholds in the Main and Control treatments the same.

[15]The regression is in Column 6 of Table 13 in Appendix D.

Condition	N obs	Mean	SE	KS exact p
Control				
t ≤ 5	42	0.345	(0.089)	
t ≥ 6	44	-0.329	(0.084)	
Main t21				
t1 ≤ 5	52	0.500	(0.179)	0.120
t1 ≥ 6	48	-0.104^{***}	(0.158)	0.006
Main t32				
t2 ≤ 5	48	0.229	(0.186)	0.100
t2 ≥ 6	52	-0.635^{**}	(0.213)	0.020
Main t43				
t3 ≤ 5	53	0.509^{**}	(0.165)	0.033
t3 ≥ 6	47	-0.553^{***}	(0.163)	0.005

Table 6: Means of differences in thresholds in the Control and Main treatments. Only data points with thresholds in the interval $[3, 7]$ are considered (the support of the thresholds in the Control treatment). Thresholds 8 and 9 in the Main treatment were discarded. The last column shows p-values of the Kolmogorov-Smirnov tests that compare distributions of differences in thresholds in the Main treatment and differences in the Control treatment. ** – $p < 0.05$; *** – $p < 0.01$.

the choice of threshold which cannot be explained simply by the regression to the mean. The non-parametric analysis in Table 6 demonstrates the same result. While the mean differences in thresholds in the Control treatment are 0.345 (for thresholds ≤ 5) and -0.329 (for thresholds ≥ 6), between Conditions 4 and 3 the differences are 0.509 and -0.553. Both are higher in absolute value than in the Control treatment. The distributions of differences are significantly different for thresholds below/equal to 5 ($p = 0.033$), and above/equal to 6 ($p = 0.005$). Thus, communication influences the choices of our subjects significantly more than would be expected from the regressions to the mean. The fitted regression on the left graph of Figure 3 shows that the difference in the thresholds between Conditions 4 and 3 are such that the threshold of 5 is targeted by all risk loving subjects and the threshold of 6 by all risk averse subjects. This is consistent with the average risk preferences in the population (see Figures 4 and 5 in Appendix D).

To substantiate the result concerning mutual responsibility even further we provide analogous analysis for the differences between Condition 4 and Conditions 1 and 2. Columns 6 and 7 in Table 5 show the regressions. The coefficient on the interaction term $\mathsf{control}\cdot\mathsf{hl}$ is significant in both cases. The same is true for the Kolmogorov-Smirnov tests reported in Table 14 in Appendix D. They show significant difference of distributions of differences of thresholds. This demonstrates that the effect of mutual responsibility is indeed strong and works equally well for the risk averse and the risk loving subjects. Both groups choose closer to the average risk preferences when mutually responsible.

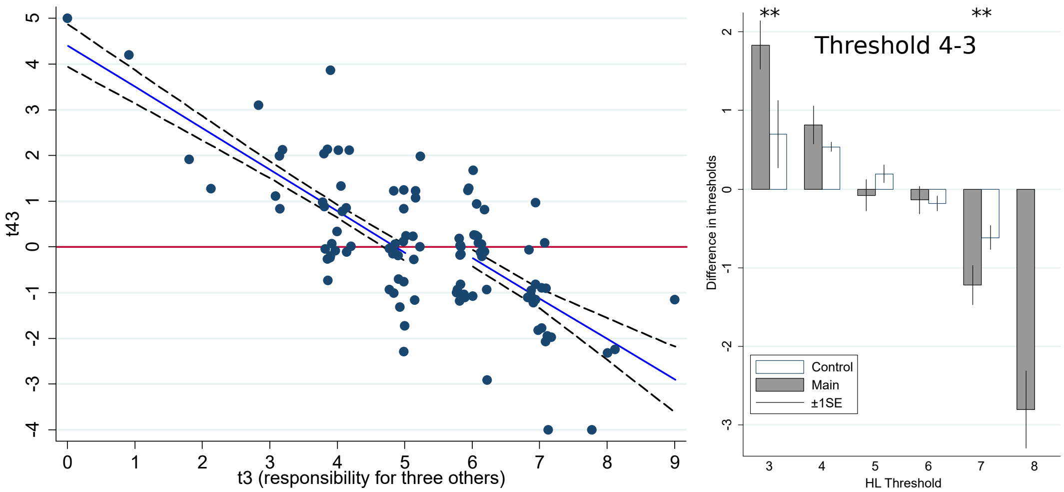

Figure 3: **Left graph.** Regression of t43 on t3, dummy for t3 being higher or equal to 6 and their interaction. The data points are jittered. **Right graph.** Average differences of the thresholds between Conditions 4 and 3, and the differences in the Control treatment. Stars stand for the significance of the rank-sum tests. ** – $p < 0.05$.

4.4 Power Analysis

Regressions in Table 5 show that the effects of responsibility for one other and three others are not significantly higher than the effect of the regression to the mean. However, the coefficients on the variable hl are still higher than the coefficient in the Control treatment: regression to the mean gives coefficient -0.355 whereas the slopes of responsibility for one other and three others are -0.404 and -0.555 respectively. This suggests that the effects of responsibility might be there, but we cannot detect them due to a small sample size.[16]

The goal of this section is to estimate the amount of data that should be collected in order to make small effects of responsibility in Conditions 2 and 3 significant. At least two ways of such power analysis are possible: 1) make parametric assumptions about the heterogeneity of preferences in the population and, assuming effect sizes from Table 5, estimate the sample size needed to make these effects significant; 2) make no assumptions and perform bootstrapping with the existing data. We decided to follow the latter approach since the former requires many assumptions that eventually would be ad hoc, given that we do not have enough information on the distributions of traits in our subject pool.

In the regressions in Columns 2 and 3 of Table 5 we are interested in the coefficient on the interaction term control·hl, which is insignificant in both regressions. Assuming that the original estimations are unbiased, the question we ask is How many observations do we need in order to

[16]Our experiment was particularly expensive. On average we paid around €27 per subject. This was necessary since when subjects chose for others they needed to be sure that the others actually received the money. Thus, each subject received four payments for herself (in four conditions) plus at least three or four payments from being passive money receiver for other subjects.

obtain the 5% significance of these treatment effects? We draw various numbers of observations with replacement from our data and, for each dataset thus generated, estimate the coefficients in the regression. The procedure is repeated 1000 times. For each repetition we obtain the *t*-statistic for the coefficient on the interaction control·hl and calculate the proportion of repetitions in which it is above 1.96, or, in other words, the percentage of cases when we can reject the null hypothesis that the coefficient on the interaction term is zero.

For the regression in Column 2 of Table 5 (the effect of responsibility for one other as compared to individual choice) we determined that we need 75 times more observations than we have (around 7500) in order to reject the null hypothesis of no treatment effect in 96% of cases (if the effect is there at all). For the regression in Column 3 (the effect of responsibility for three others as compared to the choice for oneself and one other) we need 5 times more observations (500 subjects) to reject the null hypothesis in 96% of cases. These estimations show that it is not reasonable to assume that the effect of responsibility for one other is real, whereas the absence of effect of three others might stem from small sample size. Future experiments can clarify this issue.

4.5 Choice in Mutual Responsibility Condition

In this section we analyze the influence that the choices of the two negotiators in Condition 3 have on the number of safe choices they end up agreeing on in Condition 4. The question is, given the number of safe choices of the two negotiators in Condition 3, Which number of safe choices do they end up choosing in Condition 4? We hypothesize that the negotiation process should lead to a consensus so that each pair's choice lies in between the individual choices in Condition 3. Indeed, 81% of the pairs choose the threshold in Condition 4 in between their individual choices in Condition 3. Notice that this is true even though half of the pairs choose in Condition 4 before Condition 3.

We analyze the data further to support this finding. We consider three variables: t4, the agreed number of safe choices of each pair; lowt3 and hight3, the lowest and the highest choices in Condition 3 (t3) in the pair. For example, if in Condition 3 one partner in a pair chose 4 safe choices and another, say, 7, then lowt3 is equal to 4 and hight3 is equal to 7. We would like to know how the choice of the pair depends on the partners' choices when they choose alone.

Column 1 in Table 7 reports the means of these variables. The average choice in Condition 4 lies in between the two choices of the negotiators in Condition 3. Column 2 of Table 7 reports the regression of t4 on lowt3 and hight3. Notice that the coefficients add up to 0.933, which is very close to 1. This implies that the negotiators in the pair in most cases (81% of the data) agree on a number of safe choices that lies in between their number of safe choices in Condition 3. Moreover, this choice is closer to the more cautious negotiator (around 60% weight). This means that it is "simpler" for the risk averse subjects to talk risk loving subjects into choosing

	Mean (1)	**t4** (2)
t4	5.307 (0.161)	
lowt3	4.654 (0.180)	0.344** (0.153)
hight3	6.192 (0.158)	0.589*** (0.124)
N pairs	52	52
R^2		.96

Table 7: **Left column.** Means of the choices in Condition 4 and the pair of negotiators' highest and lowest number of safe choices in Condition 3. **Right column.** OLS regression of the choice under mutual responsibility (t4) as dependent on the pair's highest and lowest number of safe choices in Condition 3. Constant term is suppressed. Errors are robust. ** – $p < 0.05$; *** – $p < 0.01$.

closer to their risk preference than other way round. However, we should not forget that mutual responsibility makes the choices of risk averse subjects much more riskier in Condition 4 than in Condition 3 as the regression analysis in Section 4.3 confirms. All these results are in line with the blame hypothesis: it is easier to make risky choices for the group when there are two people to blame and not one.

5 Conclusion

Our results can be summarized as follows. We did not find an effect of responsibility for one other person on the choices in Holt and Laury tasks, as compared with individual choice, once regression to the mean was controlled for. We did find some support for the hypothesis that the number of others affected by the choice matters: risk averse subjects seem to choose significantly riskier options (beyond the regression to the mean) when the number of others grows. Importantly, we find strong effect of mutual responsibility. When two subjects communicate in order to choose for themselves and three others they demonstrate both cautious and risky shifts towards the population mean risk preferences.

The mutual responsibility result is in line with our blame hypothesis: the cautious and risky shifts are symmetric and stronger than in the case of individual choice for others. Unlike previous studies, we find risky shift when we compare the choices for one and three others, but not cautious shift. The reason for this might be the control for the regression to the mean, which is crucial in order to make correct conclusions about the changes in behavior. Most likely, it is also responsible for the absence of the effect of responsibility for one other. More experiments with good controls are necessary in order to clarify the effects of responsibility on risky choices.

References

ANDERSSON, O., H. J. HOLM, J.-R. TYRAN, AND E. WENGSTRÖM (2014): "Deciding for Others Reduces Loss Aversion," *Management Science*.

BAKER II, R. J., S. K. LAURY, AND A. W. WILLIAMS (2008): "Comparing Small-Group and Individual Behavior in Lottery-Choice Experiments," *Southern Economic Journal*, 75(2), 367–382.

BOLTON, G. E., A. OCKENFELS, AND J. STAUF (2015): "Social responsibility promotes conservative risk behavior," *European Economic Review*, 74, 109–127.

CHAKRAVARTY, S., G. W. HARRISON, E. E. HARUVY, AND E. E. RUTSTRÖM (2011): "Are You Risk Averse over Other People's Money?," *Southern Economic Journal*, 77(4), 901–913.

CHARNESS, G., AND M. O. JACKSON (2009): "The role of responsibility in strategic risk-taking," *Journal of Economic Behavior and Organization*, 69, 241–247.

DAVIS, J. H. (1992): "Some Compelling Intuitions about Group Consensus Decisions, Theoretical and Empirical Research, and Interpersonal Aggregation Phenomena: Selected Examples, 1950-1990," *Organizational Behavior and Human Decision Processes*, 52, 3–38.

ELIAZ, K., D. RAY, AND R. RAZIN (2006): "Choice Shifts in Groups: A Decision-Theoretic Basis," *American Economic Review*, 96(4), 1321–1332.

ERTAC, S., AND M. Y. GURDAL (2012): "Deciding to decide: Gender, leadership and risk-taking in groups," *Journal of Economic Behavior and Organization*, 83, 24–30.

FISCHBACHER, U. (2007): "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10(2), 171–178.

GARDNER, M., AND L. STEINBERG (2005): "Peer Influence on Risk Taking, Risk Preference, and Risky Decision Making in Adolescence and Adulthood: An Experimental Study," *Developmental Psychology*, 41(4), 625–635.

GURDAL, M. Y., J. B. MILLER, AND A. RUSTICHINI (2013): "Why Blame?," *Journal of Political Economy*, 121(6), 1205–1247.

HARRISON, G. W., AND E. E. RUTSTRÖM (2008): "Risk Aversion in Laboratory," in *Risk Aversion in Experiments*, ed. by J. C. Cox, and G. W. Harrison, pp. 41–196. Emerald Group Publishing Limited.

HOLT, C. A., AND S. K. LAURY (2002): "Risk Aversion and Incentive Effects," *American Economic Review*, 92(5), 1644–1655.

KAHNEMAN, D., AND A. TVERSKY (1984): "Choices, values, and frames," *American Psychologist*, 39(4), 341.

MASCLET, D., N. COLOMBIER, L. DENANT-BOEMONT, AND Y. LOHÉAC (2009): "Group and individual risk preferences: A lottery-choice experiment with self-employed and salaried workers," *Journal of Economic Behavior and Organization*, 70, 470–484.

PAHLKE, J., S. STRASSER, AND F. M. VIEIDER (2012a): "Responsibility effects in decision making under risk," WZB Discussion Paper, No. SP II 2012-402.

——— (2012b): "Risk-taking for others under accountability," *Economics Letters*, 114, 102–105.

PRUITT, D. G. (1971): "Choice shifts in group discussion: An introductory review," *Journal of Personality and Social Psychology*, 20(3), 339–360.

REYNOLDS, D. B., J. JOSEPH, AND R. SHERWOOD (2009): "Risky Shift Versus Cautious Shift: Determining Differences In Risk Taking Between Private And Public Management Decision-Making," *Journal of Business and Economics Research*, 7(1), 63–78.

ROHDE, I. M. T., AND K. I. M. ROHDE (2011): "Risk attitudes in a social context," *Journal of Risk and Uncertainty*, 43, 205–225.

SHUPP, R. S., AND A. W. WILLIAMS (2008): "Risk Preference Differentials of Small Groups and Individuals," *The Economic Journal*, 118(525), 258–283.

STONER, J. A. F. (1961): "A comparison of individual and group decisions involving risk," Master's thesis, Massachusetts Institute of Technology.

VIEIDER, F., C. VILLEGAS-PALACIO, P. MARTINSSON, AND M. MEJÍA (2015): "Risk taking for oneself and others: A structural model approach," WZB Discussion Paper, No. SP II 2015-401.

WALLACH, M. A., N. KOGAN, AND D. J. BEM (1964): "Diffusion of Responsibility and Level of Risk Taking in Groups," *Journal of Abnormal Psychology*, 68, 263–274.

Appendix (for online publication)

A Details of the Design

A.1 The Payoffs in HL Tasks

	LOTTERY A		LOTTERY B	
	Payoff A1	**Payoff A2**	**Payoff B1**	**Payoff B2**
HL task 1	2.00	1.60	3.85	0.10
HL task 2	1.85	1.40	3.45	0.10
HL task 3	2.30	1.60	4.00	0.20
HL task 4	2.45	1.20	3.55	0.30

Table 8: The payoffs used in the four HL tasks. Probabilities were fixed and went from 0.1 to 1 with 0.1 increment for outcomes A1 and B1, and from 0.9 to 0 for outcomes A2 and B2.

A.2 The Order of HL Tasks

Subject	Period 1	Period 2	Period 3	Period 4
1	1	2	3	4
2	4	2	1	3
3	2	1	3	4
4	1	3	4	2
5	2	4	1	3
6	1	4	2	3
7	4	3	1	2
8	2	4	1	3
9	1	3	4	2
10	2	4	3	1
11	3	1	2	4
12	4	1	3	2
13	2	4	3	1
14	3	1	4	2
15	3	2	4	1
16	4	2	3	1
17	4	3	2	1
18	1	3	2	4
19	2	3	1	4
20	2	4	3	1
21	4	2	1	3
22	4	3	2	1
23	3	2	1	4
24	3	4	1	2

Table 9: The time sequence of the four HL tasks that were presented to each of the 24 subjects in a session for their four choices.

A.3 The Matching of Subjects

Subject	1other	3others-1	3others-2	3others-3	Mutual-1	Mutual-2	Mutual-3
1	2	3	4	5	22	23	24
2	3	4	5	6	10	11	12
3	4	5	6	7	22	23	24
4	5	6	7	8	7	8	9
5	6	7	8	9	13	14	15
6	7	8	9	10	13	14	15
7	8	9	10	11	1	2	3
8	9	10	11	12	10	11	12
9	10	11	12	13	1	2	3
10	11	12	13	14	4	5	6
11	12	13	14	15	16	17	18
12	13	14	15	16	16	17	18
13	14	15	16	17	1	2	3
14	15	16	17	18	7	8	9
15	16	17	18	19	1	2	3
16	17	18	19	20	4	5	6
17	18	19	20	21	19	20	21
18	19	20	21	22	19	20	21
19	20	21	22	23	4	5	6
20	21	22	23	24	4	5	6
21	22	23	24	1	7	8	9
22	23	24	1	2	7	8	9
23	24	1	2	3	10	11	12
24	1	2	3	4	10	11	12

Table 10: The matching of the recipients of payoffs for the 24 subjects in the session. The entries in the table represent the subjects' assigned numbers.

A.4 The Order of Conditions

Subject	Period 1	Period 2	Period 3	Period 4
1	1	2	3	4
2	1	2	4	3
3	1	3	2	5
4	1	3	6	2
5	1	4	2	3
6	1	5	3	2
7	2	1	3	6
8	2	1	5	3
9	2	3	1	7
10	2	3	8	1
11	2	6	1	3
12	2	7	3	1
13	3	1	2	8
14	3	1	7	2
15	3	2	1	9
16	3	2	9	1
17	3	8	1	2
18	3	9	2	1
19	4	1	2	3
20	5	1	3	2
21	6	2	3	1
22	7	2	3	1
23	8	3	1	2
24	9	3	2	1

Table 11: The time sequence of conditions presented to each of the 24 subjects in the session. Here 1, 2, and 3 mean Conditions 1, 2, and 3. For mutual responsibility condition the pairs of negotiators are $(4,5)$, $(6,7)$, and $(8,9)$.

B Additional Graphs

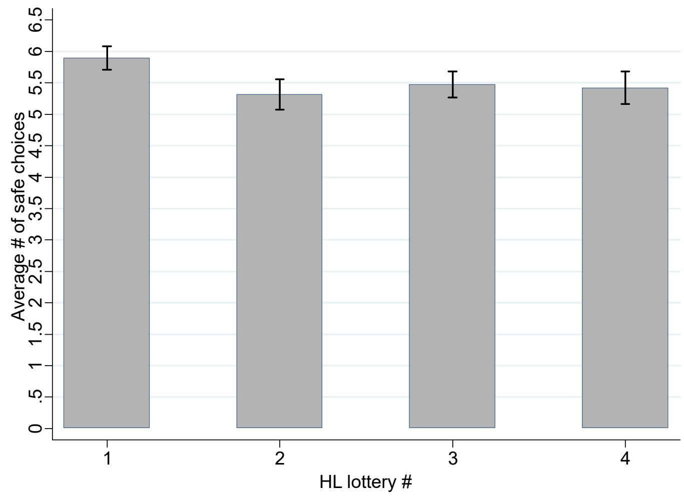

Figure 4: Average number of safe choices in the four HL tasks in the Control treatment. The spikes are ±1 SE.

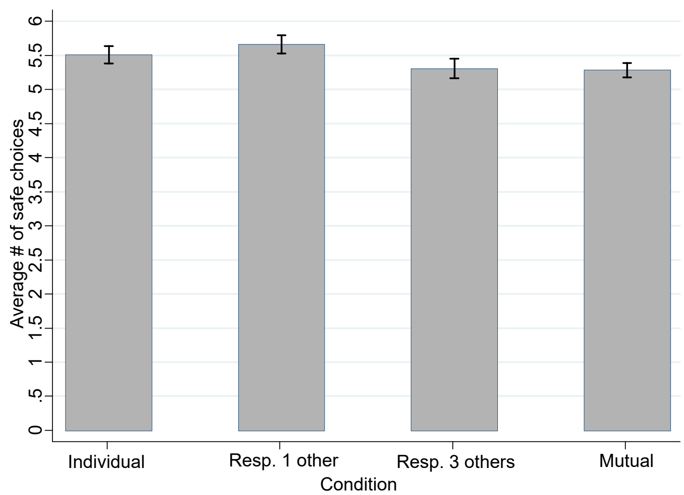

Figure 5: Average number of safe choices in the four conditions in the Main treatment. The spikes are ±1 SE.

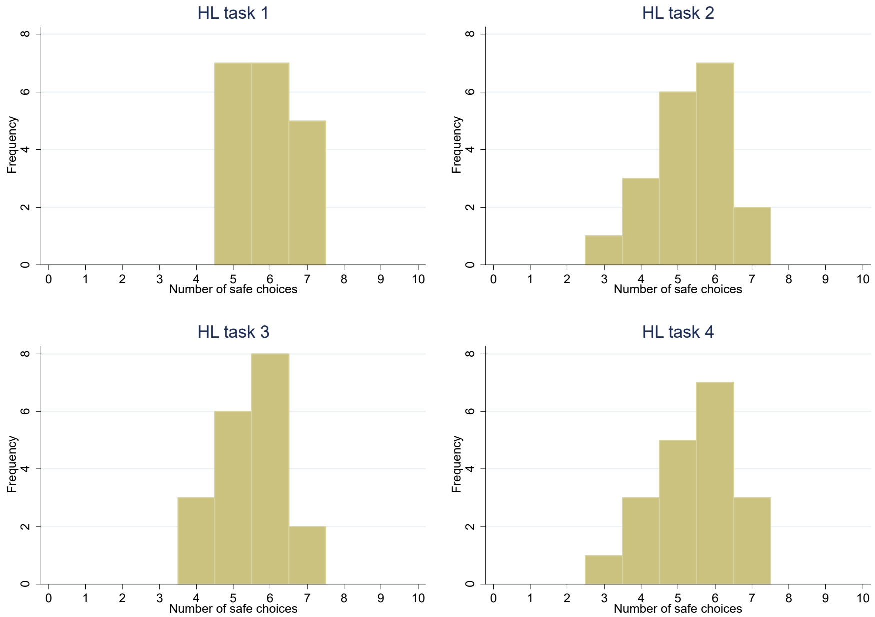

Figure 6: The histograms of the number of safe choices in four HL tasks in the Control treatment.

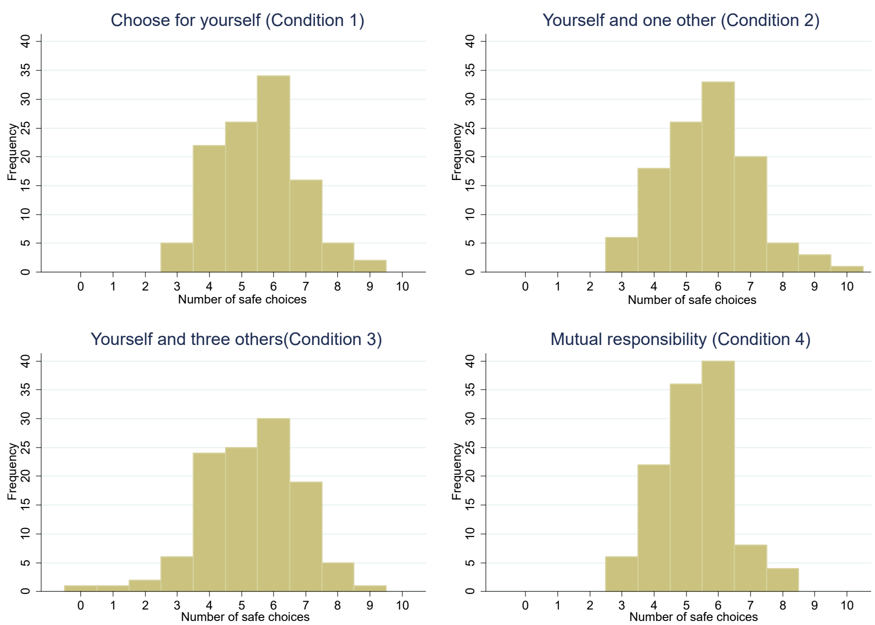

Figure 7: The histograms of the number of safe choices in four conditions of the Main treatment.

C Regression Variables

Variable	Range	Definition
hl1-4	$\{0, ..., 10\}$	The number of safe choices in four HL tasks in the Control treatment
hl	$\{0, ..., 10\}$	The number of safe choices in different conditions depending on the context of the dependent variable
t1-4	$\{0, ..., 10\}$	The number of safe choices in the HL tasks in Conditions 1-4 in the Main treatment
control	0/1	Dummy that indicates the Control treatment
tij	$\{-10, ..., 10\}$	Equals ti - tj, where *i* and *j* are the numbers of Conditions in the Main treatment. The difference in the numbers of safe choices for each subject
geq6	0/1	The dummy that is one if a subject chose the number of safe choices greater or equal to 6
lowt3	$\{1, ..., 10\}$	In Condition 4 (mutual responsibility), the lowest threshold of the two negotiators
hight3	$\{1, ..., 10\}$	In Condition 4 (mutual responsibility), the highest threshold of the two negotiators

Table 12: Variables used in the regressions.

D Additional Analyses

	Control	t21	t21	t32	t32	t43	t43
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
geq6	0.844	–0.604	–0.604	1.372	1.372	1.443	1.443
	(1.326)	(1.939)	(1.943)	(2.007)	(2.011)	(1.730)	(1.734)
hl	–0.276*	–0.672**	–0.672**	–0.654***	–0.654***	–0.898***	–0.898***
	(0.161)	(0.279)	(0.280)	(0.228)	(0.229)	(0.107)	(0.108)
geq6·hl	–0.159	0.203	0.203	–0.142	–0.142	–0.107	–0.107
	(0.237)	(0.361)	(0.362)	(0.359)	(0.360)	(0.280)	(0.280)
control			–1.859		–1.515		–2.814***
			(1.439)		(1.213)		(0.868)
control·geq6			1.448		–0.529		–0.599
			(2.350)		(2.407)		(2.180)
control·hl			0.396		0.378		0.622***
			(0.323)		(0.280)		(0.193)
control·geq6·hl			–0.362		–0.017		–0.052
			(0.433)		(0.431)		(0.367)
cons	1.589**	3.448***	3.448***	3.104***	3.104***	4.403***	4.403***
	(0.742)	(1.231)	(1.233)	(0.959)	(0.961)	(0.453)	(0.454)
N	86	107	193	108	194	110	196

Table 13: OLS regressions of differences in thresholds between conditions in the Main treatment (Columns 2-7) and the Control treatment (Column 1). Errors are robust. * – $p < 0.1$; ** – $p < 0.05$; *** – $p < 0.01$.

Condition	N obs	Mean	SE	KS exact p
Control				
t $\leq$ 5	42	0.345	(0.089)	
t $\geq$ 6	44	-0.329	(0.084)	
Main t31				
t1 $\leq$ 5	50	0.120^{*}	(0.168)	0.055
t1 $\geq$ 6	50	-0.280^{**}	(0.159)	0.014
Main t41				
t1 $\leq$ 5	52	0.423^{*}	(0.141)	0.076
t1 $\geq$ 6	49	-0.551^{**}	(0.160)	0.017
Main t42				
t2 $\leq$ 5	49	0.429^{**}	(0.146)	0.044
t2 $\geq$ 6	51	-0.746^{***}	(0.128)	0.001

Table 14: Means of differences in thresholds in the Control and Main treatments. Only data points with thresholds in the interval $[3, 7]$ are considered (the support of the thresholds in the Control treatment). Thresholds 8 and 9 in the Main treatment were discarded. The last column shows p-values of the Kolmogorov-Smirnov tests that compare distributions of differences in thresholds in the Main treatment and differences in the Control treatment. * – $p < 0.10$; ** – $p < 0.05$; *** – $p < 0.01$.

E Instructions

General Instructions

You are now participating in a decision making experiment. If you follow the instructions carefully, you can earn a considerable amount of money depending on your decisions and the decisions of the other participants. Your earnings will be paid to you in CASH at the end of the experiment. In addition to your earnings in the experiment you will receive a €2 show-up fee. During the experiment you are not allowed to communicate with anybody. In case of questions, please raise your hand. Then we will come to your seat and answer your questions. Any violation of this rule excludes you immediately from the experiment and all payments. This research is funded by the Marie Curie action of the EU.

Click OK when you are ready to go on.

***

Example Scenario

In this experiment you will be presented with four scenarios. In each scenario you will need to choose 10 times between the Left and Right option. Each option is a lottery. A lottery consists of two monetary outcomes each of which can happen with certain probability.

On the right side of the screen, you see an example scenario. There are 10 rows which correspond to the 10 choices described above. Every row represents a choice between two lotteries.

You can make a choice in each row by clicking on the lottery that you prefer. The lottery that you have chosen becomes red.

After you have made the 10 choices, the computer will select one of them randomly. Then, the lottery chosen by you will be 'played' by the computer, and you will receive the corresponding payoff.

To demonstrate how this works in this example scenario, please make 10 choices and press OK button to validate. Next, you will see a screen that shows the amount of money that you earned in this scenario. SINCE THIS IS AN EXAMPLE, YOU WILL NOT ACTUALLY RECEIVE THE AMOUNT YOU SEE ON THE NEXT SCREEN. You will NOT see the amount of money you earn after each scenario. You will be informed about your earnings in the end of the experiment.

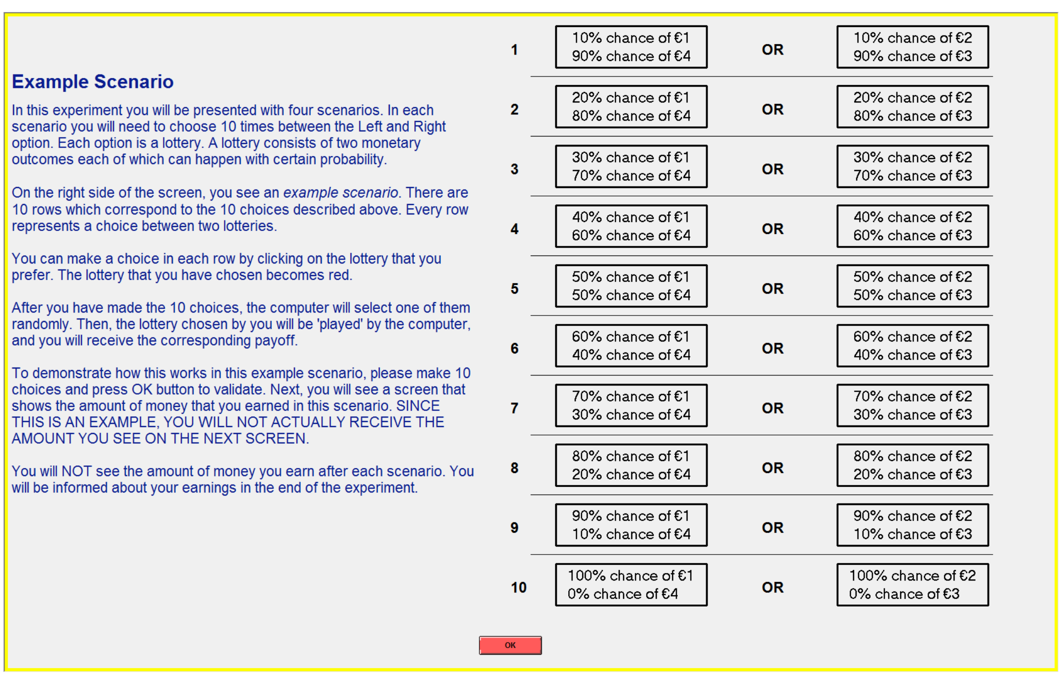

Figure 8: A screenshot of the example scenario.

***

Choosing for Yourself
In this scenario, you are the only one who is affected by your choice. You are the only person that receives the payoff of one of the lotteries that you choose. This is exactly the same as the example scenario.

***

Choosing for Yourself and One Other Person
In this scenario, there is one other person in this room who is affected by your choices. So, you are not the only person who receives the payoff of one of the lotteries that you choose. You, and the other person will receive the same payoff (the payoff is not divided). For example, if computer randomly determines your payoff to be €2, then the other person will also receive €2. After the experiment the person affected by your choice will be asked how satisfied he/she is with the amount of money he/she received.

***

Choosing for Yourself and Three Other People
In this scenario, there are three other people in this room who are affected by your choices. So, you are not the only person who receives the payoff of one of the lotteries that you choose. You, and the three other people will receive the same payoff (the payoff is not divided). For example, if computer randomly determines your payoff to be €2, then the three others will also receive €2. After the experiment the others affected by your choice will be asked how satisfied they are with the amount of money they received.

***

Choosing with Someone Else for Both of You and Three Other People
In this scenario, you decide together with someone else and three other people are affected by your choice. Therefore you decide together which choices you would like to make.

You can communicate with your partner via the chatbox below. You can type the messages to your partner in the text line and press enter to send. You should agree on each of the 10 choices. Both of you should click on exactly the same lotteries in order to proceed to the next scenario. After you both have made your choices, press the OK button to proceed.

After you click OK, the computer will randomly determine the payoff. You, your partner and three other people will receive the same amount of money. For example, if computer randomly determines the payoff to be €2, then you, your partner and the three others will each receive €2.

After the experiment the others affected by your choice will be asked how satisfied they are with the amount of money they received.

***

In every scenario a reminder was shown which repeated the general explanation of the experiment.
Reminder
In this scenario you need to choose 10 times between the Left and Right option. Each option is a lottery. A lottery consists of two monetary outcomes each of which can happen with certain probability.

On the right side of the screen, you see 10 rows which correspond to the 10 choices described above. Every row represents a choice between two lotteries.

You can make a choice in each row by clicking on the lottery that you prefer. The lottery that you have chosen becomes red.

After you have made the 10 choices, the computer will select one of them randomly. Then, the chosen lottery will be 'played' by the computer, and you will receive the corresponding payoff.