



HAL
open science

Can collective conditionality improve agri-environmental contracts? From lab to field experiments

Kristin Limbach, Anne Rozan, Philippe Le Coent, Raphaële Préget, Sophie S. Thoyer

► To cite this version:

Kristin Limbach, Anne Rozan, Philippe Le Coent, Raphaële Préget, Sophie S. Thoyer. Can collective conditionality improve agri-environmental contracts? From lab to field experiments. Review of Agricultural, Food and Environmental Studies, In press, 10.1007/s41130-023-00198-2 . hal-04114539

HAL Id: hal-04114539

<https://brgm.hal.science/hal-04114539v1>

Submitted on 21 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Can collective conditionality improve agri-environmental contracts? From lab to field experiments

March 2023

Abstract

Individual subsidy payments that are conditional on a collective contribution threshold could provide a viable resolution to the insufficient and dispersed adoption of agri-environmental contracts aiming at attaining environmental quality targets. Indeed, in a decontextualized laboratory experiment based on a threshold public good game (TPGG), Le Coent et al. (2014) offer promising results regarding a conditional subsidy compared to an unconditional subsidy (i.e. the standard subsidy in existing agri-environmental schemes). In this article, we propose to improve the external validity of these results by transposing this laboratory experiment to the "field" with farmers. To do so, we carry out a contextualized lab-in-field experiment with farmers by explicitly mentioning agri-environmental contracts and water quality. Our results show that farmers cooperate even more successfully than students and sustain more efficient outcomes over time. Considering a between-subject, both in the lab with students, and in the "lab-in-field" experiment with farmers, our results indicate that average group contributions are not significantly different under the conditional subsidy mechanism and under the unconditional subsidy mechanism. Indeed the conditional subsidy introduces two behavioural responses (perceived risks and initial beliefs on others' contributions) having opposite effects on contributions, and compensating each other. The conditional incentive mechanism thus shows promising potential as a tool for agri-environmental policy since it avoids the pay-for-nothing trap of the unconditional subsidy mechanism without discouraging contributions.

Keywords: *Agri-Environmental Schemes, Threshold Public Good Game, Collective Conditionality, Lab-in-Field Experiment.*

JEL classifications: *Q58, C92, C93*

1 Introduction

The growing consensus on the adverse effects of agricultural intensification on the environment in Europe has led to the introduction of a diversity of policy instruments to incentivize more environmentally-friendly farming practices. The Common agricultural policy (CAP) proposes agri-environmental schemes (AES), which are based on voluntary contracts signed with individual farmers, in which they commit to adopt or maintain pro-environmental practices in return for a compensatory payment. Evaluations of these programs reveal that adoption rates remain relatively low and that environmental impacts are weaker than expected (Oréade-Brèche, 2005; Barbut & Baschet, 2005; Cour des Comptes Européenne, 2011; Uthes & Matzdorf, 2013, Grüner et al., 2022).

Several explanations are advanced to explain these unsatisfying outcomes. One of them is that most contracts are individual contracts which do not include any collective dimension. Land managers are paid based on the pro-environmental actions they undertake (e.g. reduction of input use, land set-aside) independent of the ultimate environmental outcome obtained. This individual take-up may be particularly problematic in situations with environmental threshold effects, i.e. when the production of environmental benefits does not increase linearly with environmental efforts but presents discontinuities (Perrings & Pearce, 1994; Muradian, 2001) or convexities. For example, risks of eutrophication of water bodies are reduced only if the water concentration of phosphorus and nitrogen falls below a certain threshold. A sufficient large area of the watershed must be cultivated with agricultural practices reducing nitrate and phosphorus leakage to ensure that this concentration threshold is attained. If the sum of efforts by individual farmers to reduce their use of fertilizer is not sufficient, the water quality does not improve and public funds are spent without obtaining the desired ecological outcomes (Dupraz et al., 2007).

The thresholds inherent to the production of some environmental public goods (e.g. groundwater quality improvement, habitat reinstatement, protection of river banks) necessitate collective efforts at a landscape scale that often stretches across individual farmland. By neglecting the collective effort needed to reach the environmental threshold, public funds are often invested in vain. In this article, we intend to test agri-environmental contracts in which individual payments are calculated on the basis of individual efforts (e.g., per hectares enrolled in the AES), but are conditioned to the collective attainment of a participation threshold (in terms of the total number of hectares enrolled in the area of interest), guaranteeing that the environmental improvement is obtained.

The main aim of this article is to analyze whether these agri-environmental contracts with collective conditionality on participation, can be more effective in terms of environmental outcomes

and more efficient in terms of public spending compared to the existing unconditional contracts. The issue at stake is therefore to measure whether the deterrent effect on potential participants of a no-payment outcome (if the threshold is not attained) is stronger or weaker than the incentive to coordinate at the threshold level. Aversion to risk and beliefs on the behavior of others are therefore key factors to consider in the success of such mechanism.

Le Coent et al. (2014) have tested this conditional subsidy (CS) mechanism in a decontextualized laboratory experiment with students, based on a threshold public good game scenario, and compared it to the classic unconditional subsidy (US) paid to each contributor even if the public good is not produced, and to a no-subsidy situation. Subjects investing in the public good cannot be certain that the public good will be effectively produced, since it depends on the total level of contributions in the group. The CS treatment adds another source of loss when the threshold is not attained since the subsidy then is not paid to any of the contributors. Yet these two treatments have the same Nash predictions (Le Coent et al, 2014). The results of the experiments indicate that both subsidy schemes are effective in incentivizing the attainment of the threshold, compared to the no-subsidy situation. More interestingly, the conditional subsidy performs as well as the unconditional subsidy. From a regulator point of view, the financial efficiency of the conditional subsidy is even superior, mainly due to the fact that subsidies are not spent when the public good is not produced, which is politically attractive especially when budget constraints are tight. However, Le Coent et al. (2014) show that the average results hide heterogeneous patterns of behavior and require further investigations on individual behavior.

In this paper, we mobilize their methodology, data and results to enrich the analysis in two directions:

1) We compare the results found in the decontextualized laboratory experiment (hereafter referred to as the lab experiment) of Le Coent et al. (2014) to the results of a contextualized lab-in-field setting (hereafter referred to as the field experiment) that we conducted with farmers, in order to test the external validity of their conclusions. For this purpose, we used the same protocol but we transposed the initial economic experiment, based on a threshold public good game (TPGG), into the context of an agri-environmental scheme with the goal to improve water quality at the scale of the watershed. Our artefactual sample is composed of farmers of the French region of Alsace invited to participate to the experiment;

2) We analyze farmers' and students' strategies at the individual level, and not just at group level, in order to gain a better understanding of what drives contribution behavior and to understand why

some groups manage to cooperate at or above the threshold, while others fail. In particular, we include in the analysis the elicited risk preferences and beliefs about others' contributions, both in the lab and in the lab-in-field experiments. Since TPGG are coordination games, the assurance problem, as described by Sen (1967) and Runge (1984), might arise and explain individual behavior. Each player has a strategic uncertainty on the contributions of the other group members. Below a certain value of beliefs about others' contributions, there is no interest to contribute to the public good because the expected probability that the threshold will be reached is too low. When this value is reached, it is optimal to contribute to reach the threshold. Runge (1984) argues that the purpose of political and economic institutions, dealing with public goods, is the coordination of expectations and beliefs. Institutions should therefore provide assurance regarding the behavior of others, and thereby help mitigating the assurance problem. Our hypothesis is that subsidy schemes can play that role. The conditional subsidy scheme, by strengthening the psychological focus on the threshold, could even play a stronger role on beliefs. On the opposite, considering that the subsidy is conditioned to the behavior of others in the CS treatment, subjects may be more cautious when they choose their contribution in the first period, and risk averse subjects may be more reluctant to contribute.

Our results highlight that early stages of implementation are fundamental for the success of a conditional subsidy mechanism and that participants' risk aversion can limit its effectiveness, suggesting the importance of accompanying its introduction with facilitation activities. We also show (albeit with a small sample of farmers) that in the field setting, average contribution rates are sustained well above the threshold and, across all periods, and are strictly exceeding those observed in the lab. In addition, farmers being less risk averse than students on average, introducing a collective conditionality on the payment of subsidies appears to be a promising tool for incentivizing the adoption of agricultural practices when a participation threshold needs to be collectively reached.

This article therefore brings two types of contributions: first, to the literature on lab-to-field generalizability; and second, to the literature on agri-environmental contracts with an aim to provide insights on payment incentives for steering collective participation.

The paper is organized as follows: section 2 provides an overview of relevant findings from the literature in relation to the two aims of our study, section 3 describes the methodology underlying the experiments and section 4 presents the analysis of the experimental data. Section 5 discusses the results and section 6 concludes.

2 Related Literature

Previous research has addressed the environmental threshold dynamics relevant to AES theoretically (e.g. Dupraz et al., 2009) and has proposed new elements for the design of AES that include a collective dimension or that require coordination among farmers : minimum participation rates (Zavalloni et al., 2019), conditional bonuses rewarding collective participation (Kuhfuss et al., 2016), and more globally spatial coordination incentives for landscape-scale environmental management (see Nguyen et al. (2022) for a systematic review). To our knowledge only Le Coent et al. (2014) have studied a collectively conditional subsidy scheme in a threshold public good game so far.

With respect to our first research aim, the following review focuses on a discussion of lab-to-field generalizability looking at the external validity for behavior observed in TPGG, and comparing risk and cooperation attitudes for various artefactual samples.

2.1 From lab to field: generalizability of experiments

For academics and policy-makers alike, the question of generalizability of experimental results is relevant: do behavioral patterns as tested in the lab reproduce in the real world? Can one draw reliable implications and credible recommendations for real life settings from findings collected in the lab with decontextualized protocols played with students (List, 2007; Levitt & List, 2007a)?

Previous literature has already addressed some of the questions on lab-to-field generalizability using different experimental settings and subject pools. Experiments on cooperation behavior evoking different degrees of context as collected in Camerer (2011) seem to attest quite strong context dependencies for cooperation in lab-field contrasts (e.g. Benz & Meier (2008) in a donation setting or Stoop et al. (2012) in a common pool resource setting). Nevertheless, Camerer (2011) claims that “there is no replicated evidence that experimental economics lab data fail to generalize to central empirical features of field data (when the lab features are deliberately closely matched to the field features)”.

There are several reasons why the issue of contextualization is not trivial. When contrasting cooperative, competitive and neutral Prisoner’s Dilemma settings, Engel and Rand (2014) find that the tendency of participants in the neutral setting is to assume a cooperative frame by default. That is, the decisions made under the neutral frame resemble those in the cooperative setting, whereas the competitive frame significantly reduces cooperation relative to neutral. Indeed, a ‘clean’ decontextualized setup may not exist (Engel & Rand, 2014, Harrison & List, 2004). It is shown that participants tend to implicitly project their own story onto ‘neutrally’ framed experiments (Engel &

Rand, 2014). Thus, even carefully designed lab instructions are not immune to subjectively projected frames that are outside of the control of the experimenter. To that extent, a comparison between findings from several context settings may be a large advantage as to identify common mechanisms in decision making and to enrich context specific inferences (Levitt & List, 2007b).

When changing the context, notably to make the context more credible to artefactual samples, the challenge clearly lies in the degree of comparability of conditions created in lab and field (Camerer, 2011; Viceisza, 2016; Alekseev et al, 2017; Thoyer & Préget, 2019). Take the time frame to decision making, for instance. While several years of enrollment with an AES contract leave long periods of reflection between yearly decisions, the lab handles 20 decision rounds within 2 hours. Moreover, information asymmetries due to unknown threshold levels, efficient effort levels, real transaction costs and heterogeneity in group composition, may provoke dynamics over time in the field context that are left unaccounted for in a lab setting. When extrapolating results from the lab to field situations, it is also shown that simple tasks seem to produce coherent behavioral patterns while behavior in complex tasks does not reproduce in the field (Saldarriaga-Isaza et al., 2019). Therefore, lab experiments with students should be understood as a “test bed” for behavioral tendencies of the general population (Saldarriaga-Isaza et al., 2019).

In the specific case of TPGG experiments, studies on lab-to-field comparisons find that on average, an artefactual sample contributes more at group level across all treatment variations and is more successful in sustaining cooperation at or above the threshold than a student sample (punishment and external advisor with frame and miners in Saldarriaga-Isaza et al. (2019), membership fee without frame and farmers in Bchir (2014), baseline without frame and nurses in Cadsby and Maynes (1998)). Saldarriaga-Isaza et al (2019) show that lab experiments on social dilemmas can be externally valid. In a framed threshold public good game with university students and artisanal gold miners, they find more behavioural convergences than divergences between students and miners. Yet, their experimental game is quite different from ours since they do not introduce any individual subsidy conditional on a collective contribution threshold.

Both subject pools show oscillation around the efficient Nash equilibrium (NE) in a TPGG without a rebate¹ as found in Saldarriaga-Isaza et al. (2019) and Marks and Croson (1998). For treatments that include a rebate rule, this changes. Coordination towards the Pareto-efficient outcome is observed in the field with contributions well above the threshold (Saldarriaga-Isaza et al., 2019).

¹ Rebate rules are used to compensate subjects for their excess contributions when aggregate contributions are above the threshold.

2.2 Risk Attitudes and cooperation: students versus artefactual samples

In economic experiments, risk attitudes are usually assessed using the Holt and Laury (2002) lottery procedure, upon which Vollmer et al. (2017) conclude that it is externally valid in assessing farmer's risk preferences. Saldarriaga-Isaza et al. (2019) have found students to portray similar attitudes as compared to the artefactual sample of miners when assessed in risk attitudes, empathy and trust. Grüner et al. (2022) vary the size of stakes in the Holt and Laury (2002) procedure and find near identical risk attitudes between agricultural students and farmers across all variations. Likewise, other studies find similar risk preferences using a constant relative risk aversion (CRRA) function on foresters and forestry students (Sauter et al., 2015) and using rank dependent utility theory with consumers. In the case of Maart-Noelck and Musshoff (2014)'s study, conducted in Germany, however, students and farmers differ significantly in their assessed risk preferences, with students being more risk-averse than farmers. As pointed out by Drichoutis and Koundouri (2012), this result is different under the assumption of rank-dependent utility: they show that using this method to reveal risk preferences, students exhibit risk neutrality while the general population shows more risk-loving.

As a conclusion, we can hypothesize that students and farmers will display similar risk attitudes on average, although farmers may be more trusting and act more cooperatively than students in the TPPG, thus leading to a greater rate of success in reaching the threshold. What remains unknown to us is the effect of contextualization: will farmers playing the TPPG game with a conditional subsidy setting act as cooperatively as in the unconditional subsidy setting?

3 The experiments

3.1 The game and the subsidy treatments

The game underlying the experiment is a threshold public good game (TPGG). In both the lab and field experiments, we build groups of 4 players. Each player i is endowed with 20 units and must decide how many units to contribute (C_i) to a public account which benefits all members of the group, but only if the group contribution threshold is reached. The other tokens ($20 - C_i$) are invested in a private account. The threshold in the public account is set at 40 units, i.e. 50% of the group's total endowment ($0.5 \times 4 \times 20$). The marginal per capita return (MPCR) for contributions to the public good is set at 0.3. A linearly increasing production function of the public good is assumed, as in Isaac et al. (1989): it keeps increasing beyond the provision point. This standard TPGG with no subsidy has several Nash equilibria (NE):

- an *inefficient* NE: $\sum_{i=1}^4 C_i = 0$
- *multiple efficient* NE: $\sum_{i=1}^4 C_i = 40$ (they pareto-dominate the *inefficient* NE)

The inefficient Nash equilibrium exists when all contributions are zero. Here the threshold cannot be reached by a subject alone. The efficient Nash equilibria are located exactly at threshold level. This game presents a social dilemma as the Pareto optimum holds for full contribution levels ($\sum_{i=1}^4 C_i = 80$) which maximizes the total payoff but is not an equilibrium.

Le Coent et al (2014) test two variations of incentive schemes: an unconditional subsidy (US) and a conditional subsidy (CS). They show that both subsidy schemes are welfare-improving compared to the above no-subsidy situation. In this paper, we are primarily interested in comparing the performance of the two subsidy schemes, so we choose to focus only on two treatments, CS and US. Individual payoffs, π_i , depend on the individual's contribution, C_i , the subsidy scheme and the group contribution, $\sum_{i=1}^4 C_i$, which can either be above or below the threshold of 40. They are described as follows:

Unconditional subsidy (US)

$$\pi_{iUS} \begin{cases} 20 - 0.7C_i & \text{if } \sum_{i=1}^4 C_i < 40 \\ 20 - 0.7C_i + 0.3 \sum_{i=1}^4 C_i & \text{if } \sum_{i=1}^4 C_i \geq 40 \end{cases}$$

Conditional subsidy (CS)

$$\pi_{iCS} \begin{cases} 20 - C_i & \text{if } \sum_{i=1}^4 C_i < 40 \\ 20 - 0.7C_i + 0.3 \sum_{i=1}^4 C_i & \text{if } \sum_{i=1}^4 C_i \geq 40 \end{cases}$$

As emphasized in Le Coent et al (2014), theoretical predictions for the US and CS treatments are the same as for the no-subsidy scenario, i.e. a strong free-riding equilibrium and a multiplicity of efficient equilibria for which $\sum_{i=1}^4 C_i = 40$.

3.2 From lab to field: the contextualization of the protocol

In order to embed the game into a context mimicking the decisions that farmers have to make when deciding to enroll into an agri-environmental scheme, we transpose the lab experimental protocol into an agricultural frame. The stylized experimental design features are summarized in table 1, which shows how we have contextualized the protocol described in Le Coent et al. (2014), yet maintaining the exact same structure of the underlying game.

Real context of AES	Transposition to the laboratory (decontextualized)	Transposition to the field (contextualized)
Threshold environmental public good such as water quality or biodiversity conservation	Generic version of threshold public good	Framed version of threshold public good: groundwater catchment area, water quality, nitrates issue
Decisions: - Ha enrolled in AES (pro-environmental practices) - Ha not enrolled in AES (conventional practices)	Unit = token Initial endowment of 20 tokens - Tokens invested in the public account - Tokens invested in the private account	Unit = hectare Hypothetical scenario where each farmer has a 20 ha farm - Ha enrolled in pro-environmental low input practices - Ha with conventional practices
Traditional agri-environmental scheme: payment to each farmer per ha enrolled independent of other farmers' enrollment rates.	Unconditional subsidy scheme (US): individual subsidy per token contributed to the public account, independent of other subjects' contribution.	Unconditional subsidy scheme (US): individual subsidy per hectare enrolled in good low input practices, independent of other farmers' enrollment.
Agri-environmental scheme with collective conditionality: payment to each farmer per ha enrolled, provided the sum of ha enrolled by all farmers is at least equal to the collective threshold necessary for ensuring an improvement of environmental quality	Conditional subsidy scheme (CS): individual subsidy per token contributed to the public account, provided the sum of tokens in the public account is at least equal to the collective threshold of 40 tokens	Conditional subsidy scheme (CS): individual subsidy per hectare enrolled in pro-environmental low input practices, provided the sum of ha enrolled by all farmers of the group is at least equal to the collective threshold of 40 ha
Stakeholders involved in the implementation of AES: farmers	Subject sample: mainly students	Artefactual sample: farmers
Choice environment: the farmer community where individuals are socially embedded in, see and communicate with each other, within the same water basin	The university laboratory: participants are seated in isolated computer cubicles, no visual contact and no communication	Farmers seated in the same room with individual iPads, visual contact but no communication allowed

Table 1: Contextual transposition from lab to field.

Table 1 shows that the decision frame for contributions varied according to lab or field setting. Students in the decontextualized lab experiment were asked to decide how many tokens of their initial endowment they wanted to contribute to a public account. Depending on the treatment and threshold

attainment, contents of the public account were then redistributed among the four group members. Farmers in the contextualized field setting were asked to choose the number of hectares from their endowed agricultural land they wanted to enroll in pro-environmental low-input practices. Each group represented a collective of four farmers whose farm is in the recharge area of the same groundwater body. The environmental public good is the improvement of the groundwater quality, which has to reach a given threshold to be allowed into the potabilization process. Earnings in the experiments are expressed in points, later converted into euros with a pre-announced conversion rate. The treatment US represents the traditional AES in which farmers receive an individual subsidy proportional to the agricultural surface enrolled in the scheme. The treatment CS represents the new collectively conditional subsidy scheme, which grants an individual subsidy proportional to each farmer's agricultural surface enrolled in the scheme only if the aggregate enrolled surface by all the farmers of a group meets or exceeds the threshold that is necessary to ensure an improvement in environmental quality, i.e. the provision of the public good. We simplify the reality of AES by assuming that farmers adopting pro-environmental low-input practices incur implementation costs (they may be due to additional investments or to foregone revenue associated, for example to lower yields) which are identical across farmers. The subsidy scheme compensates 30% of these costs. The return from the public good account is the positive utility that each individual perceives from an improvement of the overall groundwater quality.

3.3 Experimental Procedure

In both lab and field sessions, participants arriving to the sessions were seated individually and asked to restrain from communication throughout the experiment.

We first elicit risk preferences with a series of lottery combinations (adapted from Holt and Laury, 2002). Participants were informed that, at the end of the experiment, one of the games would be randomly drawn and its inherent gamble would determine the earnings for this part of the experiment. Subjects' risk aversion was characterized by their "switching point", i.e. the first game for which they chose the uncertain option instead of the safe option. The later they switched, the higher their risk aversion score. Individuals with multiple switching points were considered to have an undetermined risk aversion.

We then formed groups of 4 participants through random and anonymous composition. The groups were maintained throughout the end of the experiment. We run a between-within subject design, i.e. every subject played successively two sequence treatments. However, in this article, we will only

consider data from the first sequence in a between-subject analysis.¹ Each sequence consisted of a threshold public good game over 10 periods, with a different subsidy scheme. At each period, two choices had to be indicated by participant: their estimation (or belief) of the other three group members' contributions for the upcoming period, and their personal contribution to the public good. Both choices were monetary incentivized. To incentivize subjects' beliefs, we used a protocol similar to that of Fischbacher and Gächter (2010): if the subject's estimation of the sum of the other three group members' contribution was accurate, he earned 5 points; if it was 1 unit away from the actual contribution of others, he earned 4 points; if it was 2 units away, he earned 3 points; if it was more than 2 units away, he earned nothing. These points were added to the points earned in the contribution game. At the end of each period, feedback on the aggregate group contribution to the public account and on individual payoff was displayed to each participant. At the end of the 10 periods, the earnings for that sequence were displayed as an aggregate over all the periods of that sequence. Participants had been informed that, at the end of the experiment, one of the two sequences would be chosen at random for payment, in addition to the lottery outcome and the participation fee. Corresponding instructions were distributed separately at the beginning of each part (risk elicitation, sequence 1 and sequence 2).

The instructions (in English) for the field experimental sessions can be found in the Appendix.

3.4 Participants and Earnings

For the laboratory experiment, as presented in Le Coent et al. (2014), 220 students were recruited in 2013 and 2014 via ORSEE (Greiner, 2004) and all the experimental sessions were conducted at the Laboratoire d'Economie Expérimentale de Montpellier (LEEM) in the University of Montpellier. All participants self-selected into both the database and the experiment. In the decontextualized university laboratory, participants were seated in isolated computer cubicles where they had no visual contact with other participants. The sessions lasted a maximum of 2 hours and the average earning was €15.90 with a standard deviation of €3. In addition, a show-up fee of €2 was provided for students of the lab campus and of €6 for students from a different campus.

For the field experiment, conducted in 2020, farmers were invited via telephone and e-mail to participate in sessions at three different locations in the French region of Alsace, according to a least

¹ It would be interesting to study the impact of the introduction of a conditional subsidy in a context in which an unconditional subsidy already exists. Unfortunately, we did not run the necessary control treatment sequence with unconditional subsidy in both successive sequences to control for potential order effects or learning impact across sequences.

distance criterion given their postal address. A database of 350 addresses had been obtained from the INSEE (Institut national de la statistique et des études économiques). Due to Covid restrictions; we only succeeded to gather a sample of 24 participants, which were clustered by geographical area, thus implying that some farmers knew each other and that all self-selected into the experiment. Facilities for the experimental sessions were provided by the regional agriculture Chamber (Chambre d’Agriculture d’Alsace) and composed of individual desks (one per participant), each equipped with an iPad from the mobile laboratory of the Laboratoire d’Economie Expérimentale de Strasbourg (LEES). All farmers were seated in the same room and could make visual contact with each other. The sessions lasted a maximum of 2 hours and the average earning was €16.38 with a standard deviation of €2.34 in addition to the show-up fee of €20.

4 Results

Section 4.1 focuses on the analysis of group contributions while section 4.2 provides parametric estimations of individual contributions including treatment dummies, elicited beliefs and risk aversion as predictors.

4.1 Comparison of average group contributions

As explained before, we are only interested in the comparison of the two subsidy treatments (US and CS) played in the first sequence. Therefore, we do not consider the lab sessions of the Le Coent et al (2014) paper where the no subsidy treatment was played in the first sequence. We keep only the first sequence of sessions where the unconditional subsidy treatment (US) and the conditional subsidy treatment (CS) were played.

Table 2 depicts the number of participants, the number of groups and the average group contributions with standard deviations for each treatment (US, CS), distinguishing the lab and the field settings. It shows that average contributions are higher in the field than in the lab experiment and display lower standard deviations.

Treatment	Number of participants	Number of groups	Average group contribution (standard deviation)
Lab experiment			
Unconditional subsidy (US)	68	17	41.62 (22.02)
Conditional subsidy (CS)	72	18	42.01 (23.50)
Lab	140	35	
Field experiment			
Unconditional subsidy (US)	12	3	59.57 (13.72)

Conditional subsidy (CS)	12	3	62.37 (4.90)
Field	24	6	

Table 2: Averages (and standard deviations) for group contribution. Note: Lab data are from Le Coent et al. (2014) and field data are new original data.

Group contribution averages for each of the 10 periods are graphically portrayed in figure 1. To better visualize the type of subsidy, the group contribution is shown as a solid line for the unconditional subsidy (US) and as a dashed line for the conditional subsidy (CS). All field treatment averages (pictured as \times) show higher contributions as compared to the lab (pictured as \bullet). Moreover, we can see from the graph that average group contributions in the laboratory evolve slightly above the threshold in the first periods, but decline below the threshold in the very last periods of the sequences. The frequency of success of the public good production in the lab is about 65%. In the field setting, this frequency of success is well above 95%. More precisely, group contributions in the field experiment are always above threshold (around 70% of endowment), apart for one observation: in the last period for one group in the US treatment. Nevertheless, farmers' group contributions also do decrease on average with later periods of the game.

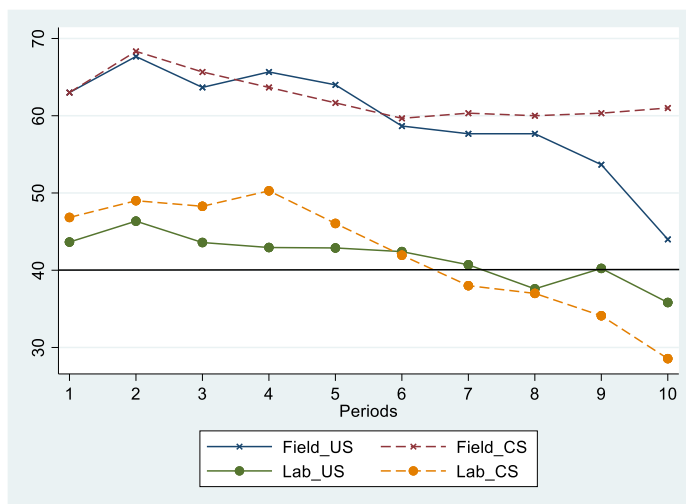


Figure 1: Laboratory and Field experiments' average group contributions per treatment (US: unconditional subsidy, CS: conditional subsidy).

Wilcoxon-Mann-Whitney (WMW) tests are performed to assess non-parametrically whether the observed differences in group contributions between treatments are significant (at a 5% threshold). Similar to Stoop et al. (2012)'s cooperation contrast between fishermen and students in lab-in-field experiments, farmers seem to be consistently more prosocial than students. The difference between lab and field settings shows highly significant (WMW, $p < 0.001$). Within the lab and the field setting,

Wilcoxon-Mann-Whitney tests do not show significant difference in group contributions between the two subsidy treatments (US vs. CS).

In other words, the conditional subsidy is as effective as the unconditional subsidy in both settings. This result confirms the potential interest of this subsidy scheme for public policies since the conditional subsidy scheme avoids the risk of paying subsidies when the public good is not provided.

These results at the group level are however quite variable across groups. In the field setting, all the groups reached the threshold right from the first period. However, in the lab experiment, the graphical representation of the dynamics of the group contributions seems to reveal that the result of the first period is key in explaining group behavior (see Figure 2). We observe that intra-group variability is rather low while inter-group variability is high. For the CS treatment, two types of groups emerge: groups that manage to coordinate over the threshold and groups that do not. Put in a simple way, the graph seems to indicate that if a group manages to coordinate in the first period, its aggregate group contribution remains over the threshold for most of the sequence, until end-game effects start appearing. However, if the group does not manage to coordinate at the threshold in the first period, its contributions rapidly converge at the zero-contribution equilibrium. The US presents similar characteristics, however groups that do not coordinate above the threshold do not converge as rapidly towards the zero-contribution equilibrium. Intermediate levels below the threshold are maintained, probably because subjects' losses are limited when the threshold is not reached, thanks to the unconditional subsidy mechanism.

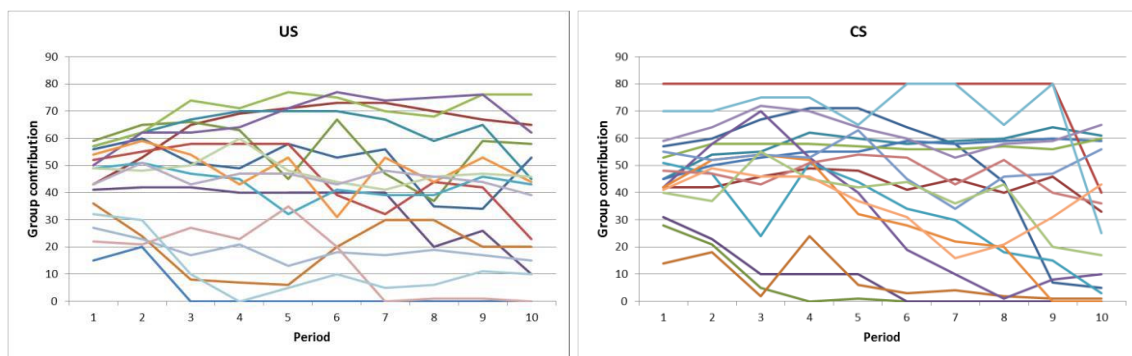


Figure 2: Aggregate group contributions for the two subsidy treatments (1 group=1 data series) of the lab experiment

In the following sub-section, we investigate individual contributions to better understand the role of the first period and how behavioral factors (individual risk aversion and beliefs about the behavior of other members of the group) may underpin the performance of this threshold public good game with subsidies.

4.2 Determinants of Individual Contributions

We propose a parametric estimation to analyze the determinants of individual contributions C_{it} . The explanatory variables are summarized in table 3.

Variables	Values	Interpretation
field	Binary {0, 1}	Field experiment = 1 (ref. lab experiment = 0)
CS	Binary {0, 1}	Conditional Subsidy treatment CS = 1 (ref. unconditional subsidy treatment US = 0)
period	Discrete {1:10}	indicates period
success	Binary {1, 0}	Threshold reached in period 1 = 1 (ref. threshold not reached in period 1 = 0)
belief	Discrete {0:60}	Belief about the aggregate contribution of the 3 other members in the upcoming period
risk	Discrete {1:10}	Risk aversion ¹

Table 3: Description of variables used in econometric analysis

The participants' belief elicitation over group members' contributions (adapted from Fischbacher and Gächter, 2010) show that, compared to students, farmers indicate more optimistic beliefs over their group members' intentions to cooperate. The observed difference may be attributed to lower past contributions observed in the lab, but also to a higher control of anonymity in the lab. As mentioned, farmers enrolled as subjects in the field experiment could know each other as they came from the same geographical area, although they did not know the exact composition of their group.

Table 4 summarizes the regression results. Models (1) - (5) of the random effects panel regression have individual contributions as its dependent variable and the generic specification is given below. The index t corresponds to the time/period at which subjects' decisions were made, ranging from 1 to 10. The standard errors for individual observations are clustered by group.

Generic Model:

$$C_{it} = \alpha_0 + \alpha_1 field_i + \alpha_2 CS_{it} + \alpha_3 period_t + \alpha_4 success_i + \alpha_5 belief_{it} + \alpha_6 risk_i + \gamma_i + \epsilon_{it}$$

We base our panel data analysis on the random effects (RE) estimator, which yields a weighted average of the between and within estimates. Contrary to the fixed effects (FE) model, the time invariant effects are not eliminated and individual-specific effects that are not explicitly controlled for rest in the error term. That is, differences among individuals are considered random (instead of fixed

¹ Participants' aversion to risk is characterized by the rank of the switching point (see section 3.3). For individuals that switched several times the risk aversion value was considered missing (12 missing values in the lab, and 2 missing values in the field). The risk aversion indicator is spread from 1 (less risk averse) to 10 (most risk averse) with an average of 7.03 and a standard deviation of 1.44 for 128 subjects in the lab and an average of 5.63 and a standard deviation of 2.48 for 22 farmers in the field setting.

and possible to estimate) and are accounted for by the individual-specific component of the error term, γ_i . Since our subjects (students and farmers) were randomly assigned to treatments, we assume that all individual unobserved confounders (such as trusting behavior) are equally distributed across the two treatments and therefore cannot explain differences in outcomes between the two treatments.

Individual contribution	(1)	(2)	(3)	(4)	(5)
Intercept	12.488*** (1.120)	12.533*** (1.206)	5.598*** (0.812)	5.286*** (0.835)	13.775*** (2.108)
field	4.790*** (1.103)	4.487** (1.796)	2.423** (0.992)	2.658*** (0.517)	5.036*** (0.993)
CS	0.185 (1.487)	-0.097 (1.126)	-1.007 (0.879)	-0.340 (0.723)	0.534 (1.512)
period	-0.387*** (0.082)	-0.387*** (0.082)	-0.464*** (0.093)	-0.256*** (0.061)	-0.429*** (0.086)
field*CS		0.603 (2.199)			
success			10.454*** (0.766)		
belief				0.200*** (0.013)	
risk					-0.177 (.277)
Nb. of obs.	1,640	1,640	1,476	1,640	1,500
Nb. of subjects	164	164	164	164	150

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses, clustered by group.

In Model 3, first period is excluded because success cannot be observed

Table 4: Random Effects Panel Regression. Average contribution at individual level.

Consistent with previous non-parametric results, we observe from model (1) that the dummy variable *field* is positive and significant and the dummy for the subsidy treatment *CS* is not significant. In addition, model (2) shows that the interaction term *field*CS* is not significant suggesting there is no specific impact of the CS treatment in the field experiment. The period coefficient being negative and highly significant formally attests the decline in contribution levels over periods. Model (3) confirms that reaching the threshold in the first period (variable *success*) has a strong significant positive effect on subsequent individual contributions. If subjects contribute above the threshold in the first period, they are likely to keep on doing so in the following periods. On the contrary, if the public good is not produced in the first period, contributions are significantly lower in subsequent periods and rapidly converge to the zero-contribution equilibrium (see also Figure 2).

In model (4), the beliefs over other group members' contributions in the upcoming period are strong predictors of individuals' choices for contribution. Most participants are conditional

contributors. They adjust their own contribution in the same direction as their belief: if they expect that others will contribute less, then they also contribute less.

Finally, we could conjecture that the more risk averse subjects are, the less they will contribute for fear of losing their contribution (or part of their contribution in the US treatment). Risk aversion differs on average between the two samples, with students being more risk-averse than farmers ($\chi^2(1) = 35.29$; p -value < 0.000). Similar results have been obtained with French farmers by Bocquého et al. (2014). This difference might explain why farmers manage to coordinate better than student subjects. Yet, according to model (5) in Table 4, participants' risk aversion does not significantly influence contributions.

Given the importance of the belief variable and the success dummy, we investigate in more detail the period 1. As indicated in section 2.3, we hypothesize that the CS treatment can help to strengthen the psychological focus on the threshold and to reinforce beliefs on others' participation, thereby increasing the chances of success in reaching the threshold. In this respect, what happens in the first period is crucial as it sets out beliefs for the following periods.

In order to verify these hypotheses, we analyze individual contributions and beliefs on others' contribution in the first period. The value of the variable *belief* cannot be influenced by the other members of the group since we focus only on the very first period of the experiment. In Table 5, we compare first period's average beliefs in lab and field groups, using a t-test. We also run a regression analysis with *belief* (in period 1) as dependent variable (table 6 column (a)). Results show that beliefs on others' contribution are indeed significantly higher with the conditional subsidy than with the unconditional subsidy. These higher beliefs in CS than in US seem to counterbalance the *a priori* greater risk in CS than in US. This positive effect of the conditional subsidy on the expectation about others' behavior before the first period is played is probably one of the explanations of its performance.

Subsidy treatment	Number of subjects	Average belief on others' contributions in period 1	Standard Deviation	T-test
US	80	32.04	1.45	belief CS > belief US**
CS	84	36.76	4.39	

Table 5: T-test comparison of beliefs about others' contribution between subsidy treatments in the first period, ** significant at 5%

Period 1	belief (a)	individual contribution (b)	individual contribution (c)
Intercept	27.476***	12.699***	5.800***

	(4.768)	(2.1350)	(1.6950)
field	6.086*	4.546***	3.018**
	(3.110)	(1.392)	(1.172)
CS	4.984**	0.757	-0.495
	(2.112)	(.9456)	(.8004)
risk	0.604	-0.222	-0.373
	(0.6035)	(.2926)	(.2438)
belief			0.251***
			(.0310)
Nb. of obs./ subjects	150	150	150

*** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses.

Table 6: Linear regression of belief and individual contribution on the first period. Dependent variable: individual contributions.

Regression (b) shows that individual contributions in the first period are significantly higher in the field setting than in the lab setting but are not significantly different in US and CS treatments. Yet, regression (c) shows that beliefs in the first period have a positive significant impact on individual contributions of the first period. Finally, as in Table 4, regression (c) of Table 6 shows that the variable *risk* does not affect either *belief* or *individual contribution*.

5 Discussion

Based on the above analysis, four key results can be formulated from this study.

Result 1: Farmers contribute more than students do in the lab. This result is coherent with findings from other contexts of experimental collective action games, such as the findings reported by Stoop et al. (2012), Saldarriaga-Isaza et al. (2019), Bchir (2014), and Cadsby and Maynes (1998) who also observed higher provision rates and more sustained cooperation in the field with artefactual participants as compared to the lab with student participants. One explanation could be that farmers in the same geographical area identify more with the economic situation of the members of their group and the environmental impact of their choices.

Result 2: Groups of farmers in the field experiment reach a higher rate of success in attaining the threshold than students in the laboratory experiment, whose average group contribution drops below the threshold in the last periods despite the subsidy mechanism. This difference could be explained by two factors. First, the average group contribution is generally higher among farmers, who might just be better cooperators (Bchir, 2014). Second, different motivations may be introduced in the two different samples due to the respective context (Levitt & List, 2007b). Students most likely seek to reap free riding benefits in the absence of subsequent punishment during the last periods of the experiment. Farmers in contextualized frames may be motivated by other elements than just payoffs, such as

environmental concern and solidarity with their group members. Agribashing, the tendency of the media to blame the agricultural sector for environmental dilemmas, has also led to a detriment in the perceived image of farmers. It can thus be expected that farmers, when investigated in experiments, surveys or interviews, will try to polish their image with respect to environmental matters. Farmers may also anticipate subsequent relationships with their fellow farmers outside the experiment which limits end-game effects in anticipation of potential future punishment.

Result 3: On average, the conditional subsidy scheme leads to about the same contributions in both lab and field as compared to the unconditional subsidy scheme. This result is consistent with theoretical model predictions. However, we could have expected lower contributions with a conditional subsidy due to a higher risk of not coordinating to reach the threshold.

Result 4: Under the CS mechanism, the reduction of contributions induced by the higher perceived risk of not reaching the threshold is compensated by increased contributions associated to higher beliefs about the contributions of others in the first period.

This confirms that belief in others' cooperative intentions is an important factor in the success and durability of cooperation for threshold attainment in environmental TPGG. Miners are found to be less forgiving than students in a TPGG with punishment possibility (Saldarriaga-Isaza et al., 2019) as they continue to punish free-riders over more periods. Reasons that have been suggested for conditional cooperation are numerous. For example, social norms can explain trust and cooperation through (i) a desire for social conformity, (ii) reciprocity motives and (iii) signaling the importance of a public good (Frey & Meier, 2004). Policy will do wisely to take those social aspects into account in farmers' decision-making for AES with a collective threshold. Next to financial rewards, trust and social norms show to have strong modulating influence for farmers' behavior within a group (Dessart et al, 2019).

By gradually approaching the field context, these results inform us about the external validity of findings from the lab. In our examination of the transposition from decontextualized lab to framed lab-in-field experiment, we do find support for the argument by Levitt and List (2007a) that decisions in experiments are more complex than a simple financial assessment of choice prospects and rather depend on social factors such as others' behavior, the context and the selection procedure. Any experimental result thus needs to be interpreted with respect to its potential biases and shortcomings imposed by the experimental setting. Moving between lab and field settings can thus enhance and complement our understanding of behavior from particular situations and provide a more comprehensive picture.

There are three important aspects to the experimental design that need to be discussed in light of interpretation of the results obtained.

First, with the current experimental setup (i.e. in absence of controlling for environmental preferences) it is difficult to say what participants' true underlying motivation is. In particular, it would be of interest to disentangle the two motives of (i) empathy and sense of group belonging and (ii) environmental concern. Empathy could be important, especially for farmers of the same region. Environmental concern can be induced by the context (water quality) brought in the field experiment, and because farmers are expected to have a stronger link to nature. Explicitly assessing these motives in field settings could inform policy makers about the existing intrinsic willingness of farmers to enroll in AES with collective threshold.

Second, it is noteworthy for the interpretation of our results and the related discussion on generalizability that, in this experimental setting, the public good is measured by monetary stakes. Therefore, public externalities to the provision of the public good are for the sole benefit of the group as only the financial returns to the public account are relevant. Natural field experiments on conditional subsidy schemes would be useful to investigate behavior facing real environmental stakes that are relevant for society at large.

Third, the experiment uses a relatively low marginal per capita return (MPCR = 0.3). In a way this may reflect farmer's conventional perception of a rather low individual benefit from an improvement of environmental quality in return to changing their agricultural practices as outlined in the AES. This may become more intuitive when considering that any environmental improvement happens only gradually over the medium to long term. This is particularly true for the transfer of pollutants to aquifers that may take several years. Farmers' investments in changing their agricultural practices, however, are immediate, so that delayed benefits are measured against relatively high transaction and investment costs and potentially lower or less profitable yields. The experimental literature shows significant variations in contributions in response to variations of the MPCR. Most studies find an increase in contributions when augmenting the return to the public good (e.g. Isaac et al., 1989), as long as group size is kept stable and rather small. It may thus be that increasing the MPCR of this experiment could introduce even more pronounced group contributions.

6 Conclusion

Agri-environmental schemes are criticized because of their limited impact and their low cost-efficiency. One of the reasons for this disappointing outcome is that contracts are implemented without taking into account potential threshold environmental effects and the need to coordinate environmental efforts at a pertinent scale. We used the lab experimental results of Le Coent et al. (2014) to test a new form of agri-environmental contracts for which the payment is conditioned to the

collective attainment of the environmental threshold. In order to strengthen external validity, these results, obtained with students in the laboratory, are confirmed using field experiment with farmers. In this study we have transposed the TPGG framework of the decontextualized lab experiment into a framed field experiment to study the reproducibility of behavior under a choice setting that imitates the social dilemma inherent to environmental threshold public goods.

In the lab, a positive facilitating impact of both subsidy schemes on cooperation has been demonstrated (Le Coent et al., 2014). In the field, we observe average group contributions that are consistently higher than those of the analogue decontextualized lab treatment and that remain above the threshold value across all periods. For both samples, the conditional subsidy scheme elicits about the same contribution levels as compared to the unconditional subsidy scheme, and guarantees that no public money is spent when participation is insufficient to generate environmental benefits. These results indicate that the conditional subsidy scheme is worth exploring further as a payment rule that could avoid unnecessary public spending without discouraging participation by farmers. Yet, more field data is needed to draw better and more robust inferences (Camerer, 2011, Levitt & List, 2007b) upon (1) the effectiveness of the conditional subsidy scheme when subjects (farmers) can talk to each other and share their intentions and (2) farmers' preferences in order to verify and disentangle risk attitudes, trust, as well as pro-social and environmental motives in the decision to participate in the AES.

Transposing strictly the experimental TPGG to the real-world context would mean that farmers would receive a subsidy for their pro-environmental efforts, only in cases when the environmental collective threshold is attained. Despite the encouraging results found in our lab-in-field experiment, we anticipate that this system may be considered unfair by farmers (since it does penalize those who have accepted to provide a costly effort when others have not done so). It may not be easily accepted by farmers. An alternative option would be to ask farmers to declare their intentions to participate in anticipation and to open the contractualization process only when the targeted participation threshold is reached (Limbach & Rozan, 2022). This situation would alleviate the risk for farmers not to receive either the contract payment nor the benefits of the public good. Another option would be to condition only a proportion of the subsidy to the attainment of the environmental threshold. This conditional bonus, paid as an additional incentive for all participating farmers when the collective participation threshold is reached, has been already tested in a choice experiment by Kuhfuss et al. (2016) with promising results. Indeed their results show that farmers have a preference for contracts offering a conditional bonus, and are therefore prepared to enroll for lower individual subsidies, leading to lower public spending even when the bonus is paid. There are all types of variants that can be tested around

the idea of this bonus payment such as a contract with a lower subsidy paid when participation remains below the threshold, and a discretely increasing payment above the threshold. Such payment systems could help boost participation, including by encouraging farmers to convince their neighbors to sign up as well in order to increase their respective subsidy levels. This would be an extension of the present article, that could be tested directly with farmers, either with a qualitative survey or with a choice experiment comparing their stated intentions to enroll for different versions of progressive payment rules.

References

- Alekseev, A., Charness G., Gneezy, U (2017) Experimental methods: when and why contextual instructions are important, *Journal of Economic Behaviour and Organization*, 134, 48-59
- Barbut, L. and Baschet, J.F. (2005). L'évaluation de la politique de soutien à l'agroenvironnement. *Notes et études économiques*, 22, 37-68.
- Bchir, M. A. (2014). From the lab to the field: an experimental investigation of the provision of a club good. *Journal of Behavioral and Experimental Economics*, 51:57-67.
- Benz, M. and Meier, S. (2008). Do people behave in experiments as in the field?—evidence from donations. *Experimental economics*, 11(3):268-281.
- Bocquého, G., Jacquet, F., and Reynaud, A. (2014). Expected utility or prospect theory maximisers? assessing farmers' risk behaviour from field-experiment data. *European Review of Agricultural Economics*, 41(1):135-172.
- Cadsby, C. B. and Maynes, E. (1998). Choosing between a socially efficient and free-riding equilibrium: Nurses versus economics and business students. *Journal of Economic Behavior & Organization*, 37(2):183-192.
- Camerer, C. (2011). The promise and success of lab-field generalizability in experimental economics: A critical reply to levitt and list. *Available at SSRN 1977749*.
- Cour des Comptes Européenne (2011), L'aide agro-environnementale est-elle conçue et gérée de manière satisfaisante ? Rapport special n°7, ISSN 1831-0850, 83 pages.
- Dessart, F., 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.
- Drichoutis, A. C. and Koundouri, P. (2012). Estimating risk attitudes in conventional and artefactual lab experiments: The importance of the underlying assumptions. *Economics Discussion Paper*, (2012-19).
- Dupraz, P., Latouche, K., and Turpin, N. (2007). Programmes agri-environnementaux en présence d'effets de seuils. *Cahiers d'Economie et de Sociologie Rurales*, (82-83):6-32.
- Dupraz, P., Latouche, K., and Turpin, N. (2009). Threshold effect and co-ordination of agrienvironmental efforts. *Journal of environmental planning and management*, 52(5):613-630.
- Engel, C. and Rand, D. G. (2014). What does "clean" really mean? the implicit framing of decontextualized experiments. *Economics Letters*, 122(3):386-389.
- Fischbacher, U. and Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American economic review*, 100(1):541-56.
- Frey, B. S. and Meier, S. (2004). Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment. *American Economic Review*, 94(5):1717-1722.
- Greiner, B. (2004) "An Online Recruitment System for Economic Experiments" in *Forschung und wissenschaftliches Rechnen* by Kurt Kremer, Volker Macho, Eds., 79-93.
- Grüner, S., Lehberger, M., Hirschauer, N., and Musshoff, O. (2022). How (un-)informative are experiments with "standard subjects" for other social groups? - the case of agricultural students and farmers. *Australian Journal of Agricultural and Resource Economics* (in print).

- Harrison, G. W. and List, J. A. (2004). Field experiments. *Journal of Economic literature*, 42(4):1009–1055.
- Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655.
- Isaac, R. M., Schmidtz, D., and Walker, J. M. (1989). The assurance problem in a laboratory market. *Public choice*, 62(3):217–236.
- Kuhfuss, L., Préget, R., Thoyer, S., and Hanley, N. (2016). Nudging farmers to enrol land into agrienvironmental schemes: the role of a collective bonus. *European Review of Agricultural Economics*, 43(4):609–636.
- Le Coent, P., Préget, R., and Thoyer, S. (2014). Why pay for nothing? an experiment on a conditional subsidy scheme in a threshold public good game. *Economics Bulletin*, 34(3):1976–1989.
- Levitt, S. D. and List, J. A. (2007a). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic perspectives*, 21(2):153–174.
- Levitt, S. D. and List, J. A. (2007b). On the generalizability of lab behaviour to the field. *Canadian Journal of Economics/Revue canadienne d'économique*, 40(2):347–370.
- Limbach, K. and Rozan, A. (2022). Coordinating Farms in Collective Agri-Environmental Schemes: The Role of Conditional Incentives. Miméo.
- List, J. A. (2007). Field experiments: a bridge between lab and naturally occurring data. *The BE Journal of Economic Analysis & Policy*, 5(2).
- 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(3):336–352.
- Marks, M. and Croson, R. (1998). Alternative rebate rules in the provision of a threshold public good: An experimental investigation. *Journal of public Economics*, 67(2):195–220.
- Marks, M., Lehr, D., and Brastow, R. (2006). Cooperation versus free riding in a threshold public goods classroom experiment. *The Journal of Economic Education*, 37(2):156–170.
- Muradian R. (2001). Ecological thresholds: a survey. *Ecological Economics* 38 : 7 – 24.
- Nguyen, C., Latacz-Lohmann, U., Hanley, N., Schilizzi, S., Iftekhar, S. (2022). Spatial Coordination Incentives for landscape-scale environmental management: A systematic review. *Land Use Policy*, 114(105936):1–15.
- Oréade-Brèche (2005), Evaluation des mesures agro-environnementales.
- Perrings, C., Pearce, D. (1994). Threshold effects and incentives for the conservation of biodiversity. *Environ Resource Econ* 4, 13–28. <https://doi.org/10.1007/BF00691930>
- Runge, C. (1984). Institutions and free-rider: the assurance problem in collective action. *The Journal of politics*, 46.
- Saldarriaga-Isaza, A., Villegas-Palacio, C., and Arango, S. (2019). Chipping in for a cleaner technology: Experimental evidence from a framed threshold public good game with students and artisanal miners. *Journal of behavioral and experimental economics*, 78:10–16.
- Sauter, P., Hermann, D., and Mußhoff, O. (2015). Risk attitudes of foresters, farmers and students: An experimental multimethod comparison. Technical report, Diskussionsbeitrag.
- Sen, A. (1967). Isolation, assurance and the social rate of discount. *The Quarterly Journal of Economics*, 81:112–124.

- Stoop, J., Noussair, C. N., and Van Soest, D. (2012). From the lab to the field: Cooperation among fishermen. *Journal of Political Economy*, 120(6):1027–1056. 1722.
- Thoyer S., Préget R., (2019). Enriching the CAP evaluation toolbox with experimental approaches: introduction to the special issue, *European Review of Agricultural Economics*, Vol 46(3), pp 347-366, <https://doi.org/10.1093/erae/jbz024>
- Uthes, S., Matzdorf, B. (2013) Studies on Agri-environmental Measures: A Survey of the Literature. *Environmental Management* **51**, 251–266 (2013). <https://doi.org/10.1007/s00267-012-9959-6>
- Viceisza, A.C.G. (2016). Creating a lab in the field: Economics experiments for policymaking. *Journal of Economic Surveys* 30, 835–854.
- Vollmer, E., Hermann, D., and Mußhoff, O. (2017). Is the risk attitude measured with the holt and laury task reflected in farmers' production risk? *European Review of Agricultural Economics*, 44(3):399–424.
- Zavalloni, M., Raggi, M., and Viaggi, D. (2019). Agri-environmental policies and public goods: An assessment of coalition incentives and minimum participation rules. *Environmental and resource economics*, 72(4):1023–1040.

7 Appendix: Instructions for the lab-in-field experiment