Co-Design of innovative contract models for agri-environment and climate measures and the valorisation of environmental public goods

Environmental Cooperation at Landscape Scales: First Insights from Co-Designing Public Goods Games with Farmers in Four EU Member States

Milestone 26 / 5.2

Author:	Jens Rommel (SLU)
Co-Authors:	Lenny van Bussel (WUR), Solen le Clech (WUR), Mikołaj Czajkowski (UW), Julia Höhler (WUR), Bettina Matzdorf (ZALF), Boldizsár Megyesi (ESSRG), Julian Sagebiel (SLU), Christoph Schulze (ZALF), Vera Wechner (Humboldt-Universität zu Berlin), Katarzyna Zagórska (UW)

Contributions from:	WP3, WP4
Work Package:	WP5
Date of delivery:	Contractual: 30/04/2021 Actual: 30/04/2021
Total number of pages:	43
Keywords:	Experimental Economics; Behavioural Economics; Collaboration; Coordination; Collective Agri-environmental Contracts

This project has received funding from the European Union`s Horizon 2020 research and innovation programme under grant Agreement No. 818190.

EXECUTIVE SUMMARY

In this milestone report, we explain how we have developed public goods games to perform an ex-ante assessment of novel collective contract models in the Contracts2.0 project. Workshops were conducted in Germany, Hungary, the Netherlands, and Poland. The first data collection was completed in Germany, and an expert prediction survey was run in parallel to the public goods game with German farmers. The overall experiences from the workshops have been positive. The public goods game was met with great interest from stakeholders, albeit in all instances, there were concerns about the level of abstraction of the game. Another frequent concern was parallelism, i.e., the link between game results and real-world behaviour. We used 358 completed online responses from German farmers for an initial analysis. Farmers' behaviour in our study differed substantially from participants in the laboratory. Overall levels of cooperation among farmers were substantially higher than one would expect from previous laboratory studies. In addition, treatment effects were not in the expected direction. The only treatment that showed substantially larger contributions was to emphasize the social optimum of the game. Expert predictions were more in line with the literature from experimental laboratory studies than with the actual behaviour of farmers. Among the experts, those indicating good knowledge on the public goods game, predicted more accurately, whereas stated sector-specific knowledge (on agriculture, the common agricultural policy, or agri-environmental schemes) did not substantially improve predictions.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The Contracts2.0 project wants to **improve the environmental, economic, and social performance of agri-environmental contracts**, involving farmers, policy-makers, researchers, and other food system actors. The project is organised around Contract Innovation Labs (WP3) and Policy Innovation Labs (WP4) as major exchange platforms among these actors. More details on this process, the policy context in different member states, and the overall conceptual framework are available in other project reports (Andersen et al., 2020; Prager et al., 2020; Wanner et al., 2020).

WP5 interacts with the **Contract and Policy Innovation Labs** by providing an ex-ante assessment of novel contracts based on inputs from the labs. In task 5.2 of the project, we study farmers' willingness to cooperate as a crucial prerequisite for collective contracts. By collective contracts, we mean all agri-environmental contracts that involve the collaboration or coordination of farmers, i.e., contracts that go beyond the level of a single farm. That is, collective contracts encompass at least some form of interaction among farms. We use public goods games that are co-designed in interaction with the labs.

Collective contracts can take many forms. Farmers may receive bonus payments for jointly addressing environmental goals. They may receive bonus payments if a pre-defined environmental threshold is achieved, or agri-environmental contracts may be organised entirely around farmer-driven initiatives or organisations. For instance, in the Netherlands, agri-environmental contracts are exclusively administered through farmer collectives (cf. Bouma et al., 2020), which has fuelled a **debate on the advantages and disadvantages of collective contracts across EU member states**. For instance, the German Scientific Advisory Board on Agricultural Policy, Food and Consumer Health Protection has recently recommended that Germany should "[i]mprove the institutional prerequisites for collectively organised agri-environment-climate protection. (1) Examine the extent to which elements of the Dutch system of collective nature conservation arrangements could also be applicable in Germany; (2) improve the institutional prerequisites for the implementation of collective models of environmental and climate action; (3) in pilot projects in the current finance period, support the grouping of relevant local actors into 'biodiversity-generating communities'" (Latacz-Lohmann et al., 2019).

The scientific literature has dealt with collective and agri-environmental contracts extensively (see Kuhfuss, 2019 for an overview of the literature). Research, based on social-ecological models has emphasized the **importance of coordination among farms** to more effectively address environmental goals. For example, it could be important to reduce habitat fragmentation for achieving biodiversity targets (Parkhurst et al., 2002). There is a growing awareness of the complex drivers of farmers' behaviour and a call for more policy-oriented experimental research (Colen et al., 2016; Dessart et al., 2019). Thus, recent research has adopted behavioural and experimental approaches to provide causal evidence on the individual drivers of farmers' adoption of environmentally friendly practices (Kuhfuss et al., 2016; Thomas et al., 2019; Villamayor-Tomas et al., 2019; Dessart et al., 2021). While these studies focus on individual behaviour or preferences, studies investigating coordination and strategic interaction in the landscape rely on student participants (e.g., Banerjee et al., 2014) or field populations distant from the European farming context (e.g., Rommel and Anggraini, 2018). One notable exception is the study of Bouma et al. (2020), who conducted a public goods game – a workhorse of Experimental Economics to study cooperation – with farm management students in the Netherlands. In the Contracts2.0 project, we want to **address this paucity of research by providing**

evidence on farmers' willingness to cooperate across different European policy contexts, using the public goods game.

This milestone report **describes our approach to study farmers' willingness to cooperate**. We describe how the public goods games were co-designed in workshops with practitioners in Germany, the Netherlands, Hungary, and Poland. We also illustrate first insights from the games by presenting results from a study in Germany for which the data collection has been completed. Parallel to the German study, we ran a survey asking researchers and practitioners to predict the outcomes of the public goods game in Germany. The report addresses partners in the project interested in understanding the basis for collective contracts and the experimental approach.

This report is organised as follows. We start by providing a short overview of the public goods game and its basis in the Experimental Economics literature. This includes a short overview of treatments studied in Economics laboratories that motivated our research. We outline an expert prediction survey that ran in parallel to the experiment for the German study. Next, we provide an overview of the workshops that were conducted to design the experiments. We then present the first results from the German study and the expert predictions. In a final section, we discuss the results and present an outlook on the next steps.

2. THE PUBLIC GOODS GAME

2.1 The Basic Game

The public goods game has been developed by Isaac et al. (1984) to study free-riding behaviour and collective action problems. The game has quickly developed into a workhorse of Experimental Economics and is frequently used as a tool to study cooperation. In the simplest version of the game, n players must allocate an initial endowment e between a private and a group account. All contributions x to the group account are multiplied by a constant a, satisfying $1 < a < n$. Contributions are then distributed in equal parts to all players. Participant i's profit π_i in a simple one-shot linear voluntary contribution mechanism public goods game is then defined as follows:

$$\pi_i = \frac{a\left(\sum_{j\neq i}^{n-1} x_j + x_i\right)}{n} + e_i - x_i$$

Note that x_j and x_i denote the contributions of $n-1$ players j and player i to the public good, and e_i is the endowment of player i. A Nash equilibrium is a set of strategies under which no single player can increase their payoff from a unilateral change in strategy. Under profit maximization and no preferences over the outcomes of others or the overall distribution, contributing zero is a unique Nash equilibrium, because a player internalises only the fraction $a/n < 1$ of their own contribution. This Nash equilibrium is at odds with the social optimum where all players contribute their entire endowment, which results from the condition that $a > 1$. Typical laboratory experiments match four or five players and use a constant of $a = 2$, but there are large variations, repeated interaction etc. Fig. 1 displays the typical steps of a public goods game. In the example, each player is endowed with 10 Euro and

contributes 5 Euro to the group account. After the contributions are doubled, each player receives an equal share, leaving everyone with 15 Euro in the end.

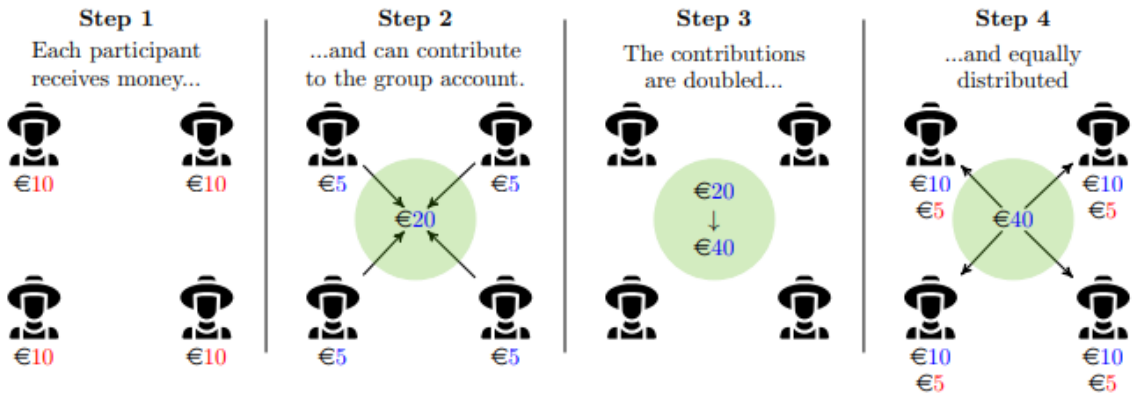

Fig. 1: Schematic representation of steps in a public goods game (Source: own design; icon from flaticon.com by freepik)

From a vast body of literature, encompassing hundreds of laboratory studies (see Ledyard, 1994 and Zelmer, 2003 for an early review and a meta-analysis), a few common patterns emerge. In a simple one-shot game, participants typically contribute about half of their initial endowment, i.e., they do not play the Nash equilibrium. In repeated games, cooperation (defined as contributions to the group account) decreases over time. Notably, if participants are matched with different participants, they start again with high initial contributions upon a restart of the game. One reason for this behaviour may be that in repeated interaction, people can be broadly categorised in three groups with similar behavioural patterns: "free-riders" who contribute little or nothing, "conditional co-operators" who reciprocate the behaviour of others, and "altruists" who contribute independently of others (Fischbacher et al., 2001).

2.2 Considered Treatments

We developed the experiments from a long list of common treatments applied in laboratory experiments. This list was inspired by the laboratory literature, particularly a review article and a meta-analysis (Ledyard, 1994; Zelmer, 2003). The list was then used in expert discussions and the workshops to arrive at treatments for the final study in each country. This section briefly describes the treatments we discussed. An English translation of the paper form we used for the German study is presented in Appendix A.

Manipulations in **group size** were discussed as an option to include. For instance, farmers could be paired in groups of four or groups of eight players. Manipulations in the **marginal per capita return** – the fraction of the public good internalised as a private benefit when contributing – were also discussed. This could mean that instead of doubling contributions to the group account ($a = 2$), one could also use a smaller (e.g., $a = 1.5$) or larger multiplication factor (e.g., $a = 3$). The public good provision could be **risky**. We discussed options in which the public good would only be provided with a certain probability. We discussed two **reward** scenarios. Participants could issue a reward to other players conditioned on their behaviour. There was a costly and a non-costly option. We also discussed **sanctioning/punishment** scenarios. A **heterogeneous endowments** treatment was discussed, meaning that farmers could either be part of a high or low endowment group. The **leading-by-example**

treatment involved sequential decisions. A leader first decides what to contribute, and followers then condition their contributions on what the leader did. We considered different versions of a **threshold** public goods game, i.e., the public good is only provided if a threshold is reached. Finally, we also considered various psychological stimuli treatments that would reframe the meaning of the game. **Social norms** could be studied by providing information that others have contributed large amounts in previous experiments. The **social optimum** or the **individual optimum/Nash equilibrium** could be emphasized. The game could be presented in a **gain or loss frame**.

2.3 Prediction Survey

Predictions of research results can help the research community to establish a baseline of prior beliefs on a subject area (DellaVigna et al., 2019). Predictions of research results are also useful for "improving the interpretation of research results, mitigating bias against null results, and improving predictive accuracy and experimental design" (ibid, p. 428). There are very few attempts to predict research results, and we are not aware of any such attempts in the agricultural social sciences. We designed a survey in which we invited academics and practitioners to predict the outcomes of the public goods game with farmers. Previous studies have primarily focused on the prediction of laboratory experiments by academics (DellaVigna and Pope, 2018a, 2018b).

The prediction survey was conducted online between January and March 2021. Respondents were recruited through various channels including research networks, email-mailing lists, and social media. In addition, various experts (e.g., at DG AGRI) were explicitly contacted to distribute the survey in their networks. The survey was also hosted on the social science prediction platform (*https://socialscienceprediction.org/*), a publicly accessible website to promote prediction studies.

The prediction survey consisted of three sections. In the first section, respondents were introduced to the topic, gave their consent to data protection and received a general explanation of public goods games. In the second section, respondents received explanations on the specific public goods games, including descriptions of the sample and all treatments, following the recommendations from the social science prediction platform. While the explanations were kept short, respondents could attain additional information and examples by clicking on various links that opened PDF files in German and English, containing graphical illustrations and explanatory text. After the explanations, respondents were asked to predict the average contribution to the group account for each of the described treatments, followed by a question on how certain they were about their answers. The section concluded with a ranking question, where respondents were asked to rank the variance of the individual contributions for each treatment. The final block included socio-demographic questions and asked respondents to assess their knowledge. As the sample consisted of experts, we asked about their area of expertise, their educational background, which field they work in, and how they assess their knowledge on public good games and agricultural policies in the European Union.

The survey was incentivised. Out of 100 respondents with a valid email address, we randomly drew one respondent to receive compensation based on the accuracy of the prediction. Selected respondents received 300 Euros minus three times the deviation of a randomly selected prediction (picked from the five treatments). At the end of the questionnaire, respondents had the option to enter their email address to be eligible to participate in the lottery and win a prize and a field to enter comments and suggestions.

3. CO-DESIGN WORKSHOPS IN FOUR MEMBER STATES

We conducted workshops in four member states: Germany, Hungary, the Netherlands, and Poland. We invited representatives with a stake in agri-environmental policy and group contracts. We aimed for about four to ten participants (out of which there should be at least one farmer or farmer representative) to achieve diverse perspectives, while ensuring a good discussion climate. In each country, seven representatives participated.

These workshops aimed to understand stakeholder perspectives on collective contracts and to develop treatments of the public goods games, which are then conducted in each of these countries. Informed consent on the use of the data was obtained from all participants. All workshops followed a common agenda. We started an introduction round of all participants and the project context. Next, we presented the basic idea of a public goods game and included a short exercise that involved playing the game. Fig. 2 displays the contributions to the group account from an initial endowment of 100 Złoty among Polish workshop participants according to the public goods game described in section 2.1.

Fig. 2: Contributions to Group Account among Workshop Participants in Poland (Source: Katarzyna Zagórska; platform: Mentimeter)

Workshop participants then shortly discussed the role of collective contracts in the overall agri-environmental policy architecture at regional, federal, and EU levels. The main part of the workshop was dedicated to discussing the treatments of the public goods game. Participants, with the help of the workshop facilitator, went through a list of treatments, voiced their opinions, and suggested additional treatments. Finally, an anonymous vote was cast after the discussion to also express views quantitatively. While the first workshop used a long list of 17 treatments (Appendix A), the other workshops pre-selected the seven to ten most relevant treatments in a discussion with the local research team. A common baseline version of the experiment will be applied in all countries (although there are some important differences in the overall design of the national case studies). In all

workshops, participants showed great interest in the heterogeneous endowments treatment. Otherwise, the selected treatments differ at country level, depending on the outcomes of the workshops. The study is completed in Germany, and in the final design stages in Hungary and the Netherlands. For the Polish study, we have not yet decided on the treatments and the experimental design. Table 1 provides an overview of the discussed and selected (if appicable) treatments.

Table 1: Overview on discussed and selected Treatments

Treatment	Included in DE Workshop	Included in DE Study	Included in HU Workshop	Included in HU Study	Included in NL Workshop	Included in NL Study	Included in PL Workshop
Baseline version	X	X	X	X	X	X	X
Group size	X		X	X	X		X
Different marginal per capita Return	X				X	X	X
Risky provision	X		X				X
Rewards	X		X		X		X
Sanctions	X		X		X		X
Heterogeneous endowments	X	X	X	X	X	X	X
Leading-by-example	X	X	X				
Threshold	X		X	X	X	X	X
Provide information on behaviour of others (social norms)	X	X	X				
Emphasise social optimum	X	X					
Highlight individual optimum/Nash equilibrium	X						
Gain/loss framing	X						

Source: own design

3.1 Germany

The German workshop was held on 23 January 2020 in Berlin. The workshop lasted approximately three and a half hours. We took advantage of the "International Green Week" – a major fair and event of the agricultural and food sector – that ensured access to a wide range of stakeholders. Participants were two farmers (both male, one organic, one conventional from Western and Northern Germany),

one administrator from the agricultural chamber involved in the administration of contracts (male, state of North Rhine-Westphalia), one representative from the farmers' union (female, Berlin), a representative from the Westphalian Cultural Landscape Foundation (male), a representative from the Rhenish Cultural Landscape Foundation (male), one scientist specialising in incentive-based nature conservation instruments (male), and the research team (four members, two male, two female). From the discussion and voting, it became clear that understanding norms (highlighting what others have done), highlighting the social optimum ("It is in everyone's best interest to contribute everything to the group account."), heterogeneous endowments and leading-by-example were the most favored treatments according to an anonymous vote.

3.2 Hungary

The workshop took place on 13 October 2020 at Őrség National Park. It involved an expert on agri-environmental policy (male), another expert who is an advisor and a farmer, but not from the local area (male), three farmers: farmer 1, an owner of a local goat milk processor of approximately 3,000 liters/day (male), farmer 2, a manager of an 80 ha farm with 40 fattening cows (male), farmer 3, a self-provisioning owner of a local guesthouse engaged in tourism (one hectare, female), a local restaurant owner who uses local products (male), and an employee of the National Park (female).

At the beginning, the facilitator shortly presented the main aim of the meeting, the plan of the day (to better understand the motivation for cooperation among farmers) and the Dutch case to illustrate the role of collective contracts elsewhere. It was decided to focus on the local park as a common pool resource managed by people in the area. The facilitator also highlighted the ecosystem services provided by the National Park to frame the overall discussion. A simple majority vote was cast on each discussed treatment. Every participant could vote for three alternatives. We discussed the following treatments: group size (five votes), risky provision of the public good (one vote), reward (two votes), sanctions (two votes), heterogeneous endowments (five votes), leading-by-example (one vote), two versions of a threshold public goods game (three and four votes), and emphasising social norms (one vote) (see Appendix B for more details).

3.3 Netherlands

The Dutch workshop took place as a Zoom meeting on 15 December 2020 from 9 am to 11 am. Six participants (5 male, one female) were farmers and members of agri-environmental collectives. One participant (female) represented the umbrella organisation of agri-environmental collectives (BoerenNatuur). To provide an additional perspective on the discussions and outcomes, we conducted an expert interview on 14 January 2021 with an advisor from the national service point on the Common Agricultural Policy. Appendix C presents the outcomes of the voting and discussion notes of the workshop in detail. The final decision was to adopt three treatments in addition to a baseline version: heterogeneous endowments, a threshold public good game, and a higher marginal per capita return.

In the Netherlands, collective contracts are very well known among farmers. Farmers can access agri-environmental measures exclusively through approximately 40 regionally organised collectives. Hence, we will adopt a different approach for the Dutch case. In the experiment, farmers will participate in different treatments of a public goods game (within-subjects design). In a second stage of the experiment, participants will have an opportunity to choose among those treatments and to play the

game again, being paired only with others who have also selected this treatment. This study of preferences for institutions (cf. Rommel, 2015) allows us to draw some conclusions on the value of choosing between different organisations.

3.4 Poland

The workshop took place on 8 April 2021 between 11:00 and 13:00 on the Zoom online meeting platform. For the workshop in Poland, 15 experts were invited, including representatives of ministries, environmental NGOs, scientists specialising in agriculture, administration and rural advice, and people from farmers' associations. Nine people responded positively to our invitations, out of which seven participated in the online workshop. The remaining two indicated an interest in participating in a future meeting.

Participants of the workshop included one representative of the Ministry of Environmental Protection, one representative of the Ministry of Agriculture, and two representatives from NGOs working on biodiversity protection (especially rural bird habitats), one researcher from Warsaw Agricultural University, two agricultural advisors who are heads of the environmental departments in rural advisory centres (from Podlaskie and Pomorskie voivodships). A summary of the results and discussions can be found in Appendix D. At the time of writing this report, the final study design has not been decided.

4. FIRST RESULTS FROM GERMANY

4.1 Farmer Sample, Experimental Design, and Participant Characteristics

Due to the ongoing SARS-CoV-2 pandemic, data were collected in an online survey format. It was not possible to match participants in real life or online in real-time. Hence, we decided for a high-stakes one-shot game with ex-post matching in a survey format. The survey was conducted from December 2020 to February 2021 in collaboration with an experienced German market research company specialized in farming and agriculture (*https://www.agri-experts.de/*). Farmers were recruited from a panel of approximately 1,000 farmer participants and through online ads in specialized farmer magazines. We aimed for a sample size of at least 70 participants per treatment (which was based on expedite power calculations with simplified assumptions). No strict filters were applied to achieve a sufficiently large sample size. Participating farmers had to be involved in the management/decision-making of a farm in Germany. Participants were endowed with 50 Euros. Every tenth farmer received a payment based on the decisions in the public goods game. The initial endowment had an expected value of five Euro, which is a fairly large incentive, given that the survey lasted approximately 10–15 minutes, and comparable to similar studies such as Bouma et al., 2020. The research team calculated the incentives and shared anonymous ID numbers with the survey company, which then matched the data with personal information available only to the company. The survey company administered payments through bank transfers. Informed consent was obtained at the beginning of the survey instrument. No deception was used in the study.

The experimental design is based on five different treatments of a one-shot linear voluntary contributions public goods game with four players (applied between subjects, i.e., participants saw only one of the treatments): BASELINE, HETEROGENEITY, LEADING, NORMS, and OPTIMUM. In BASELINE, participants were endowed with 50 Euros which they could allocate between the group and private account.

In HETEROGENEITY, participants had to decide how much they would contribute if they were endowed with either 25 or 75 Euro. Note that participants were asked to respond to both possibilities. The LEADING treatment requested participants to indicate their contribution (from an initial endowment of 50 Euro) if they were the first to decide in a group of four players (the “leader”). Participants were also asked to indicate their contributions after one person had already decided (as a “follower”).

In the HETEROGENEITY and LEADING treatments, we took inspiration from the so-called strategy method, often applied to experiments based on sequential games (e.g., the ultimatum game). In the strategy method, experimental subjects take on both roles in a sequential game with a finite strategy space. Brandts and Charness (2000) did not find a difference between the “hot” and “cold” versions of an experiment. That is, behaviour did not differ when players assumed only one of two roles compared to when they were deciding in both roles of a sequential game in a finite strategy space. In the LEADING treatment, we asked for conditional contributions in ten intervals with increments of five Euro, i.e., participants had to indicate their contributions as followers for ten intervals of leader contributions, namely 0–5 Euro, 5–10 Euro, …, 45–50 Euro.

The NORMS treatment adds a statement to the BASELINE treatment that points out that participants in similar studies have contributed large amounts to the group account. The OPTIMUM treatment adds

a statement to the BASELINE treatment emphasising that it would be in the best interest of everyone to contribute everything to the group account, i.e., the social optimum.

In total, 358 farmers completed the full questionnaire, and their responses were used for analysis. Participants were between 19 and 77 years old (median = 44; mean = 43.69; SD = 13.26). The median farm size was 87 hectares (mean = 176.97; SD = 13.26). Only 20 participants were female.

4.2 Treatment Effects

Table 2 displays summary statistics for contributions to the group account in Euro. The last row presents the aggregate numbers across all treatments. For the BASELINE, NORMS, and OPTIMUM treatments, contributions were directly taken from participants' responses. For HETEROGENEITY we calculated per person for the high and low endowment conditions. For the LEADING treatment, a weighted average was calculated as follows: one quarter of the average was based on the leader contribution; three quarters were based on the followers' contributions conditioned on the distribution of leaders' contributions.

Table 2: Summary Statistics on Contributions in Euro by Treatments

	Number of observations	Mean	Std. Dev.	Median	Minimum	Maximum
BASELINE	72	34.78	12.78	30	0	50
HETEROGENEITY	71	32.61	11.47	35	0	50
NORMS	72	33	14.93	35	0	50
OPTIMUM	72	40.64	12.46	47.5	1	50
LEADING	71	31.33	10.12	32.71	1.16	50
Pooled data across all treatments	358	34.48	12.82	35	0	50

Source: own calculations

Formal testing revealed that at least one of the treatment distributions is different from at least one other distribution (Kruskal-Wallis test; $\chi = 28.37$ with 4 degrees of freedom; $p < 0.01$). Pairwise tests did not show evidence of differences at $p < 0.05$ for any of the treatment combinations. With the only exception of the OPTIMUM treatment that is statistically significantly different from all other treatments (Two-sample Wilcoxon rank-sum (Mann-Whitney) tests; all $p < 0.01$). To illustrate the distribution of contributions, Fig. 3 displays raincloud plots relative contributions (percentage of initial endowment) by treatment.

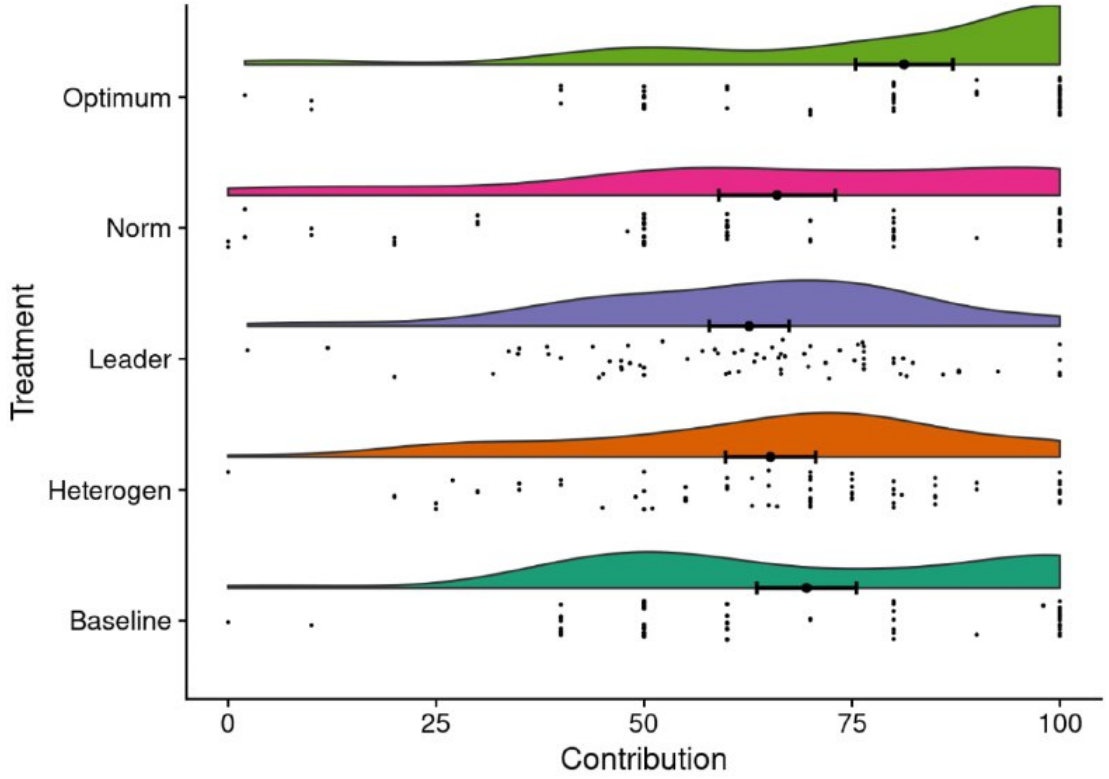

Fig. 3: Raincloud Plots of Relative Contributions by Treatments (Source: own calculations)

Cooperation levels are high. Farmers contributed more than two-thirds of their endowment on average – a number that is typically around 50% in both laboratory (Zelmer, 2003) and field studies (e.g., Rommel and Janssen, 2019). In the HETEROGENEITY treatment, farmers contributed relatively more when they have low endowments (Mean = 17.56 Euro; SD = 20 Euro; equivalent to approximately 70.3% of the total endowment on average), whereas contributions were smaller in relative terms in the high endowment condition (Mean = 47.65; SD = 18.34; equivalent to approximately 63.5% of the total endowment on average). The distributions are statistically significantly different from each other (Wilcoxon signed-rank test, within subjects; $z = -2.145$; $p < 0.05$). This pattern was also found in a laboratory study (Cherry et al., 2005; Martinangeli and Martinsson, 2020).

Fig. 4 displays the average contributions of followers for each of the possible leader contributions in the LEADING treatment, starting from 0–5 Euro up to the highest category of 45–50 Euro. A flat line would indicate that followers ignore what leaders do by not conditioning their contributions on the leader's contributions. An proportionate increase starting at zero would indicate that followers exactly match leaders' contributions to the group account (a slope of one). A steeper increase (a slope greater than one) means that folalowers disproportionally reward high leader contributions (i.e., contributions increase relative to the leader with increasing leader contributions), a slope smaller than one (but greater than zero) indicates that contributions decrease relative to the leader with increasing leader contributions.

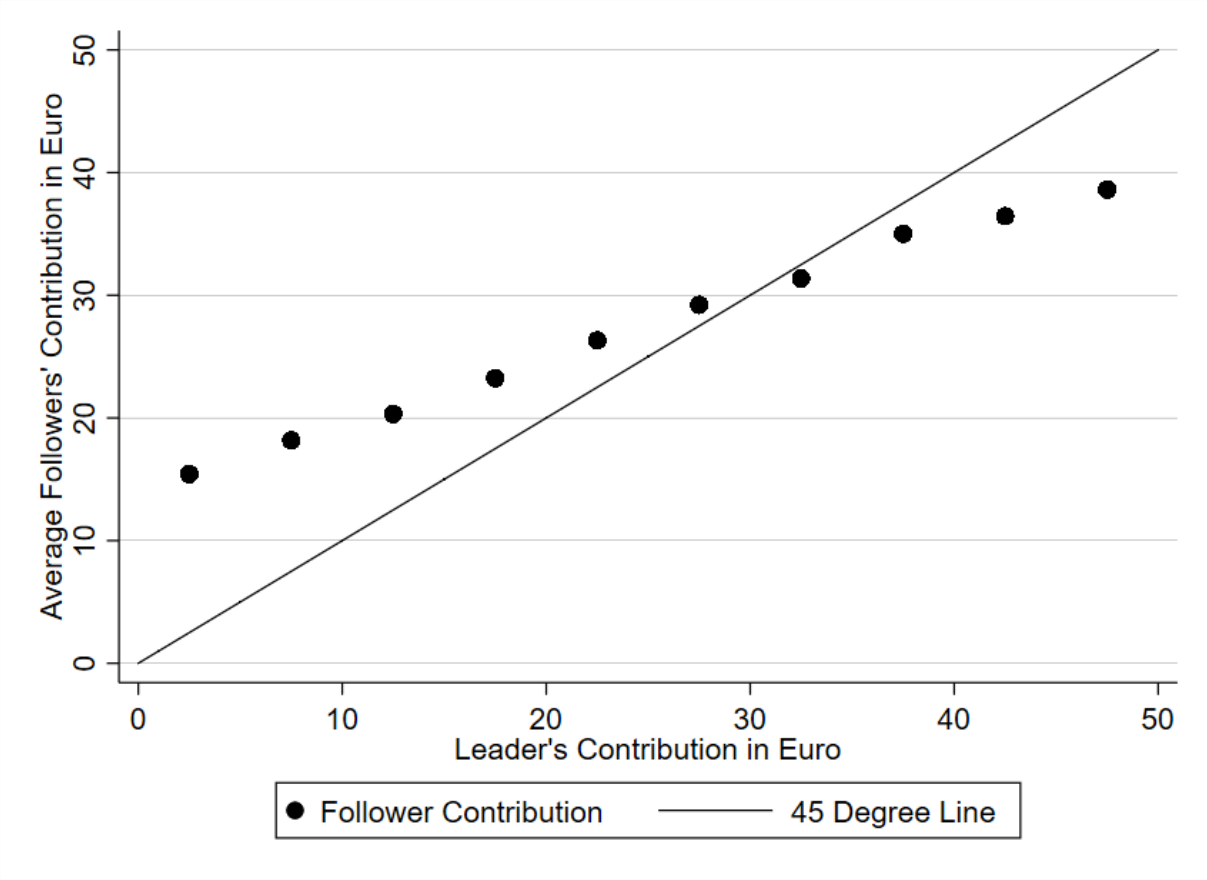

Fig. 4: Average Followers' Contributions by Leaders' Contributions (Source: own calculations)

Leaders contributed 32.32 Euro on average (median = 30; SD = 13.77). The distribution of leader contributions is not statistically different from the BASELINE treatment. Note that the slope in Fig. 4 is below one, but contributions started at a high level. In the lowest category followers match contributions by a factor of approximately six if we compare it to the mid point of the category of 2.50 Euro. As leaders' contributions increase, followers reduce this relative contribution, ending with a factor of approximately 0.8 in the highest category.

4.3 Prediction Survey

4.3.1 Survey Design and Participant Characteristics

We received a total of 212 completed expert predictions from the survey. Some respondents dropped out after the predictions, and there are a few missing variables. A large share of respondents (23%) opened a link distributed through the popular google mailing list of the Economic Science Association, the largest academic network of experimental economists. Twenty-five respondents (= 12%) opened the survey from the link distributed within the Contracts2.0 project.

Table 3 provides an overview of participant characteristics. It displays age in years and the share of female respondents. We also asked respondents to self-assess their knowledge. Respondents were asked about their level of agreeing with statements describing good knowledge of the EU's Common Agricultural Policy (KNOWLEDGE EU CAP), agri-environmental schemes and policy (KNOWLEDGE AES), Economics (KNOWLEDGE ECONOMICS), the public goods game (KNOWLEDGE PGG), and agriculture

(KNOWLEDGE AGRICULTURE) on a scale from strongly disagree (1) to strongly agree (7). The majority of respondents were working in academia (78%). A background in Economics (60%) was most common among respondents. The average age was 36. Given the large share of economists in the sample, good knowledge of the public goods game and Economics should not come as a surprise.

Table 3: Summary Statistics Participants Prediction Survey

Variable Name	Obs.	Mean	Median	Std. Dev.	Min	Max
AGE (years)	201	36.21	34	9.51	21	73
FEMALE (1= female)	209	0.36	0	n/a	0	1
KNOWLEDGE EU CAP (1 = low; 7 = high)	212	3.84	4	1.97	1	7
KNOWLEDGE AES (1 = low; 7 = high)	211	4.46	5	1.78	1	7
KNOWLEDGE ECONOMICS (1 = low; 7 = high)	211	5.49	6	1.37	1	7
KNOWLEDGE PGG (1 = low; 7 = high)	211	4.91	5	1.64	1	7
KNOWLEDGE AGRICULTURE (1 = low; 7 = high)	211	4.18	5	1.79	1	7

Source: own calculations

4.3.2 Comparing Predictions with Farmers' Game Outcomes

Across all treatments, respondents are more pessimistic regarding farmers' contributions on average. They predict farmers to contribute 38% of their initial endowment in the BASELINE condition, which is only approximately half of farmers' actual contribution (cf. Table 2). Experts are most pessimistic regarding the HETEROGENEITY treatment (34%) and more optimistic for the LEADING (50%), NORMS (55%), and OPTIMUM (64%) treatments. Deviations of these predictions are the largest for the BASELINE and HETEROGENEITY treatments (approximately 33 percentage points). For the other three treatments, deviations are lower (see Table 4 for details). Fig. 5 below overlaps the distributions of actual contributions (orange) and expert predictions (green).

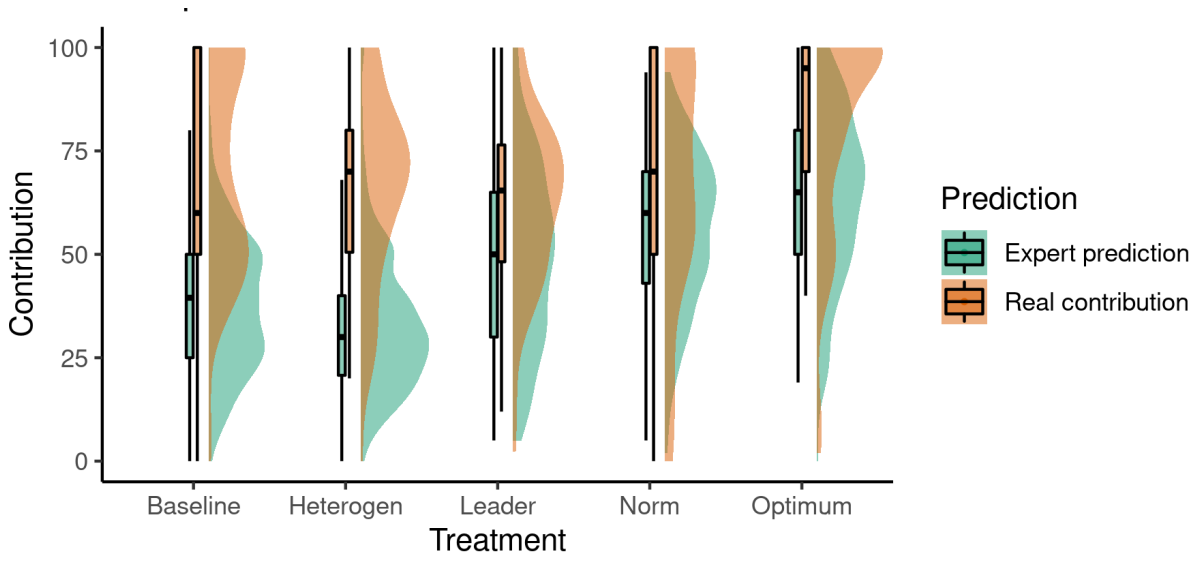

Fig. 5: Distribution of Relative Contributions of Farmers vs. Expert Predictions by Treatments (Source: own calculations)

Fig. 6 displays the accumulated deviations of expert predictions from the survey in percentage points. The graph shows that a few people to the left of the distribution predicted very successfully. The distribution is most dense at around 100, i.e., a deviation of approximately 20 percentage points per treatment on average (100/5) was common. Note that there are also a few outliers on the right part of the distribution, i.e., respondents who predicted very inaccurately and who most likely were too pessimistic.

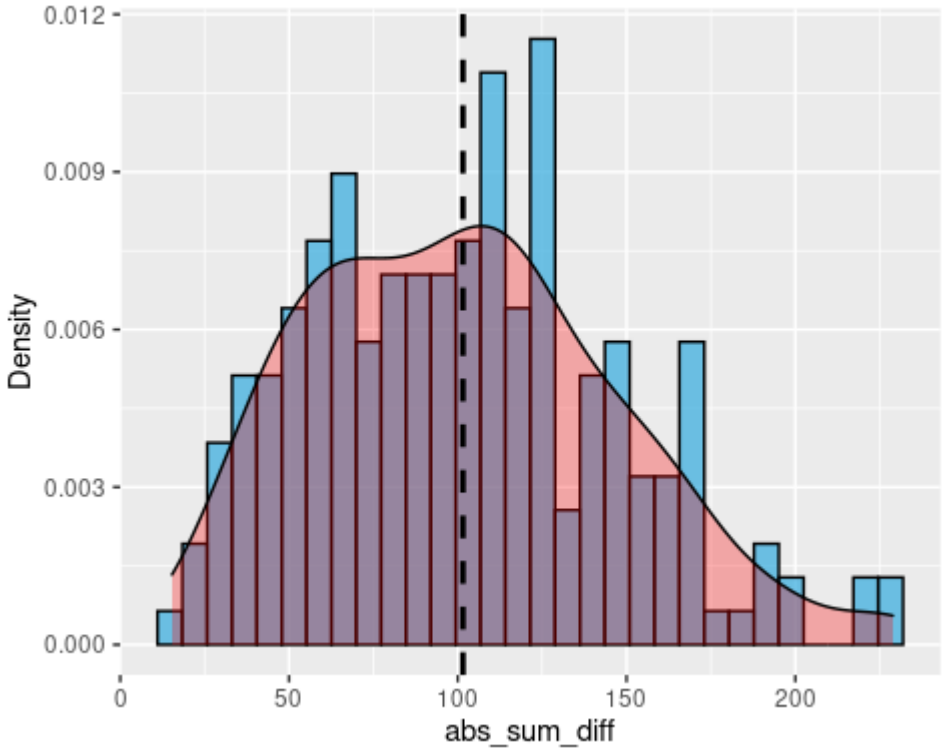

Fig. 6: Summed Deviations of Expert Predictions in Percentage Points of Endowment (Source: own calculations)

4.3.3 Who Predicts Best?

Table 4 presents results from a linear ordinary least regression (OLS) with the difference in percentage points between the actual and predicted average treatment contribution as the dependent variable. Note that these are the pooled data of all five predictions per person. Standard errors are not clustered, and one should repeat the analysis with a panel data model accounting for the nested observations. However, this model can give a good first overview of what more advanced modeling options could confirm. All independent variables are mean-centered, meaning that the treatment effects have a straightforward interpretation for the average respondent. Treatment dummies use BASELINE as the reference category, which under mean-centered covariates shows the deviation of percentage points from the prediction in the constant term of the model. Due to the large number of economists and academics in the sample, we did not include these variables to avoid poor balance. However, we included a combined variable of academic economists that represents approximately half of the sample.

The parameters of the treatment variables other than BASELINE show the effect of the treatment on prediction accuracy. The effects are large and statistically significant for LEADING, NORMS, and OPTIMUM. For these three treatments, respondents predict 11.38, 15.67, and 10.94 percentage points more accurately, respectively. Among the knowledge parameters, *knowledge of Economics* and *knowledge of the public goods game* show larger and statistically significant effects. A one step increase in self-assessed Economics knowledge (recall that this was a seven-point scale) *decreases* accuracy by approximately 1.3 percentage points. In contrast, knowledge of the public goods game *increases* accuracy by approximately the same magnitude. Hence, moving from the lowest to the highest category for this question increases accuracy by approximately eight percentage points (6 x 1.3 = 7.8). There are larger effects for being an academic economist, female, and age, but these variables may also be correlated with each other and some of the other variables. Hence, standard errors and subsequent levels of statistical significance should be interpreted with some care and need further investigation. The model shows high R^2 and adjusted R^2 values, and a large and statistically significant F-statistic, indicating large explanatory power of the independent variables.

Table 4: Regression Results on Predictive Accuracy

Variable	OLS
HETEROGENEITY	0.184
	(1.476)
LEADING	-11.383***
	(1.476)
NORMS	-15.673***
	(1.476)
OPTIMUM	-10.939***
	(1.476)
KNOWLEDGE EU CAP	-0.224
	(0.376)
KNOWLEDGE AES	-0.549
	(0.440)
KNOWLEDGE ECONOMICS	1.306***
	(0.459)
KNOWLEDGE PGG	-1.321***
	(0.361)
KNOWLEDGE AGRICULTURE	0.463
	(0.434)
AGE	-0.264***
	(0.051)
FEMALE	-2.719***
	(0.985)
ECON ACADEMIA	-1.861*
	(1.107)
Constant (BASELINE)	34.507***
	(1.266)
Observations	980
R^2	0.201
Adjusted R^2	0.191
Residual Std. Error	14.616 (df = 967)
F Statistic	20.209*** (df = 12; 967)

Note: Standard errors in parentheses; *p < 0.1, **p< 0.05, ***p<0.01; Source: own calculations

5. SUMMARY AND DISCUSSION

In this milestone report, we have explained how we have developed the public goods game as part of our goal to perform an ex-ante assessment of novel collective contract models in the Contracts2.0 project. Workshops were conducted in four member states, the first data collection was completed in Germany, and an expert prediction survey was run parallel to the public goods game with German farmers.

The overall experiences from the workshops have been positive. The public goods game was met with great interest from stakeholders, albeit there were concerns about the level of abstraction of the game. Fine-tuning the experimental designs and choosing a context that ensures high comprehension and meaningful results in experiments with farmers is not an easy task (Meraner et al., 2018). This holds particularly true in times of the SARS-SoV-2 pandemic when increased involvement and personal interaction that could address some of these challenges are more difficult to achieve.

Another concern frequently raised in the workshops was parallelism, i.e., the link between game results and real-world behaviour. The jury is still out on the best ways to achieve parallelism in experiments with stakeholders. One option is to frame the task in a familiar context (Rommel et al., 2019). Other options include working with rewards linked to the subject under investigation. For instance, Carlsson et al. (2015) funded the construction of a bridge in a Vietnamese village with contributions from a public goods game with participants from the same village. In Müller (2020), contributions to the group account were used for joint investments of machinery circles cooperative organisations in rural Tajikistan. Dessart et al. (2021) used a donation to enhance parallelism in a study on environmentally-friendly farming practices among farmers in three EU member states.

The results of the German public goods game showed that farmers' behaviour differs substantially from subjects in the laboratory. Overall levels of cooperation were substantially larger than expected despite the rather anonymous online format that supposedly could further boost egoistic behaviour. In addition, treatment effects were not in the expected direction. Typically, laboratory studies find negative effects of endowment heterogeneity and positive effects of leading-by-example treatments (e.g., Levati et al., 2007), none of which could be observed in our study. The only treatment that showed substantially larger contributions was to emphasise the social optimum.

Expert predictions were more pessimistic than the actual behaviour of farmers. Experts were also generally more in line with the literature from experimental laboratory studies. Hence, predictions were also not very accurate, albeit there were larger differences between treatments. Among the experts, those who indicated good knowledge about the public goods game performed better, whereas sector-specific knowledge (on agriculture, the common agricultural policy, or agri-environmental schemes) did not substantially improve predictions. Note that we have presented only preliminary results here.

6. REFERENCES

Andersen, E. Kelemen, E., & Megyesi, B. (2020). Catalogue on methods and tools for Policy Innovation Labs. https://www.project-contracts20.eu/wp-content/uploads/2020/05/C20_WP4_D14_D4.1_UCPH.pdf

Banerjee, S., De Vries, F. P., Hanley, N., & van Soest, D. P. (2014). The impact of information provision on agglomeration bonus performance: an experimental study on local networks. American Journal of Agricultural Economics, 96(4), 1009-1029.

Bouma, J. A., Nguyen, T. B., Van Der Heijden, E., & Dijk, J. J. (2020). Analysing group contract design using a threshold public goods experiment. European Review of Agricultural Economics, 47(3), 1250-1275.

Brandts, J., & Charness, G. (2000). Hot vs. cold: Sequential responses and preference stability in experimental games. Experimental Economics, 2(3), 227-238.

Carlsson, F., Johansson-Stenman, O., & Khanh Nam, P. (2015). Funding a new bridge in rural Vietnam: a field experiment on social influence and default contributions. Oxford Economic Papers, 67(4), 987-1014.

Cherry, T. L., Kroll, S., & Shogren, J. F. (2005). The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. Journal of Economic Behavior & Organization, 57(3), 357-365.

Colen, L., Gomez y Paloma, S., Latacz-Lohmann, U., Lefebvre, M., Préget, R., & Thoyer, S. (2016). Economic experiments as a tool for agricultural policy evaluation: insights from the European CAP. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie, 64(4), 667-694.

DellaVigna, S., & Pope, D. (2018a). Predicting experimental results: who knows what? Journal of Political Economy, 126(6), 2410-2456.

DellaVigna, S., & Pope, D. (2018b). What motivates effort? Evidence and expert forecasts. The Review of Economic Studies, 85(2), 1029-1069.

DellaVigna, S., Pope, D., & Vivalt, E. (2019). Predict science to improve science. Science, 366(6464), 428-429.

Dessart, F. J., Barreiro-Hurlé, J., & van Bavel, R. (2019). Behavioural factors affecting the adoption of sustainable farming practices: a policy-oriented review. European Review of Agricultural Economics, 46(3), 417-471.

Dessart, F., Rommel, J., Barreiro-Hurle, J., Thomas, F., Rodríguez-Entrena, M., Espinosa-Goded, M., Zagórska, K., Czajkowski, M., van Bavel, R. (2021). Farmers and the new green architecture of the Common Agricultural Policy: a behavioural experiment. JRC Science for Policy Report, forthcoming.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. Economics Letters, 71(3), 397-404.

Isaac, R. M., Walker, J. M., & Thomas, S. H. (1984). Divergent evidence on free riding: An experimental examination of possible explanations. Public Choice, 43(2), 113-149.

Kuhfuss, L., Begg, G., Flanigan, S., Hawes, C., & Piras, S. (2019). Should agri-environmental schemes aim at coordinating farmers' pro-environmental practices? A review of the literature. https://ageconsearch.umn.edu/record/289774/files/Kuhfuss-Should%20agri-environmental%20schemes%20aim%20at%20coordinating%20farmers%27%20pro-environmental%20practices_a.pdf

Kuhfuss, L., Préget, R., Thoyer, S., & Hanley, N. (2016). Nudging farmers to enrol land into agri-environmental schemes: the role of a collective bonus. European Review of Agricultural Economics, 43(4), 609-636.

Latacz-Lohmann, U., Balmann, A., Birner, R., Christen, O., Gauly, M., Grajewski, R., ... & Weingarten, P. (2019). Designing an effective agri-environment-climate policy as part of the post-2020 EU Common Agricultural Policy. Berichte über Landwirtschaft-Zeitschrift für Agrarpolitik und Landwirtschaft. https://buel.bmel.de/index.php/buel/article/view/273

Ledyard, J. O. (1994). Public goods: A survey of experimental research. https://authors.library.caltech.edu/80749/

Levati, M. V., Sutter, M., & Van der Heijden, E. (2007). Leading by example in a public goods experiment with heterogeneity and incomplete information. Journal of Conflict Resolution, 51(5), 793-818.

Martinangeli, A. F., & Martinsson, P. (2020). We, the rich: Inequality, identity and cooperation. Journal of Economic Behavior & Organization, 178, 249-266.

Meraner, M., Musshoff, O., & Finger, R. (2018). Using involvement to reduce inconsistencies in risk preference elicitation. Journal of Behavioral and Experimental Economics, 73, 22-33.

Müller, M. (2020). Leadership in agricultural machinery circles: experimental evidence from Tajikistan. Australian Journal of Agricultural and Resource Economics, 64(2), 533-554.

Parkhurst, G. M., Shogren, J. F., Bastian, C., Kivi, P., Donner, J., & Smith, R. B. (2002). Agglomeration bonus: an incentive mechanism to reunite fragmented habitat for biodiversity conservation. Ecological Economics, 41(2), 305-328.

Prager, K., Matzdorf, B., Dutilly, C., Andersen, E., Barghusen, R., Bredemeier, B., ... & Turkelboom, F. (2020). Key concepts to investigate agri-environmental contracts–shared conceptual framework. https://pub.epsilon.slu.se/17280/

Rommel, J. (2015). What can economic experiments tell us about institutional change in social–ecological systems? Environmental Science & Policy, 53, 96-104.

Rommel, J., & Anggraini, E. (2018). Spatially explicit framed field experiments on ecosystem services governance. Ecosystem Services, 34, 201-205.

Rommel, J., Hermann, D., Müller, M., & Mußhoff, O. (2019). Contextual framing and monetary incentives in field experiments on risk preferences: Evidence from German farmers. Journal of Agricultural Economics, 70(2), 408-425.

Rommel, J., & Janssen, M. (2019). Leading by example and endowment heterogeneity in local public good provision: A framed field experiment in Hyderabad, India. https://complexity.asu.edu/sites/default/files/papers/cbie_wp_2019_002.pdf

Thomas, F., Midler, E., Lefebvre, M., & Engel, S. (2019). Greening the common agricultural policy: a behavioural perspective and lab-in-the-field experiment in Germany. European Review of Agricultural Economics, 46(3), 367-392.

Villamayor-Tomas, S., Sagebiel, J., & Olschewski, R. (2019). Bringing the neighbors in: A choice experiment on the influence of coordination and social norms on farmers' willingness to accept agro-environmental schemes across Europe. Land Use Policy, 84, 200-215.

Wanner, S., Mortelmans, D., Turkelboom, F., Defrijn, S., García, M., DUH, S. S., ... & WP, B. M. (2020). Integrated evaluation report of innovative/interesting contracts. https://www.project-contracts20.eu/wp-content/uploads/2020/08/C20_WP3_D11_D3.2_INBO.pdf

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. Experimental Economics, 6(3), 299-310.

APPENDIX A: ENGLISH TRANSLATION OF LIST USED TO DISCUSS TREATMENTS IN THE GERMAN WORKSHOP

Code:

Suggested treatments

1) Group size is doubled from four to eight: “In this exercise you’ll be part of a group of eight farmers…”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

2) Public good has larger return: “Money on the group account will be tripled….”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

3) Public good provision is risky: “Money on the group account will be multiplied with three with 50% chance and multiplied with one with 50% chance.”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

4) Public good provision is very risky: “Money on the group account will be multiplied with four with 50% chance and multiplied with zero with 50% chance.”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

5) Reward: “You can send 100 Euro to another player, depending on the other player’s behaviour.”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

6) Costly reward: “You can send 100 Euro to another player, depending on the other player’s behaviour. You have to pay 50 Euro if you want to do so.”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

7) Sanction: “You can take 100 Euro from another player, depending on the other player’s behaviour.”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

8) Costly sanction: “You can take 100 Euro from another player, depending on the other player’s behaviour. You have to pay 50 Euro if you want to do so.”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

9) Heterogeneous endowments: „Two players have 300 Euro, two players have 700 Euro…"

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

10) Leading by example: "You decide first. The other participants can choose depending on what you did. / Another player decides first. You and the other two participants decide depending on the first participant."

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

11) Social norm: “Other players have often placed large amounts on the group account.”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

12) Social optimum: “If everyone contributes their full endowment of 500 Euro, then everyone receives a maximum of 1,000 Euro.”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

13) Individual optimum/Nash equilibrium: “In this exercise, it is the best strategy to keep everything for yourself.”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

14) Critical mass I: “The money on the group account will only be doubled if there are at least 1,400 Euro. If this threshold is not reached, the money on the group account will be lost.”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

15) Critical mass II: “The money on the group account will only be doubled if there are at least 1,400 Euro. If this threshold is not reached, the money will be distributed in equal parts to all participants (without multiplication).”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

16) Critical mass III: “The money on the group account will only be doubled if there are at least 1,400 Euro. If this threshold is not reached, everyone receives back their own amount.”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

17) Emphasising losses: “By contributing to the group account, you may lose money.”

-5	-4	-3	-2	-1	0	1	2	3	4	5
Not interesting at all					neutral					Very interesting
o	o	o	o	o	o	o	o	o	o	o

Justification/comments:

_____

_____

_____

_____

_____

Other suggestions?

_____

_____

_____

_____

_____

APPENDIX B: VOTING RESULTS AND DISCUSSION NOTES OF HUNGARIAN WORKSHOP

Discussions during the workshop

After a short self presentation we discussed about the factors influencing cooperation; we also had a short game (how to distribute 2,000 Hungarian Forint (approx. 6 euros) between a private and a public account).

Then we discussed about the different treatments.

Discussion about the variants

We formed pairs, who discussed the treatments among each other and then we had plenary discussions.

1. Group size is doubled from four to eight: "In this exercise you'll be part of a group of eight farmers…"

Farmer 2: causes real changes, increased risks.

Farmer 1: the increase in the number of participating farmers is good; 8 is still not a too huge number (200 would be too high).

Expert 1: it can be difficult to coordinate if there are more participants; how individual and community interests can be mediated?

Farmer 2: it is up to the goal: if it is about increasing product quantity then it is beneficial if it is about sharing the machinery, then it is making cooperation more difficult.

ecologist 1: the jump form 4 to 8 is okay; over 50 can be problemtic.

Expert 2: can be interesting; it is more difficult to coordinate if the number of participants increases.

2. Public good provision is risky: "Money on the group account will be multiplied with three with 50% chance and multiplied with one with 50% chance.

Expert 1: It can be interesting for those who can stay risks.

Farmer 2: Loosing your investment raises risk.

The participants did not like this treatment, there was a consensus that it is too risky.

3. Reward: "You can send 100 Euro to another player, depending on the other player's behavior."
4. Sanction: "You can take 100 Euro from another player, depending on the other player's behavior."

The participants preferred to discuss the two treatments together, but found sanction too offensive.

Expert 1: The amount is very important in this case; and also the share of the possible gains which can be given or taken away.

Farmer 2: It can be very harmful for the community to know that the others can reward or sanction each other.

Farmer 1: Yes it can result in tensions, conflicts, also maybe corruption or playing out the members against each other.

Farmer 3: I also do not think that it would be good, neither rewarding nor sanctioning within the community.

Expert 1: sanctioning is a serious decision, one should avoid it, as the overall goal is to make better the situation instead of making it more difficult. Rewarding can be useful though.

Expert 1: It is better to avoid such situations, but if they decided together then mabe it works.

5. Heterogeneous endowments: „Two players have 300 Euro, two players have 700 Euro…" (5 votes)

Expert 2: This is the most interesting, and it is the closest to the reality.

The participants agree that is the most typical situation, and most of them thinks that exactly for this reason it should be used as a treatment.

6. Leading by example: "You decide first. The other participants can choose depending on what you did. / Another player decides first. You and the other two participants decide depending on the first participant."

They agreed that it can be interesting as the example has a real influence, although an expert mentioned that it is more probable that they discuss their decisions together and then they decide about the sums.

7. Critical mass I: "The money on the group account will only be doubled if there are at least 1,400 Euro. If this threshold is not reached, the money on the group account will be lost." (3 votes)
8. Critical mass II: "The money on the group account will only be doubled if there are at least 1,400 Euro. If this threshold is not reached, the money will be distributed in equal parts to all participants (without multiplication)." (4 votes)

We discussed the two treatments together.

Ecologist 1: being a person who does not like risk, I cannot support critical mass 1, and risk the loss of collected money, but critical mass 2 can be okay.

Farmer 1: it is a good opportunity to count and decide; the second option is better.

Expert 2: it is an interesting suggestion, both of them.

Farmer 2: it would be important not to loose the money from the public account.

Expert 1: the second one is less risky so it is less interesting, and also less possible, but exactly for this reason the first one is interesting because if the money is gone the community cannot realize its goals. If a community initiate something but unable to realize and then the money disappears; it is a real incentive.

9. Social norm: "Other players have often placed large amounts on the group account." (1 vote)

There is a misunderstanding around the treatment (the participants assume that someone not belonging to the community places money). After clarifying it they realise that it is interesting, but suggest to have new treatment (10).

10. : Other players donate higher amounts at the group account" (1 vote)

Results

Finally we selected the treatments by a simple voting; the results can be seen in Fig. 7.

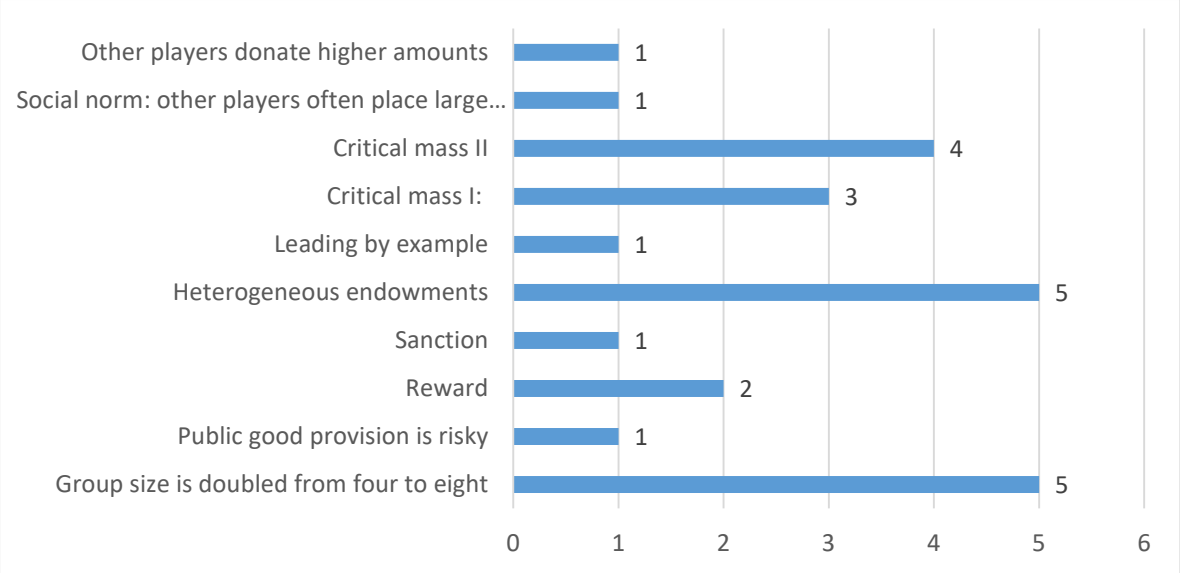

Fig. 7: Results of Voting for Variants Hungarian Workshop (Source: own calculations)

APPENDIX C: VOTING RESULTS AND DISCUSSION NOTES OF DUTCH WORKSHOP

Results of the voting

After the explanation of all the variants the participants were given the opportunity to indicate per variant individually how interesting they found the variants (ranging from -3 not interesting at all to 3 very interesting). Fig. 8 shows the total sum of points received per variant. The participants were also given the opportunity to provide an explanation for their voting (see Annex 2).

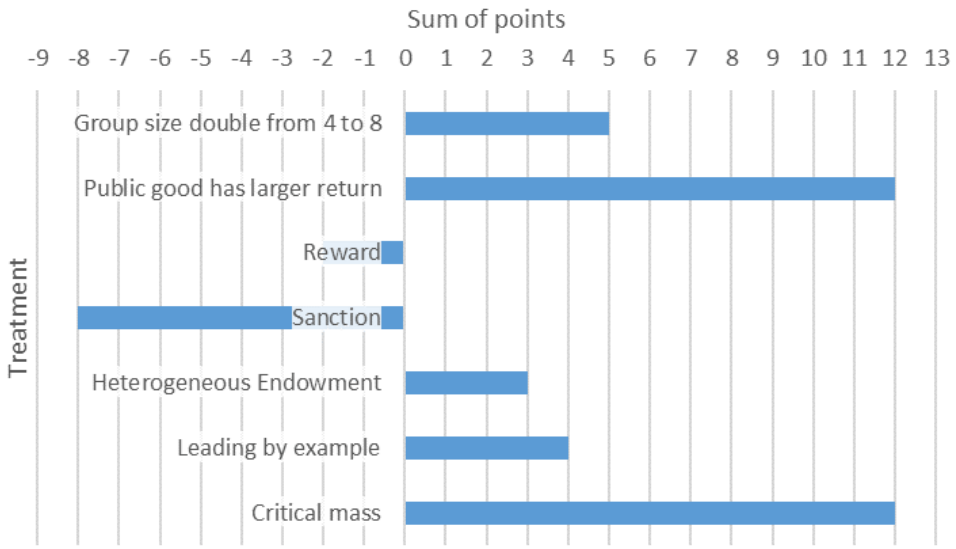

Fig. 8: Results of Voting for Variants Dutch Workshop (Source: own calculations)

Plenary reflection on filling in the form with the different variants:

Difficulty level differed greatly per variants. For example, group size: you would expect a larger group to yield more, but the larger the group, the more difficult it will be to continue to serve the group (trade-off). Not all variants fit the current working method of the collectives. Another consideration: a collective has an optimum, when does a farmer still feel like part of a group?

Question: Don't people react differently when it comes to their own money? Answer: Yes, that's right, but it would be unethical to let people put their own money in. It is possible to reflect on this by asking participating farmers after the Public Goods Games how they would use their own money.

Next steps and reflection on workshop:

Participants needed some time to get into the workshop. More context in preparation for the workshop would have been better. A clearer translation to the Dutch collective system would also have been welcome. The interactive form of the workshop was appreciated, participants felt they had really provided input.

The participants would like to receive a report of this workshop. This report and the results of the workshop will also be discussed with RVO.

Overview of the variants presented

1. Public good has larger return: "Money in the group account is tripled ...."

2. The group size is doubled from four to eight: "In this exercise you will be part of a group of eight farmers ..."

3. Reward: "You can give € 100 to another player, depending on the other player's behaviour."

4. Sanction: "You can take € 100 from another player, depending on the behaviour of the other player."

5. Differences in endowment: "Two players have € 300, two players have € 700 ..."

6. Leading the example: “You decide first. The other participants can choose depending on what you have decided. / Another player decides first. You and the other two participants decide depending on the decision of the first participant. "

7. Critical amount: “The money in the group account is only doubled if there is at least € 1,400 in it. If this amount is not reached, the money will be divided equally among all participants (without multiplication). "

Individual explanation on the presented variants

1. Public good has larger return: "Money in the group account is tripled ...."

- Collaboration pays off (earning model); this strengthens social capital

- I think it provides an incentive for both considerations, if you bet more you get more back, but also if you bet less

- Together you are stronger, does that also work out in practice?

- Rewarding is the tool to motivate farmers, at our dairy factory this has also been tested in the Caring Dairy program

- This is an extra trigger to commit to the collective and will certainly be stimulating. But in the end (almost) everyone is calculating and will wonder what the personal benefits of 'me'.

2. The group size is doubled from four to eight: "In this exercise you will be part of a group of eight farmers ..."

- Interesting, but be vigilant about the total group size in relation to the target. A sense of togetherness must remain intact

- I think it does affect people's behaviour but I wonder if the group size is big enough, maybe it is more interesting to make it much bigger, for example 20

- This in itself is interesting, you would like to know what the effect of increasing the group of farmers. However, a larger group is not necessarily ideal because as a collective you want to be able to continue to serve everyone and keep the work area clear.

- The farmer himself will not notice this much and will not be motivated.

- It is true that with larger group sizes, choices are made more for individual interest. But to have a greater impact and togetherness, a larger group is better. The amount to be distributed must increase proportionally, otherwise it is not attractive.

3. Reward: "You can give € 100 to another player, depending on the other player's behaviour."

- Don't, you can damage trust and emphasize differences (coercion and abuse of power)

- Provided it is not subtracted from your own pot of money, this would provide a lot of incentives for collaboration (kind of like the collaboration bonus that Bert talked about)

- I would like to encourage extra effort, the more you do, the higher the average compensation. Effort must be rewardable

- Farmers should not be able to give each other 'money', that creates friction. You allow one more than the other. This could only be done from a collective, then it is an option.

- I am not in favour of rewarding others based on his / her behaviour. Because that also automatically means 'penalty discount' and you don't want that in a collective in which good relationships are crucial.

4. Sanction: "You can take € 100 from another player, depending on the behaviour of the other player."

- Don't, you can harm trust and emphasize differences (coercion and abuse of power)

- I think it is more interesting to investigate rewards

- It is not nice to contribute fully to the result while someone else is cutting the corners. If a participant commits a violation, one must be able to sanction. Just like you should be able to be rewarded if you do very well

- Just like the previous question, farmers should not be able to determine each other's 'income' and certainly not in a negative form.

- Same answer as previous. You should avoid a reward or penalty discount between equal participants.

5. Differences in endowment: "Two players have € 300, two players have € 700 ..."

- Don't, you can damage trust and emphasize differences (coercion and abuse of power)

- Interesting because it is closer to reality. I wonder if those with larger capital automatically take more responsibility by investing more. I would also find it interesting to combine this variant with a threshold value.

- The "small" farmers can often simply make a greater contribution to biodiversity than the large farmers because this can be better integrated into their business operations. But the question is whether they can contribute as much financially as a big farmer. Investments should not have to be restrictive on results

- This is a realistic starting point. Because one has more to spend or wants to spend more than the other. This also means that the consequence is that the person who contributes more also gets more in return. But it must be proportionally correct. It cannot be that the most wealthy go on with the main prize. As with the nitrogen space that is open to agriculture and industry, with which the wealthy industrial sector is running off with space and agriculture is, as it were, being bought empty. Maybe there should be some limitation.

6. Leading-by-example: “You decide first. The other participants can choose depending on what you have decided. / Another player decides first. You and the other two participants decide depending on the decision of the first participant. "

- Be careful who goes first. Does everyone trust them?

- This would not be my preference at all, then everyone will wait and see, however, I am curious what effect this method has. So hence the zero

- An example could have a positive effect, but the downside is that this can also go the other way, the neighbour doesn't do it so I don't do it either.

- I think realistic. Your behaviour is nevertheless based on how someone else reacts and you will also look at each other in practice.

7. Critical amount: “The money in the group account is only doubled if there is at least € 1,400 in it. If this amount is not reached, the money will be divided equally among all participants (without multiplication). "

- Collaboration pays off (earnings model); this strengthens social capital (working on togetherness)

- Equal distribution is difficult because every participant is different just like his company. Again, one person can contribute to biodiversity more easily than another, while the financial resources may be smaller

- This will motivate farmers to do / deploy a little more than they actually would.

- I think it's a good one too. It is an incentive to reach a certain critical mass.

General remarks

- When making a choice to choose a variant, think about doing it. A certain variant may be preferred per goal (e.g., duration of the goal / reward).

- Add a variant that combines variant "differences in capital" and "threshold"

APPENDIX D: VOTING RESULTS AND DISCUSSION NOTES OF POLISH WORKSHOP

Comments and suggestions regarding PGG:

We collected comments in the discussion and in a survey on Google forms. Here we present a summary of most important suggestions:

– Environmental context shall be separated from the game instructions. The introduction should present the goal of the project, then it should be said that we prepared the game to measure cooperation/simulate cooperative decision making for common good, and at the end the game should be introduced (separate screen from the introduction). In particular, the CAP should not be too closely related to the game instructions. The game instructions are more important than the context in this case.

– The sentence of the introduction "We developed a small exercise to investigate this." could be used as a link between introduction (about the study goal, agri-environmental context etc.) and game instructions.

– The game shall be explained carefully to ensure proper understanding amongst farmers: bold the most important parts, prepare video with explanation of consequences of various decisions, introduce a trial round.

– Relationship between payments being made every 10th farmer and a group size of 4 is not presented clearly enough. Explain in more detail why and how 1 in 10 participants will be randomly selected, if there are only 4 farmers in every group. For example: “Among the entire pool of people who take part in the study we will draw 1 in 10. It may happen that no one in your group gets paid.”

– It may be interesting for all participants to learn about others behaviour of others in their group, not only to the ones who are selected to get the reward. Everyone should be informed about the result of the game, independent of whether he is selected or not.

– Please explain how the cooperation benefits everyone in the group. What is the goal of cooperation? Why would one cooperate?

– The game seemed very abstract and far from representing real-world farming decisions to most of the participants. They suggested we could develop scenarios (water scarcity, pastures) to closer simulate the decision process. In particular concerns were expressed with regard to policy conclusions resulting from the study.

– To others lack of context seemed to be more generalizable. They only suggested that the introduction and explanations must be more engaging to respondents.

– Lack of other than monetary goals was suggested as the weakness of the proposed study. It was suggested that farmers will only cooperate when they see a clear benefit to the local community.

– This method of research will not translate into the knowledge of what the farmer is inclined to do in the real situation of the implementation of agri-environmental policy. The research questions are too abstract.

Predictors of cooperation

We discussed the conditions which could enhance cooperation amongst Polish farmers, as well as farm and farmers' characteristics that could be predictors of one's willingness to enter into group contracts.

The following conditions could enhance cooperation:

– Clear and substantial benefit (profit, premium) of cooperation: a farmer must clearly see the profit from the collective implementation of certain practices.

– Leader arranging formalities and paperwork for the whole group (administrative work performed by someone else).

– Low costs of entry into cooperation.

– Acquaintances and close relationships – friends, family, and everyday met neighbors being part of the group (not strangers)

– Context – as examples water management during drought, pastures.

The following characteristics could predict farmers' susceptibility to enter into contracts:

– Current engagement in similar contracts (for example Leader programme, or cooperation with local NGOs).

Other comments

– Farmers in Poland tend to be very risk averse. Risk neutral farmers are the riskiest (real risk lovers can't be found in the farmer community).

– The game resembles casino playing, it is not a good tool to advise policy.

– Clerical approach: how to get farmers to participate? Is a more important question then the potential of novel contract design features.

– The Ministry is not planning to use group contracts under the new CAP. Mostly due to inability of control over responsibility attribution in relation to payments (for example 5 farmers enter into the contract, one of them stops fulfilling the obligations – what to do then?)

– The LEADER programme works well, but responsibility is ascribed to the local centers. It requires administration that takes over the managerial role over the group.

– In terms of environmental challenges Poland is distinct from EU countries and some of the solutions may not apply to the region. For example, it does not need restoration of environmentally valuable areas, it just requires stopping of the intensification processes.

Comments collected after the workshop:

– I think it would be very interesting to play a game of public goods "without context", i.e. without specifying a group goal, compared to a game with a specific group goal or different goals (e.g. road construction, orphanage, environmental protection).

– From my several years of experience: agri-environmental measures were first undertaken by farmers, who are generally more willing to take risks and trust other people (and the system) more than others.

– I wonder how you will invite participants and how you will interpret the results in this context. Most people are more cautious of offers that appear on the web and received by email, and might be reluctant to take part if they don’t hear about the study "in person".

– Together with several other representatives of NGOs, we have prepared a document, in which we introduced the possibility of collective involvement of farmers in the context of the newly designed CAP. This document was submitted to the Ministry of Agriculture and Rural Development.

– Concerns over big business, such as the chemical sector, being too involved in European Green Deal strategy, particularly climate neutral agriculture.