*Q Open*, 2022, **3**, 1–18
https://doi.org/10.1093/qopen/qoac023
Advance access publication date: 26 August 2022
**Article**

OXFORD



# Learning about German farmers' willingness to cooperate from public goods games and expert predictions

**Jens Rommel [1,*], Christoph Schulze [2], Bettina Matzdorf [2,3], Julian Sagebiel [4] and Vera Wechner [5]**

[1]Swedish University of Agricultural Sciences, Department of Economics, 750 07 Uppsala, Sweden
[2]Leibniz Centre for Agricultural Landscape Research, Working Group: Governance of Ecosystem Services, 15374 Müncheberg, Germany
[3]Leibniz Universität Hannover, Institut für Umweltplanung, 30419 Hannover, Germany
[4]German Centre for Integrative Biodiversity Research (iDiv), 04103 Leipzig, Germany
[5]Humboldt-Universität zu Berlin, Department of Agricultural Economics, 10117 Berlin, Germany

*Corresponding author: Swedish University of Agricultural Sciences, Department of Economics, 750 07 Uppsala, Sweden. E-mail: jens.rommel@slu.se

**Received:** June 1, 2022. **Accepted:** August 20, 2022

## Abstract

There is a growing interest in collective contracts to address agri-environmental policy goals at landscape scales. Yet, little is known about farmers' general willingness to cooperate. We developed four treatments of a linear public goods game to investigate farmers' willingness to cooperate: (1) heterogeneous endowments, (2) leading-by-example, (3) social norms, and (4) pinpointing the socially optimal solution. Based on a sample of 358 German farmers, we find that contributions reach more than two-thirds of the initial endowment across different treatments on average. Nudging the socially optimal solution is the most effective treatment. In addition to the experiment, we elicited incentivized predictions on experimental outcomes from 212 experts. Expert beliefs on treatment effects appear to be calibrated on laboratory studies, highlighting the need to conduct, communicate, and discuss experimental studies outside the laboratory. Young female academics with an Economics background most accurately predict farmers' behaviour in the experiment.

**Keywords:** Artefactual field experiment, Common Agricultural Policy, Methodology, Experimental Economics

**JEL codes:** C93, D91, Q18

## 1. Introduction

The European Union's Common Agricultural Policy is frequently criticized for its low environmental effectiveness. In spite of large tax-funded payments to farmers, environmental public goods are underprovided in agricultural landscapes (Pe'er *et al.* 2017; Hasler *et al.* 2022). Targeting and coordinating policy measures remain a major policy challenge. Under the so-called Dutch model, farmers exclusively receive second pillar support from the Common Agricultural Policy's agri-environmental schemes under collective contracts (cf. Bouma *et al.* 2020). Approximately fourty regionally organized environmental farmer collectives are currently active in the Netherlands. Advantages of organizing environmental management in collectives may include improved coordination of land management at larger scales for habitat protection and related biodiversity conservation or the rewetting of peatlands. Lower administrative costs (as only one formal contract is needed) and greater flexibility (as farmers may set aside no land, whereas others may set aside a lot of land to reach a certain target at larger scales) are additional benefits of collective contracts.

In spite of international interest in the Dutch model or experiences with collective elements in farmland nature conservation in Switzerland (Huber *et al.* 2021; Krämer and Wätzold 2018), little is known about farmers' basic propensity to cooperate. This is important, because—unlike in the Netherlands—collective contracts must co-exist and compete with individual contracts in other EU member states. The German Scientific Advisory Board on Agricultural Policy, Food, and Consumer Health Protection recommended that Germany should 'improve 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). In the current debate on Common Agricultural Policy reforms, Germany was reminded by the commission to offer greater support for collective contract models.

The scientific literature has dealt with collective and agri-environmental contracts extensively (see Kuhfuss *et al.* 2019 for an overview of the literature and Prager 2022 for an overview of diverse cases). Research, based on social-ecological models has emphasized the importance of spatial coordination among farms to address environmental goals more effectively (Nguyen *et al.* 2022). For example, it could be important to reduce habitat fragmentation for achieving biodiversity targets (Parkhurst *et al.* 2002). The shape of the landscape also impacts conservation costs and hence farmers' decisions to enrol their land in nature conservation schemes (Huber *et al.* 2021). There is a growing awareness of these complex drivers of farmers' behaviour and a call for more policy-oriented experimental research on farming practices (Colen *et al.* 2016; Dessart *et al.* 2019; Lefebvre *et al.* 2021). 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, 2021; Dessart *et al.* 2021). While these studies focus on individual behaviour or preferences, studies investigating coordination and strategic interaction with a landscape connectivity perspective mostly rely on students (e.g. Banerjee *et al.* 2014; Bouma *et al.* 2020; Kuhfuss *et al.* 2022). Although agricultural students are frequently used as subjects in behavioural experiments to proxy for farmers' behaviour, comparisons of farmers and agricultural students show that these subject pools can differ in important behavioural outcomes (Grüner *et al.* 2022; Maart-Noelck and Mußhoff 2014). We augment this research by conducting an artefactual field experiment (Harrison & List 2004) with farmers.

For an effective transfer of research results into the policy process, eliciting expert views and predictions is useful for several reasons (DellaVigna *et al.* 2019). Benchmarking beliefs can increase the acceptance of null results. Regular and accurate predictions may also fill-in for in data-scarce settings. Predictions can inform study designs (DellaVigna and Pope 2018b; Milkman *et al.* 2022) and be effective in eliciting ex-ante beliefs on a research topic, ideally leading to an update in beliefs of policy-makers or other experts (Vivalt and Coville, 2020). Whereas predictions have so far focused on abstract laboratory experiments or large-scale randomized controlled trials (DellaVigna and Pope 2018a,b; Milkman *et al.* 2022), there is only one other study predicting the behavior of farmers in an experiment (Schaak *et al.* 2022).

It is the first objective of this paper to investigate German farmers' willingness to cooperate, using public goods games. It is the second objective of this paper to study deviations from farmers' behaviour and experts' predictions of this behaviour. To address these objectives, in parallel to the experiment, we ran a survey asking researchers and practitioners to predict the outcomes of the public goods game in Germany. By addressing the first objective, we provide an ex-ante policy evaluation of farmers' general propensity to cooperate, an important pre-requisite for the successful implementation of collective agri-environmental contract models (Latacz-Lohmann *et al.* 2019). Unlike, discrete choice experiments, our approach takes into account the strategic interdependence among farmers and provides monetary incentives for decisions. On the downside, the public goods games involve abstract decision-making, and participating farmers do not directly interact with each other in real life, limiting the realism of the task. With the second objective, we investigate how well policy and economic experts' beliefs are aligned with actual behaviour of farmer stakeholders. This allows us to assess possible biases which may directly shape the debate and policy outcomes via the science-policy interface.

The paper is organised as follows. We start by providing a short overview of the public goods game and the treatments we developed in the workshop. We also outline the expert prediction survey. We then present and discuss the results.

## 2. Methodology and experimental design

### 2.1 The public goods game

The public goods game was developed to study free-riding behaviour and collective action problems (Isaac *et al.* 1984). The game has quickly developed into a workhorse of Experimental Economics to study human 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 $\pi_i$ in a simple one-shot linear voluntary contribution mechanism public goods game is then defined as follows:

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

where $x_j$ 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 $i$. In this simple setting, with full information, profit maximization, and no preferences over the outcomes of others or the overall distribution of outcomes (i.e. so-called standard preferences), the Nash equilibrium is to contribute zero for all players, because players internalise 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 one-shot laboratory experiments match four or five players and use a constant of $a = 2$, that is, contributions to the group account are doubled. Fig. 1 displays the steps of the BASELINE version of the public goods game used in our experiment. Each participant is endowed with 50 Euro and, in this example, contributes 25 Euro to the group account. After the contributions are doubled, each player receives an equal share, leaving everyone with 75 Euro in the end.

### 2.2 Considered treatments and workshops

From a vast body of literature—encompassing hundreds of laboratory studies (see Ledyard 1995 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, that is, they do not play the Nash equilibrium. In repeated games,

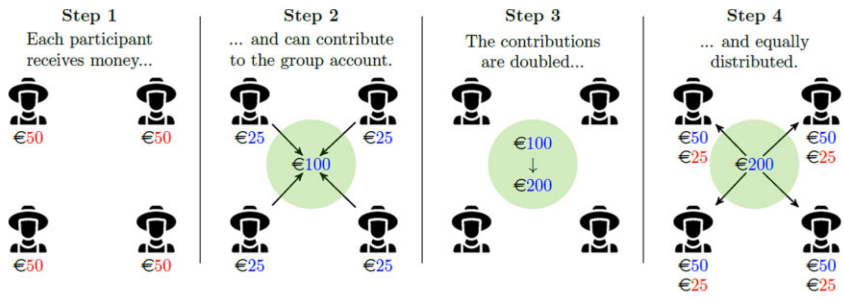


**Figure 1.** Example of steps and outcomes in the public goods game (Source: own design; icon from flaticon.com by freepik).

cooperation (defined as contributions to the group account) decreases over time. Notably, if participants are re-matched with different participants, even after experiencing such decay, they proceed 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).

We developed the experiments jointly with stakeholders from a long list of treatments commonly applied in laboratory experiments. This list was inspired by the laboratory literature, particularly a review article and a meta-analysis (Ledyard 1995; Zelmer 2003). The final treatments were the outcomes of a workshop held on 23 January 2020 in Berlin, which 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. For details on the workshop organization and related experiments in other EU member states, see Rommel *et al.* (2021).

With the help of a facilitator, workshop participants started to discuss 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 (which as a group exercise was also played with all participants).[1] Participants went through a list of treatments, voiced their opinions, and suggested additional treatments. Finally, an anonymous vote was casted after the discussion. In total, seventeen treatments were discussed (see supplementary material for an overview of the workshop material).

From the discussion and voting, it became clear that understanding norms (highlighting that others have contributed substantial amounts in similar studies), nudging 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 favoured treatments. For the exact wording of the treatments, see the English and German versions of the final survey instrument.

## 2.3 Experimental design, farmer sample, and participant characteristics

The BASELINE version of the experiment consisted of a four player one shot linear voluntary contribution mechanism public goods game (with $n = 4, a = 2, e = 50$ Euro). Four treatments were applied between subjects, namely HETEROGENEITY, LEADING, NORMS, and OPTIMUM. In BASELINE, participants were endowed with fifty, 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 twenty-five or seventy-five Euro. Note that participants were asked to respond to both of these conditions. The LEADING treatment requested participants to indicate their contribution (from an initial endowment of fifty 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 with sequential games (e.g. the ultimatum game). In the strategy method, experimental subjects take on all possible roles of 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, to limit the strategy space, we asked for conditional contributions in ten intervals with increments of five Euro, that is, participants had to indicate their contributions as followers for ten intervals of leader contributions, namely zero–five Euro, five–ten Euro, …, fourty-five–fifty 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, that is, the social optimum.

Due to the ongoing SARS–CoV–2 pandemic and the limited availability of farmers, we could not match participants in real life or online in real-time. Hence, we decided for a high-stakes one-shot game with ex-post matching for a payment of selected participants in an online survey format (see Sattler *et al.* 2022 for a discussion of the involved trade-offs). 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 which was part of a simplified pre-registration protocol[2] for the analysis).

No strict filters were applied to achieve a sufficiently large sample size. Participating farmers had to be involved in the management of a farm in Germany. Every tenth farmer received a payment based on the decisions in the public goods game (cf. Charness *et al.* 2016). The initial endowment had an expected value of five Euro and was comparable to similar studies such as Bouma *et al.* 2020. This is a large incentive for a survey, which lasted approximately 10 to 15 minutes on average. 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 presented scenario introduced participants to the context of the study ('The aim of this survey is to better understand the conditions under which farmers cooperate in agri-environmental measures.'), but we refrained from inducing more context into the scenarios and used abstract instructions with monetary amounts throughout to achieve salience and dominance (Smith 1976). Hence, our experiment can be classified as an artefactual field experiment (Harrison and List 2004).

## 2.4 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). We designed a survey in which we invited academics and practitioners to predict the outcomes of the public goods game with farmers.

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 of the European Commission) 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 goods games and agricultural policies in the European Union.

The survey was incentivised, and incentives were common knowledge. For every started 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 they were asked to predict). 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. As we received more than 200 (but less than 300) responses, we contacted three randomly selected participants, all of whom accepted the payment. The prediction part of the study was mentioned in the pre-registration, but no specific analysis was pre-registered for this part.

## 3. Results

### 3.1 Sample description and treatment effects

In total, 358 farmers completed the full questionnaire, and their responses were used for analysis. Summary statistics for farm and respondent characteristics are provided in Table 1.

If we compare the sample to the latest available statistics for the German farm population from 2016 (European Commission 2022), it becomes clear that our sample is biased towards larger farms and young male farm holders: The average utilized agricultural area per farm in Germany was 60.2 ha at the time (compared to 177 ha in our sample), the share of female farm holders was 10.2 per cent (compared to 6 per cent in our sample), and the share of farmers below 35 years was only 7.4 per cent (compared to 30.5 per cent in our data). Regarding risk attitudes, the average response to the risk attitude question of Dohmen *et al.* (2011) is close to a recent sample of farmers from the federal state of Hesse (=5.43 in Labajova *et al.* 2022 vs. 5.75 in our sample), but farmers in our sample are slightly

**Table 1.** Summary statistics for farmer participants in the public goods game experiment.

| | Median | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| = 1 of female | 0.00 | 0.06 | 0.23 | 0 | 1 |
| Age in years | 44.00 | 43.69 | 13.26 | 19 | 77 |
| Farm size in hectares | 87.00 | 176.97 | 339.35 | 0 | 2800 |
| = 1 if yearly income 50 to 100 thousand Euro | 0.00 | 0.41 | 0.49 | 0 | 1 |
| = 1 if yearly income above 100 thousand Euro | 0.00 | 0.15 | 0.36 | 0 | 1 |
| = 1 if farmers has ongoing cooperation in environmental management | 0.00 | 0.09 | 0.29 | 0 | 1 |
| Risk attitude on 11-point scale (Dohmen *et al.* 2011) | 6.00 | 5.75 | 1.97 | 0 | 10 |
| I enjoy working with other farmers. (1 = completely disagree, 7 = completely agree) | 5.00 | 5.02 | 1.45 | 1 | 7 |
| Achieving environmental goals in agriculture requires greater cooperation among farmers. (1 = completely disagree, 7 = completely agree) | 5.00 | 5.00 | 1.68 | 1 | 7 |
| Cooperation with other farmers is generally difficult. (1 = completely disagree, 7 = completely agree) | 4.00 | 3.98 | 1.57 | 1 | 7 |
| I frequently exchange ideas with other farmers on agricultural topics. (1 = completely disagree, 7 = completely agree) | 6.00 | 5.49 | 1.41 | 1 | 7 |

Source: Own calculations.

**Table 2.** Summary statistics on contributions in percent of the initial endowment by treatments.

| | Number of observations | Mean | Std. Dev. | Median | Minimum | Maximum |
|---|---|---|---|---|---|---|
| BASELINE | 72 | 69.56 | 25.56 | 60 | 0 | 100 |
| HETEROGENEITY | 71 | 65.21 | 22.93 | 70 | 0 | 100 |
| NORMS | 72 | 66 | 29.85 | 70 | 0 | 100 |
| OPTIMUM | 72 | 81.28 | 24.91 | 95 | 2 | 100 |
| LEADING | 71 | 62.65 | 20.24 | 65.42 | 2.33 | 100 |
| Pooled data across all treatments | 358 | 68.97 | 25.65 | 70 | 0 | 100 |

Source: own calculations.

more risk-seeking than the general population (see Labajova *et al.* 2022 for more data and discussion on this).

Table 2 displays summary statistics for contributions to the group account in per cent of the initial endowment. The last row presents the aggregate numbers pooled across all treatments. For the BASELINE, NORMS, and OPTIMUM treatments, contributions were directly taken from participants' responses. For HETEROGENEITY, we calculated the average 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 of a participant; three quarters were based on the followers' contributions conditioned on the distribution of leaders' contributions depending on the selected interval.

Formal testing across all treatments revealed that at least one of the treatment distributions is different from at least one other distribution (Kruskal–Wallis test; $\chi = 28.37$ with four degrees of freedom; $P < 0.01$). Pairwise tests did not show evidence of differences

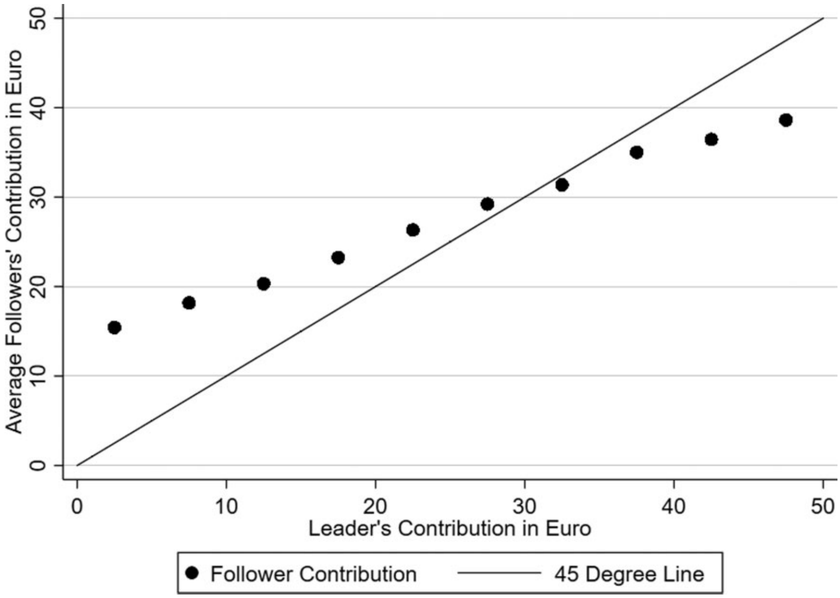


**Figure 2.** Average followers' contributions by Leaders' Contributions (Source: own calculations).

at $P < 0.05$ for any of the treatment combinations, with the only exception of the OPTIMUM treatment which was statistically significantly different from all other treatments (Two-sample Wilcoxon rank-sum (Mann–Whitney) tests; all $P < 0.01$).

Cooperation levels are high. Farmers contributed more than two-thirds of their endowment on average—a number that is typically around 50 per cent in laboratory studies (Zelmer 2003). In the HETEROGENEITY treatment, farmers contributed relatively more when they had low endowments (Mean = 17.56 Euro; SD = 20 Euro; equivalent to approximately 70.3 per cent 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 per cent 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 laboratory studies (Cherry *et al.* 2005; Martinangeli and Martinsson 2020).

Fig. 2 displays the average contributions of followers for each of the possible leader contributions in the LEADING treatment, starting from zero–five Euro up to the highest category of fourty-five–fifty Euro. A flat line would indicate that followers ignore what leaders do by not conditioning their contributions on the leader's contributions. A 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) would mean that followers 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.

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. 2 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, arriving at a factor of approximately 0.8 in the highest category.

We performed Wilcoxon matched-pairs signed-rank tests for all nine possible combinations of neighbouring contribution categories: Contributions in the zero–five Euro category are compared to the five–ten Euro category; contributions in the five–ten Euro category are compared to the ten–fifteen Euro category, and so forth. All nine tests showed large z-statistics (above 4) and yielded a very low *P*-value ($<0.01$) for the null hypothesis of equal distributions, that is, we cannot accept the null hypothesis of followers' contributions being independent of the leader's contribution. However, the empirical confidence intervals within each of the categories are very wide due to the high number of contributions of 0 per cent and 100 per cent, which results from the limited distribution of the dependent variable.

## 3.2 Exploring heterogeneity in contributions

In order to adjust for socio-demographics and to explore sources of heterogeneity in contributions, Table 3 presents three linear regressions.[3] Model (1) only includes indicators for the treatment; model (2) adds gender, age, farm size, and income category dummy variables, as well as an indicator of ongoing cooperation with organizations in the field of environmental management; model (3) adds risk attitudes (based on the 11-point scale by Dohmen *et al.* 2011) and other attitudes that shall proxy for the general willingness to cooperate with others and the specific interest in environmental cooperation. Note again that these regressions were not part of the pre-registration protocol.

One can see that the treatment effect estimates are relatively robust to the adjustment for co-variates. Among the co-variates, risk attitudes, as well as attitudes towards cooperation with other farmers in general and attitudes towards environmental cooperation have large and statistically significant effects. Notably, observables for which we detected biases of the sample in relation to the farm population (farm size, age, gender) do not have large or statistically significant impacts on cooperation levels.

In model (3), a one-point increase on the scale by Dohmen *et al.* (2011) towards a greater willingness to take risks leads to a 1.7 percentage point decrease in contributions. Moving up one point on the scale of agreement with the statement: 'Achieving environmental goals in agriculture requires greater cooperation among farmers.' even increases contributions by approximately 3.9 percentage points; whereas scepticism towards cooperation with others farmers ('Cooperation with other farmers is generally difficult.') lowers contributions by approximately 1.7 percentage points per one category movement along the scale. Overall, these results indicate a fair degree of external validity, since attitudinal statements related to the farming context appear to have a fairly large impact on behaviour in the public goods game.

## 3.3 Expert predictions

We received 1,060 completed predictions from 212 experts. Some respondents dropped out after the predictions, and there are a few missing variables, explaining the lower number of observations in some of the regression models presented below. A large share of respondents (23 per cent) 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 per cent) opened the survey from the link distributed within the Contracts2.0 research project.

Table 4 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 based on statements regarding 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

**Table 3.** Coefficient estimates on relative contributions (percentage of initial endowment) to the group account in the public goods game.

| | (1) | (2) | (3) |
|---|---|---|---|
| HETEROGENEITY | −4.34 | −4.58 | −3.55 |
| | (4.17) | (4.25) | (4.14) |
| NORMS | −3.56 | −4.72 | −5.12 |
| | (4.15) | (4.21) | (4.12) |
| OPTIMUM | 11.72*** | 13.09*** | 13.74*** |
| | (4.15) | (4.32) | (4.22) |
| LEADING | −6.90* | −6.97 | −5.53 |
| | (4.17) | (4.29) | (4.22) |
| = 1 of female | | −0.68 | −0.74 |
| | | (6.26) | (6.08) |
| Age in years | | 0.13 | 0.11 |
| | | (0.10) | (0.10) |
| Farm size in hectares | | −0.01 | −0.01 |
| | | (0.00) | (0.00) |
| = 1 if yearly income 50 to 100 thousand Euro | | 4.12 | 4.17 |
| | | (2.98) | (2.90) |
| = 1 if yearly income above 100 thousand Euro | | 3.57 | 3.97 |
| | | (4.00) | (3.90) |
| = 1 if farmers has ongoing cooperation in environmental management | | 8.69* | 5.36 |
| | | (4.91) | (4.78) |
| Risk attitude on 11-point scale (Dohmen *et al.* 2011) | | | −1.70** |
| | | | (0.70) |
| ***Included statements on cooperation (1 = completely disagree, 7 = completely agree)*** | | | |
| I enjoy working with other farmers. | | | −0.38 |
| | | | (1.20) |
| Achieving environmental goals in agriculture requires greater cooperation among farmers. | | | 3.88*** |
| | | | (0.86) |
| Cooperation with other farmers is generally difficult. | | | −1.67* |
| | | | (0.92) |
| I frequently exchange ideas with other farmers on agricultural topics. | | | −0.38 |
| | | | (1.15) |
| Constant | 69.56*** | 61.66*** | 63.27*** |
| | (2.94) | (5.81) | (9.85) |
| *N* | 358 | 338 | 333 |
| $R^2$ | 0.066 | 0.096 | 0.179 |
| F | 6.20 | 3.49 | 4.60 |
| Adj. $R^2$ | 0.06 | 0.07 | 0.14 |

Standard errors in parentheses; number of observations varies due to list-wise exclusion for missing values.
*$P < 0.10$, ** $P < 0.05$, *** $P < 0.01$.

agriculture (KNOWLEDGE AGRICULTURE) on a scale from strongly disagree (1) to strongly agree (7). The majority of respondents were working in academia (78 per cent). A background in Economics (60 per cent) 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.

Across all treatments, respondents are more pessimistic regarding farmers' contributions on average. They predict farmers to contribute 38 per cent of their initial endowment in the BASELINE condition, which is only approximately half of farmers' actual contri-

**Table 4.** 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.

**Table 5.** Regression results on predictive accuracy.

| Variable | OLS |
|---|---|
| HETEROGENEITY | 0.184 |
| | (1.001) |
| LEADING | −11.383*** |
| | (1.067) |
| NORMS | −15.673*** |
| | (1.030) |
| OPTIMUM | −10.939*** |
| | (1.184) |
| KNOWLEDGE EU CAP | −0.224 |
| | (0.613) |
| KNOWLEDGE AES | −0.549 |
| | (0.832) |
| KNOWLEDGE ECONOMICS | 1.306* |
| | (0.705) |
| KNOWLEDGE PGG | −1.321* |
| | (0.625) |
| KNOWLEDGE AGRICULTURE | 0.463 |
| | (0.809) |
| AGE | −0.264*** |
| | (0.078) |
| FEMALE | −2.719* |
| | (1.494) |
| ECON ACADEMIA | −1.861 |
| | (1.870) |
| Constant (BASELINE) | 34.507*** |
| | (1.626) |
| Observations | 980 |
| $R^2$ | 0.201 |
| Adjusted $R^2$ | 0.191 |
| F Statistic | 20.209*** (df = 12; 967) |

Note: Standard errors (clustered for respondents) in parentheses; *$P$ < 0.1, **$P$< 0.05, ***$P$<0.01; Source: own calculations.

bution (cf. Table 2). In comparison to the BASELINE predictions, experts are most pessimistic regarding the HETEROGENEITY treatment (34 per cent) and more optimistic for the LEADING (50 per cent), NORMS (55 per cent), and OPTIMUM (64 per cent) 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 smaller (see Table 5 for details). To further illustrate that experts

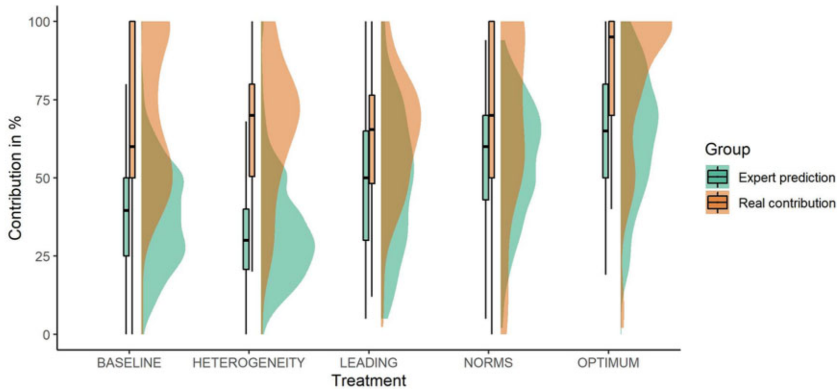


**Figure 3.** Distributions and box plots of relative contributions of farmers vs. expert predictions by treatments (Source: own calculations).

substantially underestimate farmers' cooperation levels, Fig. 3 overlaps smoothened distributions and box plots of actual contributions (orange) and expert predictions (green).

Table 5 presents results from a pooled linear ordinary least regression (OLS) with the absolute deviation in percentage points between the actual and predicted average treatment contribution as the dependent variable.[4] For instance, if the average farmer participant in the public goods contributed 80 per cent of her initial endowment and the expert predicted this average to be 55 per cent, this number would be 25 per cent. This has the advantage that the estimated coefficient of an independent variable in a linear model predicts the impact of a unit change in this variable to the predictive accuracy in percentage point deviation. For example, if the coefficient of FEMALE was estimated to be −3, it would mean that predictions of participants who identified as female are 3 percentage points more accurate than the reference category on average.

Note that these are the pooled data of all five predictions per person (standard errors are clustered per respondent to account for correlated errors of the five predictions per person). All independent variables are mean-centred (but not standardized) to allow for an easier interpretation for the average respondent.[5] Treatment dummies use BASELINE as the reference category, which under mean-centred covariates shows the deviation of percentage points from the prediction in the constant term of the model (which is the baseline accuracy). 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 (i.e. participants are off by approximately 34.5 percentage points on average). 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 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 × 1.3 = 7.8). There are larger effects for being an academic economist, female, and age. The model shows high $R^2$ and adjusted $R^2$ values, and a large and statistically significant F-statistic.

## 4. Discussion

Famers' contributions to the public good in our experiment substantially exceeded the 50 per cent of the initial endowment typically found in laboratory studies. This happened in spite of the anonymous online format that could have further boosted selfish behaviour. The directions of treatment effects were not always 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). We did not find statistically significant effects for these two treatments, albeit there was a drop in contributions of approximately four percentage points for the heterogeneity treatment.

To our surprise, the leading-by-example treatment also showed negative average effects. However, this effect was largely driven by the flat slope of the conditioned contributions of followers (see Fig. 2), that is, other things being equal, leading-by-example would be an effective cooperation-enhancing treatment if baseline contributions are small. In the heterogeneity treatment, farmers also contributed relatively more when they had low endowments. Against our expectations, highlighting high contributions of others (the social norm treatment) also showed negative effects, but this behaviour is consistent with the leading-by-example and heterogeneity treatments in the sense that large contributions of others lead people to decrease their relative contributions.

The only treatment that led to substantial and statistically significant increases in contributions was to emphasise the social optimum. Although one should not be tempted to jump to bold agricultural policy conclusions from the rather abstract setting of an online experiment with farmers, communicating joint benefits, that is, the gains from cooperation, might be a good strategy when setting up collective contracts.

Fine-tuning 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). In this context, it is a good sign that behaviour in the game was correlated with real-world behaviours and attitudes. As the regression results showed, engagement in environmental cooperation with organizations and positive attitudes towards cooperation with other farmers increased contributions to the group account in the game. A concern frequently raised in the design workshops of the game treatments was parallelism, that is, the link between game results and real-world behaviour. While one option is to frame the task in a familiar context (e.g. Rommel *et al.* 2019), another option is to link rewards more saliently with actions in experiments. 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, whereas Dessart *et al.* (2021) used a donation to an environmental organization to enhance parallelism in a study on environmentally-friendly farming practices among farmers in three EU member states. Future research may explore such options to delve deeper into the important question of parallelism in economic experiments with real-world decision-makers.

Our experiment involved delayed payments and an ex-post random matching of participants. While these conditions may not be ideal, and we used them only because of the ongoing SARS–CoV–2 pandemic, they also allowed us to gather a relatively large sample of farmers. Finding 358 farmers for participation in real-time experiments would have been very challenging. On the other hand, the asynchronous matching limited our abilities to study the evolution of cooperation over time. We can also not rule out that the anonymous matching affected results relative to a real-time interaction. Methodological tests of such 'strategy version' elements in experiments and greater insights on the impact of delayed payments in experiments with farmers would be an important route for future research.

Elicited expert predictions were substantially more pessimistic regarding farmers' cooperation levels when compared to the actual behaviour in the game. Experts were also generally more in line with the literature from experimental laboratory studies, where contributions reach approximately half of the endowment in a simple game (Zelmer 2003). Hence, predictions were not very accurate, albeit there were larger differences between treatments. Adjusting for the baseline, experts were far off with their average predictions of the baseline and heterogeneity treatment effect, as indicated by the large constant and small coefficient for HETEROGENEITY in Table 5. For all other treatments, expert predictions improved by more than ten percentage points on average (relative to the baseline). Our models could be extended by taking into account the variation in predictions or by distinguishing between overly optimistic and overly pessimistic predictions (as suggested in Schaak *et al.* 2022).

## 5. Policy implications and concluding remarks

The effective design and implementation of collective agri-environmental contracts is a challenging task. While the basic propensity of farmers to cooperate with each other is a necessary condition for setting up such contracts, it is not a sufficient condition. Important details and the complexity of a specific context can lead to failure of collective models even when the willingness to cooperate is large (Prager 2022). While it is a weakness of the abstract experimental approach not to acknowledge these real-world complexities, there are also important strengths. Experiments allow us to isolate small fragments of the real world and to manipulate and study them under controlled conditions. Here, we have seen that German farmers are more likely to cooperate when being reminded of the gains from cooperation. By comparing groups of participants, we can also gain important insights about heterogeneity in behaviour. In our study, participating farmers were substantially more likely to cooperate than students in laboratory experiments. Farmers with positive attitudes towards environmental cooperation were also more likely to cooperate. Policy-makers should communicate objectives and potentials of new approaches to farmers before implementing novel policies, taking into account such attitudes and behavioural barriers which play a role across a wider range of contexts (Meijer *et al.* 2015). In this sense, our results are evidence against an important barrier of collective contracts—the perceived lack in the willingness of farmers to cooperate with each other. Other methods such as social-ecologic modelling and case studies can complement our experiment, pick up on the results, and take them further by taking into account wider social and environmental contexts. Along these lines, one might investigate how cooperation differs under and interacts with varying biophysical conditions in peatlands or protected areas, taking into account additional farm and farmer characteristics.

Our study has shown that experts are overly pessimistic regarding farmers' willingness to cooperate. This scepticism is mirrored in the German agricultural policy debate (Latacz-Lohmann *et al.* 2019; Feindt *et al.* 2020). We need an increased willingness to test new approaches to break path dependencies. Living lab approaches could be a useful step in this direction (McPhee *et al.* 2021). Not all experts are alike. Women who indicated good knowledge about the public goods game performed better. Sector-specific knowledge on agriculture, the common agricultural policy, or agri-environmental schemes did not substantially improve predictions, whereas a sound understanding of public goods games characterized experts with more accurate predictions. Whether or not these features of experts are important in other contexts remains to be seen, because the prediction of research results is still new to the agricultural economics field (Schaak *et al.* 2022). It would be instructive to follow up on how experts would update beliefs after confrontation with the evidence (cf. Vivalt & Coville 2020). According to our results, more optimism and openness regarding novel collective approaches to agri-environmental policy in Germany is pertinent.

## Supplementary material

Supplementary data, including data and replication code, are available at *Q Open* online.


## Funding

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 818190. Julian Sagebiel gratefully acknowledges the support of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig funded by the German Research Foundation (FZT 118).

## Acknowledgments

We thank all workshop participants for sharing their insights and contributing to the design of the experiment. We are grateful for helpful and encouraging comments from two anonymous reviewers and the guest editor Nadja El Benni.


## End Notes

1 Among the discussed treatments, manipulations in group size were one option. 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 use smaller (e.g. $a = 1.5$) or larger (e.g. $a = 3$) multiplication factors. 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 See https://aspredicted.org/PKN_ABL for the pre-registration. We only pre-registered basic analysis in the form of non-parametric tests for group and pairwise comparisons. All other analysis should be considered explorative.

3 Formally, the models are defined as follows: $Y = a + bT + cX + \in$, where Y is the contribution to the public good in percentage points, T is a set of dummies indicating the four treatments (BASELINE is the reference category), X are the farmer characteristics and attitudes, b and c are parameter vectors to be estimated, a is a constant term to be estimated, and $\in$ is an independently and identically distributed error term.

4 More formally, we estimate a model of the form $Y = a + bT + cX + \in$, where Y is the absolute deviation in percentage points, T is a set of dummies indicating the (to-be-predicted) treatments, X are the covariates of the experts, as indicated in Table 4, b and c are parameter vectors to be estimated, a is a constant term to be estimated, and $\in$ is an independently and identically distributed error term (we cluster standard errors at the level of the respondent to enhance independence, as predictions are likely correlated within the expert).

5 Formally, for any independent variable x from Table 4 we use $x - \bar{x}$ (instead of just x), where $\bar{x}$ is the arithmetic mean of variable x. Note that this is different from standardizing coefficients (which would also imply dividing this difference by the standard deviation) and which would result in standardized coefficient estimates. The sole reason for this procedure is to impose an average of zero on all independent variables, so that the constant becomes the average of the dependent variable, and hence the

BASELINE deviation between the prediction and the real average in percentage points. Not doing this calculation would only change the estimate for the constant, but coefficient estimates on covariates would be exactly the same.



## References

Banerjee S., De Vries F. P., Hanley N. and 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*, 96:1009–29.

Bouma J. A., Nguyen T. B., Van Der Heijden E. and Dijk J. J. (2020) 'Analysing Group Contract Design Using a Threshold Public Goods Experiment', *European Review of Agricultural Economics*, 47:1250–75.

Brandts J. and Charness G. (2000) 'Hot vs. Cold: Sequential Responses and Preference Stability in Experimental Games', *Experimental Economics*, 2:227–38.

Carlsson F., Johansson-Stenman O. and Khanh Nam P. (2015) 'Funding a New Bridge in Rural Vietnam: A Field Experiment on Social Influence and Default Contributions', *Oxford Economic Papers*, 67:987–1014.

Charness G., Gneezy U. and Halladay B. (2016) 'Experimental Methods: Pay One or Pay All', *Journal of Economic Behavior & Organization*, 131:141–50.

Cherry T. L., Kroll S. and Shogren J. F. (2005) 'The Impact of Endowment Heterogeneity and Origin on Public Good Contributions: Evidence from the Lab', *Journal of Economic Behavior & Organization*, 57:357–65.

Colen L., Gomez y Paloma S., Latacz-Lohmann U., Lefebvre M., Préget R. and 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*, 64:667–94.

DellaVigna S. and Pope D. (2018a) 'Predicting Experimental Results: Who Knows What?', *Journal of Political Economy*, 126:2410–56.

DellaVigna S. and Pope D. (2018b) 'What Motivates Effort? Evidence and Expert Forecasts', *The Review of Economic Studies*, 85:1029–69.

DellaVigna S., Pope D. and Vivalt E. (2019) 'Predict Science to Improve Science', *Science*, 366:428–9.

Dessart F. J., Barreiro-Hurlé J. and van Bavel R. (2019) 'Behavioural Factors Affecting the Adoption of Sustainable Farming Practices: A Policy-Oriented Review', *European Review of Agricultural Economics*, 46:417–71.

Dessart F. J., Rommel J., Barreiro-Hurlé J., Thomas F., Rodríguez-Entrena M., Espinosa M. and van Bavel R. (2021) *Farmers and the new green architecture of the EU common agricultural policy: a behavioural experiment*. Publication Office of the European Union.

Dohmen T., Falk A., Huffman D., Sunde U., Schupp J. and Wagner G. G. (2011) 'Individual Risk Attitudes: Measurement, Determinants, and Behavioural Consequences', *Journal of the European Economic Association*, 9:522–50.

European Commission (2022) *Germany: Agriculture statistical factsheet*. https://agriculture.ec.europa.eu/cap-my-country/performance-agricultural-policy/agriculture-country/eu-country-factsheets_en.

Feindt P. H., Dietze V., Krämer C., Thomas F., Lukat E. and Häger A. (2020) *Kooperationsprogramm Natur und Landwirtschaft—Eine Möglichkeit zur Stärkung des Naturschutzes und der Zusammenarbeit von Akteuren im ländlichen Raum*. https://ageconsearch.umn.edu/record/305584.

Fischbacher U., Gächter S. and Fehr E. (2001) 'Are People Conditionally Cooperative? Evidence from a Public Goods Experiment', *Economics Letters*, 71:397–404.

Grüner S., Lehberger M., Hirschauer N. and Mußhoff O. (2022) 'How (Un)informative Are Experiments with Students for Other Social Groups? A Study of Agricultural Students and Farmers', *Australian Journal of Agricultural and Resource Economics,* 66:471–504.

Harrison G. W. and List J. A. (2004) 'Field Experiments', *Journal of Economic Literature*, 42:1009–55.

Hasler B., Termansen M., Nielsen H. Ø., Daugbjerg C., Wunder S. and Latacz-Lohmann U. (2022) 'European Agri-Environmental Policy: Evolution, Effectiveness, and Challenges', *Review of Environmental Economics and Policy,* 16:105–25.

Huber R., Zabel A., Schleiffer M., Vroege W., Brändle J. M. and Finger R. (2021) 'Conservation Costs Drive Enrolment in Agglomeration Bonus Scheme', *Ecological Economics,* 186:107064.

Isaac R. M., Walker J. M. and Thomas S. H. (1984) 'Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations', *Public Choice*, 43:113–49.

Krämer J. E. and Wätzold F. (2018) 'The Agglomeration Bonus in Practice—An Exploratory Assessment of the Swiss Network Bonus', *Journal for Nature Conservation,* 43:126–35.

Kuhfuss L., Begg G., Flanigan S., Hawes C. and 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.

Kuhfuss L., Préget R., Thoyer S. and Hanley N. (2016) 'Nudging Farmers to Enrol Land into Agri-Environmental Schemes: The Role of a Collective Bonus', *European Review of Agricultural Economics*, 43:609–36.

Kuhfuss L., Préget R., Thoyer S., de Vries F. P. and Hanley N. (2022) 'Enhancing Spatial Coordination in Payment for Ecosystem Services Schemes with Non-Pecuniary Preferences', *Ecological Economics*, 192:107271.

Labajova K., Höhler J., Lagerkvist C. J., Müller J. and Rommel J. (2022) 'Illusion of Control in Farmers' Investment and Financing Decisions', *Agricultural Finance Review,* 82:675–89.

Latacz-Lohmann U., Balmann A., Birner R., Christen O., Gauly M., Grajewski R. and 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*, 227: 1–100.

Ledyard J. (1995) 'Public Goods: A Survey of Experimental Research', In: Kagel, John H., Roth, Alvin E. (eds.) *The Handbook of Experimental Economics*, pp. 111–94. Princeton: Princeton University Press.

Lefebvre M., Barreiro-Hurlé J., Blanchflower C., Colen L., Kuhfuss L., Rommel J. and Thoyer S. (2021) 'Can Economic Experiments Contribute to a More Effective CAP?', *EuroChoices,* 20:42–9.

Levati M. V., Sutter M. and Van der Heijden E. (2007). Leading by Example in a Public Goods Experiment with Heterogeneity and Incomplete Information', *Journal of Conflict Resolution*, 51: 793–818.

Maart-Noelck S. C. and Mußhoff O. (2014) 'Measuring the Risk Attitude of Decision-Makers: Are There Differences Between Groups of Methods and Persons?', *Australian Journal of Agricultural and Resource Economics,* 58:336–52.

Martinangeli A. F. and Martinsson P. (2020) 'We, The Rich: Inequality, Identity, and Cooperation', *Journal of Economic Behavior & Organization*, 178:249–66.

McPhee C., Bancerz M., Mambrini-Doudet M., Chrétien F., Huyghe C. and Gracia-Garza J. (2021) 'The Defining Characteristics of Agro-ecosystem Living Labs', *Sustainability,* 13:1718.

Meraner M., Mußhoff O. and Finger R. (2018) 'Using Involvement to Reduce Inconsistencies in Risk Preference Elicitation', *Journal of Behavioural and Experimental Economics*, 73:22–33.

Meijer S. S., Catacutan D., Sileshi G. W. and Nieuwenhuis M. (2015) 'Tree Planting by Smallholder Farmers in Malawi: Using the Theory of Planned Behaviour to Examine the Relationship Between Attitudes and Behaviour', *Journal of Environmental Psychology,* 43:1–12.

Milkman K. L., Gandhi L., Patel M. S., Graci H. N., Gromet D. M., Ho H. and Duckworth A. L. (2022) 'A 680,000-Person Megastudy of Nudges to Encourage Vaccination in Pharmacies', *Proceedings of the National Academy of Sciences*, 119:e2115126119.

Müller M. (2020) 'Leadership in Agricultural Machinery Circles: Experimental Evidence From Tajikistan', *Australian Journal of Agricultural and Resource Economics*, 64:533–54.

Nguyen C., Latacz-Lohmann U., Hanley N., Schilizzi S. and Iftekhar S. (2022) 'Spatial Coordination Incentives for Landscape-Scale Environmental Management: A Systematic Review', *Land Use Policy,* 114:105936.

Parkhurst G. M., Shogren J. F., Bastian C., Kivi P., Donner J. and Smith R. B. (2002) 'Agglomeration Bonus: An Incentive Mechanism to Reunite Fragmented Habitat for Biodiversity Conservation', *Ecological Economics*, 41:305–28.

Pe'er G., Zinngrebe Y., Hauck J., Schindler S., Dittrich A., Zingg S. and Lakner S. (2017) 'Adding Some Green to the Greening: Improving the EU's Ecological Focus Areas for Biodiversity and Farmers', *Conservation Letters,* 10:517–30.

Prager K. (2022) 'Implementing Policy Interventions to Support Farmer Cooperation for Environmental Benefits', *Land Use Policy,* 119:106182.

Rommel J., van Bussel L., le Clech S., Czajkowski M., Höhler J., Matzdorf B. and Zagorska K. (2021) 'Environmental Cooperation at Landscape Scales: First Insights from Co-Designing Public Goods Games with Farmers in Four EU Member States', https://pub.epsilon.slu.se/23419/.

Rommel J., Hermann D., Müller M. and Mußhoff O. (2019) 'Contextual Framing and Monetary Incentives in Field Experiments on Risk Preferences: Evidence from German Farmers', *Journal of Agricultural Economics*, 70:408–25.

Sattler C., Rommel J., Chen C., García-Llorente M., Gutiérrez-Briceño I., Prager K. and Kelemen E. (2022) 'Participatory Research in Times of COVID-19 and Beyond: Adjusting your Methodological Toolkits', *One Earth*, 5:62–73.

Schaak H., Rommel J., Sagebiel J., Barreiro-Hurlé J., Bougherara D., Cembalo L., Cerjak M., Čop T., Czajkowski M., Espinosa-Goded M., Höhler J., Lagerkvist C-J., Rodriguez-Entrena M., Tensi A., Thoyer S., Tomić Maksan M., Vecchio R. and Zagórska K. (2022) *How well can experts predict farmers' risk preferences?* https://econpapers.repec.org/paper/agsaaea22/322194.htm.

Smith V. L. (1976) 'Experimental Economics: Induced Value Theory', *The American Economic Review*, 66:274–9.

Thomas F., Midler E., Lefebvre M. and Engel S. (2019) 'Greening the Common Agricultural Policy: A Behavioural Perspective and Lab-in-the-Field Experiment in Germany', *European Review of Agricultural Economics*, 46:367–92.

Villamayor-Tomas S., Sagebiel J. and 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*, 84:200–15.

Villamayor-Tomas S., Sagebiel J., Rommel J. and Olschewski R. (2021) 'Types of Collective Action Problems and Farmers' Willingness to Accept Agri-Environmental Schemes in Switzerland', *Ecosystem Services*, 50:101304.

Vivalt E. and Coville A. (2020) 'How Do Policymakers Update Their Beliefs?', Mimeo, http://evavivalt.com/wp-content/uploads/How-Do-Policymakers-Update.pdf.

Zelmer J. (2003) 'Linear Public Goods Experiments: A Meta-Analysis', *Experimental Economics*, 6:299–310.