Received: 20 September 2022 | Accepted: 12 January 2023

DOI: 10.1111/1477-9552.12529

ORIGINAL ARTICLE

JAE Journal of Agricultural Economics

External validity of economic experiments on Agri-environmental scheme design

Marie Ferré[1,2] | **Stefanie Engel**[3] | **Elisabeth Gsottbauer**[4]

[1]UMR Innovation, French Agricultural Research Center for International Development (Cirad), Montpellier, France

[2]University of Montpellier, Montpellier, France

[3]School of Business Administration and Economics and Institute for Environmental Systems Research, University of Osnabrück, Osnabrück, Germany

[4]Institute of Public Finance, University of Innsbruck, Innsbruck, Austria

Correspondence

Marie Ferré, UMR Innovation, French Agricultural Research Center for International, Development (Cirad), Montpellier, France.
Email: marie.ferre@cirad.fr

Funding information

Alexander von Humboldt-Stiftung; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 406840-143145; Swiss National Science Foundation; Alexander von Humboldt Foundation; Federal Ministry of Education and Research

[Correction added on 2nd March 2023, after first online publication: The affiliation for the second author has been updated in this version.]

Abstract

The use of laboratory experiments to study issues in agricultural policy has grown in prominence within the fields of agricultural and environmental economics. Such experiments are often conducted with university students and framed in an abstract manner. This raises questions about whether the findings of these experiments provide reliable insights on the behaviour of actual agents in real settings. We contribute to this methodological debate by analysing the impacts of sample population and framing on behaviour in the experiment and on two policy effects: the direction and the magnitude of the policy impact. We also examine the channels through which differences in results may occur. For this, we test if behaviour is correlated with a set of covariates collected from our samples, including socio-demographics, social and risk preferences. Our main finding is that the type of subject significantly affects the magnitude of the policy impact. The two populations differ substantially in the representation of key characteristics and preferences, which in consequence affects behaviour in the experiment. We find no significant impact of framing on behaviour.

KEYWORDS

agglomeration payment, agricultural policy, generalizability, laboratory experiment, social preferences

JEL CLASSIFICATION

C12, C91, D22, D91, Q12, Q20, Q57

1 | INTRODUCTION

The use of laboratory experiments to study issues in agricultural policy has grown in prominence within the fields of agricultural and environmental economics (Colen et al., 2016; Kesternich et al., 2017; Voors et al., 2016). Such experiments are often conducted with university students and framed in an abstract manner (e.g., Banerjee, 2017; Banerjee et al., 2012, 2014, 2017; Le Coent et al., 2014; Parkhurst et al., 2002; Warziniack et al., 2007). Yet, there is an ongoing methodological debate about whether the findings of such experiments provide reliable insights on the behaviour of actual agricultural policy addressees (Cason & Wu, 2017; Roe & Just, 2009; Thoyer & Préget, 2019; Viceisza, 2016).[1] This paper contributes to this debate by analysing the impacts of two aspects of experimental design for agricultural policy-relevant research: (i) sample population, that is, whether the experiment is conducted with students or 'real' policy actors, and (ii) framing, that is, whether the experimental decision situation is described in an abstract or context-specific manner.

The key advantage of the laboratory lies in the exogenous and controlled variation of policy interventions to study their causal impact on conservation outcomes. Student subjects are also easily accessible at low-cost as compared to farming professionals. Laboratory experiments thus provide a cost-effective alternative to randomised control trials (RCTs) often regarded as a gold standard in policy impact evaluations, because they are able to examine if a policy intervention could work in the real world (Behaghel et al., 2019; Ferraro, 2009),[2] and allow comparison of alternative policy designs. Especially, conducting experiments with students can constitute a useful first step towards formulating ex-ante predictions on how a population may react to a specific policy. In the context of agricultural policy, several studies have used laboratory experiments with university students to study specific informational and design aspects of agri-environmental policy. For example, some have examined the design of an agglomeration bonus as an incentive mechanism to promote the coordination of land use decisions and enhance the provision of ecosystem services at the landscape level. Usually, these experiments make use of a coordination game to explore how best to incentivise landowners to obtain an efficient coordination level. Parkhurst et al. (2002) and Warziniack et al. (2007) find that such a bonus helps mitigate coordination failure akin to conservation at the landscape level, and that the effect is reinforced by pre-play communication or 'cheap-talk' in groups. Banerjee et al. (2014, 2017) analyse the impact that the type of information available to subjects has on spatial coordination success. They find that providing information about the behaviour of 'neighbouring landowners' leads to improved coordination outcomes when introduced as a stand-alone policy. Yet, such information is found to be less effective in increasing coordination rates than a bonus policy without such information. Others have examined factors that limit the effectiveness of an agglomeration bonus mechanism, such as group size and transaction costs (Banerjee et al., 2012, 2017). Other studies used economic laboratory experiments with students to assess the effect of alternative payment for ecosystem services (PES) designs on collective action outcomes. Le Coent et al. (2014), using a public goods game framework, show that a PES improves the provision of a public good compared to a situation without such a subsidy. They also found that a collective-group PES scheme is more effective in promoting public good provision than an individual-user subsidy. Alpízar et al. (2017) study rules for optimal targeting of environmental services to decide who to include in a subsidy scheme to promote the provision of ecosystem services. They find that those included in a subsidy scheme increase their contributions to a public good, yet those excluded significantly decrease theirs, which can jeopardise the gains of the subsidy scheme. While providing important insights on the impact of different design aspects of conservation contracts, these experimental studies raise an important question: Do farming professionals who have been the main target population of conservation contracts react in the same way to alternative contract designs as conventional students? This question touches on the need for generalising behavioural

[1]Similarly, a growing number of papers in economics are questioning what can be learned from laboratory studies (e.g., Alm et al., 2015; Gneezy & Imas, 2017; Kessler & Vesterlund, 2015; Levitt & List, 2007; Schram, 2005).
[2]Some recent studies also highlight limitations of RCTs (see e.g., Deaton & Cartwright, 2018).

results from the laboratory to the real world, also referred to as 'external validity' of the experimental results.

There are two main concerns about external validity. First, results obtained from a sample of university students might not be generalisable to non-student populations (Omar & List, 2013; Snowberg & Leeat, 2021). For instance, students may have no experience with the problem at hand and their socio-economic and demographic background usually differ from the targeted population. We are only aware of three studies that compare the strategic behaviour or preferences of students and farmers using a laboratory experiment (Carpenter & Seki, 2011; Maart-Noelck & Musshoff, 2013; Waichman & Ness, 2012). These studies focus, however, on abstract designs, not directly related to concrete insights on agri-environmental policy design. However, non-student subject pools are increasingly used. With respect to conservation policies, several studies conduct field experiments with professional farmers or forest users to study the optimal design of PES aimed to limit deforestation in the developing world (e.g., Andersson et al., 2018; Gatiso et al., 2018; Handberg & Angelsen, 2015, 2019; Kaczan & Swallow, 2019; Loft et al., 2019; Reutemann et al., 2016; Salk et al., 2017). Few experimental studies also tackle the design of agri-environmental schemes with farming professionals in Europe and in the USA (e.g., Ferré et al., 2018; Palm-Forster et al., 2018; Peth et al., 2018; Suter & Vossler, 2013; Thomas et al., 2019). None of them compares results to the alternative of using students. Thus, we aim in this paper to investigate how the experimental results and resulting policy insights from a student sample compare to those from a sample of farming professionals. For this, we use an experiment developed to study alternative conservation policy designs for promoting sustainable use of peatlands in Switzerland, which we run with university students and with farm apprentices. Because it is difficult and costly to conduct experiments with actual farmers in developed countries, working with farm apprentices is a viable alternative and of policy interest as they have a farm background and represent the next generation of farmers.[3]

A second concern is that the abstract context in which decisions are made in the laboratory are not necessarily present in the real world, which also threatens the external validity of the results. Many experimental studies tackling agricultural and environmental policy questions follow the convention used in experimental economic research: context-free and neutrally framed instructions in order to retain experimental control (Durlauf & Blume, 2009). Despite this, experimental results are interpreted with respect to environment-related behaviour and policy insights including design issues for agri-environmental schemes (e.g., Banerjee, 2017; Banerjee et al., 2012, 2014, 2017; Le Coent et al., 2014; Parkhurst et al., 2002; Warziniack et al., 2007), conservation auctions (e.g., Cason & Gangadharan, 2004; Rolfe et al., 2009; Schilizzi & Latacz-Lohmann, 2007), and emission permit markets (e.g., Cason & Gangadharan, 2003, 2006, 2011). Only a few laboratory experiments in agricultural and environmental economics use context-loaded instructions where the experimental task is framed with the environmental problem at hand, such as forest and biodiversity conservation and climate change mitigation (e.g., Brick & Visser, 2015; Dörschner & Musshoff, 2015; Hasson et al., 2010; Milinski et al., 2006; Rommel et al., 2017; Tavoni et al., 2011). Despite discussions on the use of heavily contextualised instructions, systematic comparisons regarding policy insights are rare (Cason & Raymond, 2011). A second aim of this paper is therefore to provide such a comparison. We use a highly contextualised version of our experiment that provides specific case information with respect to the nature of the conservation problem and contextual visualisation within the experimental software, and we compare this to a context-free, abstract experiment.

[3]No such comparison exists for farm apprentices. We only know one study that makes use of professional students. Beck et al. (2014) use car mechanics apprentices to study whether professional car mechanics react the same way in a laboratory credence good market experiment as standard students. They find similar behaviours between the subject pools. Yet, car mechanics overtreat more, which is explained by their professional decision heuristics.

Our experiment was developed to study economic incentives promoting the sustainable management of organic (peat) soils in locations where these soils are used for high value crop production and therefore the opportunity cost of adopting sustainable practices is high. The case study region, Seeland in Western Switzerland (French speaking part), provides an example of such a setting, and we parameterised the experimental payoffs based on production data from this region (Ferré et al., 2018). Although draining of organic soils is necessary for conventional agricultural activities (here vegetable production), drained organic soils generate negative environmental externalities (greenhouse gas emissions). The experiment tests the effectiveness of an agglomeration payment to promote a peat-conserving alternative land use. Because rewetting requires spatial coordination among a group of farmers, the game has the structure of a coordination game. Thus, to study population effects we compare the effect of the policy treatment between farm apprentices and a standard university student sample. To study framing effects, we compare the contextualised and framed version of the experiment to an abstract and context-free one, with university students only. We also check if behaviour is correlated with a rich set of covariates, which we collect from our samples including socio-demographics, social and risk preferences as well as environmental attitudes.

In Section 2, we provide a review of the experimental literature on population and framing effects. Section 3 describes the experimental design. Section 4 presents and discusses the results and Section 5 concludes.

2 | POPULATION AND FRAMING EFFECTS IN ECONOMIC EXPERIMENTS: STATE OF THE ART

2.1 | Population effects

Although the majority of broader behavioural experimental economic studies are based on university students' samples, a growing literature analyses experimental results from other populations, including samples of large representative populations and samples of professionals or specialists (for a meta-study see Fréchette, 2015).[4] A robust finding from this literature is that in typical social preference games, students usually behave less prosocially than representative populations and professionals (e.g., Anderson et al., 2013; Bellemare & Kröger, 2007; Belot et al., 2015; Exadaktylos et al., 2013; Fehr & List, 2004). Falk et al. (2013) note that student samples likely provide a lower bound measurement of (social) preferences. In the realm of agricultural economics, Carpenter and Seki (2011) compare the social preferences of students to resource users (i.e., fishermen) via a public goods game. They also find that students contribute significantly less than real fishermen do. With respect to strategic behaviour, Waichman and Ness (2012) find no differences in a bargaining experiment between students and farmers. Maart-Noelck and Musshoff (2013) measure risk attitudes and find that the behaviour of students is more risk-averse than that of the German farmers. Table 1 provides a summary of studies that systematically assess population (and framing) effects in environmental and agricultural economics.

Various reasons have been suggested to explain such differences. First, the results may reflect actual differences in the general distribution of (social) preferences across populations (Carpenter & Seki, 2011; Falk et al., 2013). A second potential explanation is self-selection because participants of a laboratory experiment are volunteers, and the behaviour of non-volunteers cannot be

[4]Problems of external validity might arise because of the artificiality of the subject (as we test in our case) or the experimental setting, which may not reflect real-life conditions (Schram, 2005). In our experiment, the latter is not an issue since our experiment was developed to closely follow the decision-making situation of a vegetable farmer on organic soils in the Seeland region, and was parameterised accordingly.

TABLE 1 Studies on population and framing effects in agricultural and environmental economics.

Reference	Task/game	Measurement	Differences
Population			
Carpenter and Seki (2011)	Public good game	Cooperation	Yes
Waichman and Ness (2012)	Bargaining game	Strategic behaviour	No
Maart-Noelck and Musshoff (2013)	Holt-and-Laury lottery	Risk preferences	Yes
Framing			
Cason and Raymond (2011)	Emission market	Compliance behaviour	Yes
Pevnitskaya and Ryvkin (2013)	Public bad game	Production/Pollution	Yes
Bernold et al. (2015)	Public good game	Cooperation	No
Rommel et al. (2019)	Holt-and-Laury lottery	Risk preferences	Yes

assessed in such experiments. Eckel and Grossman (2000) find that volunteers act more prosocially in a dictator game than non-volunteers. However, several other studies do not find a difference in preference types between volunteers and non-volunteers (Anderson et al., 2013; Cleave et al., 2013; Falk et al., 2013). Third, and most relevant for this study, differences in the familiarity of the subject with the experiment is another possible underlying reason for the differences in experimental results. This may induce professionals to assume the presence of certain features of their work environment and behave accordingly, or some of their behaviour may only be triggered by specific signals being present in the experiment or not (Fréchette, 2015, 2016). This argument is in line with Cason and Wu (2017) who highlight that the suitability of students as the sample population depends on the objectives of the experiment. Field professionals (or 'real' decision-makers) are more appropriate in the context of measuring behaviour or preferences for a specific population, such as in Maart-Noelck and Musshoff (2013) measuring risk attitudes, or measuring the impact of a specific agricultural policy on a population, such as in Suter and Vossler (2013). They argue that when aiming to discover something about the preferences of the specific people who would be most affected by a policy intervention, it is preferable to use a professional sample.

2.2 | Framing effects

Abstract experiments combined with neutral instructions are most common in experimental economics. The main arguments that support the use of such format include the strengthening of internal validity and control of the experiment as well as easier replicability (Alekseev et al., 2017). This is based upon experimental evidence that framing can affect individual decisions and preferences (Elliott & Hayward, 1998; Elliott et al., 1998; Gächter et al., 2009; Homar & Cvelbar, 2020; Hossain & List, 2012; Tversky & Kahneman, 1981) as well as the interpretation of others' behaviour (Ellingsen et al., 2011).[5] Findings are mixed: some find that experimental subjects may associate a context with a certain kind of behaviour, i.e. greater self-interest in an economic context (Koneberg et al., 2010); others find no such context effects (Meier, 2006; Rege & Telle, 2004).

A few studies provide evidence on the effect of 'environmental' framing on experimental outcomes. Results are again mixed. Cason and Raymond (2011) study compliance behaviour in an emission market experiment and find unexpectedly that compliance rates are significantly lower, and participants under-reported pollution levels in the treatment that used contextual

[5]For an overview of various types of framing approaches used in the experimental economics literature, see Levin et al. (1998).

language with references to pollution and emission trading. Pevnitskaya and Ryvkin (2013) and Bernold et al. (2015) use a public bad and public good game, respectively, to study the impact of contextual, environmental framing. Although the former find that contextual information significantly decreases production and pollution in the expected direction when compared to a neutral treatment, the latter find no differences between a neutral and a contextually framed treatment on cooperation levels. Rommel et al. (2019) investigate the predictive power of differently framed lotteries measuring risk attitudes of farmers in a field experiment. They find that participants are more risk-seeking in an agricultural framing. To our knowledge, only one study addresses the impact of abstract versus contextualised framing in the context of the design of conservation incentives. Reddy et al. (2020) use a field experiment to investigate the role of emphasising the economic and environmental benefits of conservation in the framing of a conservation agriculture programme. They find no positive effect of this framing on conservation as compared to a neutral framing. Although contextualised experiments may induce less control over the experiment and are potentially less easily replicable, a context-free (neutral) experiment gives room for a diversity of interpretations to subjects. This effect, referred to as the 'construal' effect, can lead to difficulties in understanding behavioural outcomes of unframed experiments (Levitt & List, 2007; Paluck & Shafir, 2017). In this respect, the framing and contextualisation of experiments can act as a way to guide the player to a certain context and therefore narrow their spectrum of possible interpretations of the experimental set up. In other words, providing a frame and meaningful context to the experiment can exert control over how players approach the task in the experiment (Cronk & Wasielewski, 2008; Harrison & List, 2004). This would seem particularly relevant when experiments are conducted for the purpose of assessing context-specific policy outcomes or to guide policy design for specific contexts, as in our study.

2.3 | Our contribution

Our paper contributes to the literature on population and framing effects in several ways. First, we compare the extent to which the behaviour of students differs from that of farm apprentices in an experiment that is contextualised by an agricultural issue. In the literature, experimental studies rarely engage with a population close to farmer professionals. Second, in addition to comparing behaviour in the main experiment, we also conduct experiments measuring social preferences,[6] and examine the impacts of the population regarding the effect of these preferences and other attitudinal variables on behaviour. Thereby, we contribute to the evidence on potential differences in preferences and other characteristics between students and professionals, which has not been analysed for the case of farm apprentices. Moreover, we shed light on whether the degree to which such preferences or characteristics actually influence behaviour also differs between populations. Third, we aim to provide policy insights from the experimental results, and test whether the choice of sample population indeed matters for the policy-relevant results, as suggested by Cason and Wu (2017). We differentiate three different types of policy insights: the direction of the policy effect,[7] the size of the policy impact and the effect of behavioural

[6]There are an increasing number of studies that evidence the role of prosocial preferences in explaining individual decisions. In fact, Heinz and Koessler (2021) find that activating and underlining other-regarding preferences in the design of an intervention could promote pro-environmental behaviour.

[7]We do not expect the direction of the policy effect to differ with the type of framing and the type of population in this study. As the treatment deals with a monetary incentive, it will most likely induce a positive direction of the policy impact. However, while most studies on monetary incentives and effect on behaviour show a clear positive effect on behaviour (e.g., increased cooperation), a smaller part of the literature has shown that monetary incentives can have counter-productive effects and a negative impact on behaviour (referred to as crowding out). Crowding-out effects can be explained by a combination of various interlinked socio-psychological mechanisms that can motivate the individuals besides the financial incentive. For example, sensitiveness of the shifting of a frame (Kieninger et al., 2018) and feelings of reduction in autonomous decision-making induced by the payment (Akers & Yasué, 2019) are factors reported to potentially generate crowding out.

drivers on the policy impact. Fourth, we test the effect of framing for the behaviour of university students, introducing a high degree of contextualisation that includes asking subjects to take on the role of vegetable farmers in the Seeland region. Following the approach of Janssen et al. (2010, 2014), the visual interface of the experiment was highly contextualised to the Seeland region, and the experiment simulated the dynamic of the resource management problem as well as the asymmetry of the resource endowment among 'farmers'. Finally, because our experiment was designed to closely reflect the decision situation of farmers on organic soils in Switzerland, the game structure is somewhat more complex than that applied in previous studies on framing or populations effects. This complexity could potentially also induce differences in the importance of framing and of working with a sample population more familiar with this type of decision situation as compared to previous studies that have used more standardised game structures.[8]

3 | METHODS

3.1 | Experimental design

3.1.1 | General setup

We make use of a laboratory experiment originally conducted with farm apprentices by Ferré et al. (2018) and conduct additional sessions with university students to address subject pool and framing effects.[9] The experiment captures the following main components of the decision situation of farmers on organic soils in Switzerland. First, sustainable use of organic soils is considerably less profitable than the conventional management practice (vegetable farming) that relies on intensive drainage of these soils. Second, the adoption of sustainable land use requires farmers to coordinate, because several farmers depend on a joint drainage system. Specifically, sustainable land use requires rewetting the soils, which in turn requires unanimous agreement among neighbouring farmers. Third, farmers differ in their costs of conservation. The conventional land use reduces the peat soil layer and eventually makes farmers dependent on the underlying mineral soil layer. This underlying soil layer spatially differs in quality, implying differences in production potentials of the conventional activity as organic soils get depleted. These differences translate into differences in opportunity costs: a farmer with a high-quality underlying mineral soil layer has relatively little to lose from depleting the peat layer, implying a higher opportunity cost of switching to the sustainable use.

Our experiment captures the dynamics of the management problem over time inspired and calibrated by realistic production functions from the Seeland region.[10] Groups of two players were formed and consist of a High (H) and a Low (L) opportunity cost player. In the experiment, subjects were randomly assigned one player type (i.e., L or H type). The H player is characterised by having a high quality underlying mineral soil layer, meaning that depletion of the peat will not affect productivity of their land. The L player is characterised by having a low quality underlying mineral soil layer, meaning that depletion of the peat through vegetable farming will substantially affect productivity of their land. More precisely, the dynamics of the game unfold as follows. H's payoff from conventional land use is constant over time. L's payoff declines as they get closer to the underlying mineral soil layer, which is proxied in the experiment by the number of rounds spent in conventional land use. This means that players face different opportunity costs and there-

[8]We check for the general understanding of the experiment with the help of quiz questions. We find that across treatments and conditions, all subjects demonstrated a good understanding of the experiment.

[9]Ferré et al. (2018) examine the relative effectiveness of different policy incentives on sustainable land use in a contextualised experiment conducted with farm apprentices.

[10]Such a dynamics approach is similar to experimental environments that include spatial and temporal dynamics to study social-ecological systems. For an overview, see Janssen et al. (2008).

fore different incentives for adopting the sustainable land use over time. Moreover, it means that the opportunity costs of the L players are endogenous and dynamic: they depend on the player's choice in terms of the number of previous rounds spent in the conventional land use, which in turn may also be affected by the other player's cooperativeness in terms of rewetting the land.

3.1.2 | Decision procedure and payoff matrix

The decision procedure of the experiment is as follows. In every round, each player needs to decide between conventional and sustainable land use. The decision process is modelled as two stages. In stage 1, each group member votes in favour or against rewetting the soils (remember that rewetting is a necessary condition to sustainable land use). If at least one of the two players of a group rejects rewetting, the sustainable land use is not feasible and both players receive their payoff from the conventional land use by default in stage 2. If both players vote in favour of rewetting, players move to stage 2. In stage 2, players can either adopt the sustainable land use (which we refer to as 'cooperate') or instal a private drainage system and continue the conventional land use. Figure 1 shows the corresponding payoff matrices in stage 2, depending on the results of the vote in stage 1.[11]

Heterogeneous opportunity costs between group members create different incentives to cooperate in the experiment. Therefore, our experimental setup allowed two things to players within a group: (i) communication via a chat interface before making their decision; and (ii) making side

		Player H
		Conventional land use
Player L	Conventional land use	$\pi^L_{n,t}$ **π^H = 800**

(a) Payoffs at time t if at least one player votes against rewetting in stage 1

		Player H	
		Sustainable land use	Conventional land use
Player L	Sustainable land use	*$40 + UA - S^L + S^H$* **$40 + UA - S^L - S^H$**	*$40 + S^H$* **$\pi^H - 25 - S^H$**
	Conventional land use	*$\pi^L_{n,t} - 25 - S^L$* **$40 + S^L$**	*$\pi^L_{n,t} - 25$* **$\pi^H - 25$**

(b) Payoffs at time t if both players vote for rewetting in stage 1

FIGURE 1 Payoff matrices in stage 2. (a) Payoffs at time t if at least one player votes against rewetting in stage 1. (b) Payoffs at time t if both players vote for rewetting in stage 1. Payoffs of the H player are indicated in bold print, player L's payoffs in italics. Profit under sustainable land use is equal to 40. π^H and $\pi^L_{n,t}$ are defined as the payoffs of L and H players in conventional land use at time t given a number n of previous rounds used in conventional land use ($n = \{0, \ldots, 9\}$). $\pi^H = 800$ (constant) and $\pi^L_{n,t} = 800$ if $n = 0, 1, 2$; $\pi^L_{3,t} = 550$; $\pi^L_{4,t} = 160$; and $\pi^L_{n,t} = 0$ if $n = 5, \ldots, 9$ (that is, L's profit is negatively affected by conventional land use: it declines from the fourth round in conventional land use and falls to 0 from the sixth round in such land use, i.e., after five previous rounds in conventional land use [$n \geq 5$]). In the baseline treatment $UA = 0$. In the policy treatment, a uniform agglomeration payment treatment $UA = 770$ is allocated only if both players adopt the sustainable land use. S^i: side-payment offer made by player i.

[11]Payoffs were calibrated according to realistic production profits as obtained from the Seeland region: the manifold differences in profits between conventional and sustainable land use do reflect real world profits and values. More specifically, farm profits were calculated based on actual farm profit data as the present value of cumulated future farm gross margin over 5 years (with a discount rate of 4%), and rounded up (the unit of these numbers is kCHF) (Ferré et al., 2018).

payment offers (S^i) to each other.[12] Thus, ahead of stage 1, players can make a side payment offer to their group members. The offer is binding and conditional on other player's cooperation (i.e., choosing the sustainable land use).

The experiment includes a baseline phase (10 rounds) that captures the situation without a policy intervention, followed by a treatment phase (10 rounds) that simulates the introduction of an agri-environmental payment promoting sustainable land use. Players are rematched with a different player between the baseline and the treatment phase. This means that groups' composition changes between baseline and treatment. Besides, every experimental session starts with an incentivised[13] elicitation of social preferences, for which we use the Social Value Orientation (SVO) slider (Murphy et al., 2011); and ends with a short exit survey collecting socio-economic characteristics and environmental attitudes of subjects.[14] Experimental instructions, SVO test, and the survey are provided in the Appendix S1.

3.1.3 | Policy treatment

In the treatment phase, we introduce an agri-environmental payment motivated by the fact that the conventional activity implies an external cost to society (induced by greenhouse gas emissions from the peat layer depletion). We test a uniform agglomeration payment treatment (UA), which is conditional on both players adopting sustainable land use. The payment scheme is uniform, which means it pays an equal price for conservation to both players. The uniform payment is chosen such that total payoffs from cooperation always exceed the total payoffs from the conventional land use. The payment is set at UA = 770 and is constant over time.

The choice of this scheme is motivated by ongoing debate on the potential of agglomeration payments to motivate coordinated action of land users (e.g., Banerjee et al., 2014; Bell et al., 2016; Parkhurst & Shogren, 2007) and on whether payments should be differentiated by opportunity costs or be uniform across land users (e.g., Armsworth et al., 2012; Wünscher et al., 2008). The policy option considered raises interesting issues in the context of this paper. First, because the uniform payment treatment implies unequal economic incentives for cooperation, side payments are introduced in our experimental framework and may influence whether an agreement on mutual cooperation is reached. Second, farm apprentices may be more familiar than students with the agglomeration payment idea of rewarding spatial coordination (e.g., the 'Network Bonus' payment in Switzerland). Thus, it is interesting to compare the effect of an agglomeration payment between these two populations.

3.1.4 | Behavioural predictions

We provide a discussion of behavioural predictions for the baseline and treatment phase, starting with standard economic assumptions and then widening perspective to alternative assumptions. We developed the predictions by considering what players would do if they had both voted to rewet (Figure 1b). From there, we derived predictions on whether players would vote for rewet-

[12]Side payments are commonly assumed a viable option in the agglomeration payment literature (e.g., Bell et al., 2016) as farmers who aim to pursue conservation activities may use them to encourage participation of neighbouring farmers. In our setting, side payments are modelled as simple wealth transfer between players. We implement side payments via a numeric chat function on the interface of the experiment that enables players to make an offer to their group member. The offer is only transferred if the beneficiary adopts the sustainable land use option. In reality, side payments may not necessarily constitute monetary transfers but may also include risk sharing agreements and non-monetary benefits including reputation.

[13]Incentivised means that participants are paid according to their choices they make in the Social Value Orientation test. Please refer to the supporting material (1.1) of this paper for detailed information on how the SVO test was conducted.

[14]The experiment includes a quiz at the beginning of the session, after participants watched the instruction. It aimed at controlling for subjects' understanding of the experiment.

ting in the first place, or not (Figure 1a). A more detailed discussion of behavioural predictions can be found in Ferré et al. (2018). In the baseline phase (UA = 0), the payoff matrix in case of rewetting (Figure 1b) is such that payoff-maximising players would choose to instal the individual drainage system and revert to conventional land use. However, this is costly. Therefore, in the first rounds, both players are better off not voting to rewet (resulting in the payoff matrix in Figure 1a), and thus maintaining the conventional land use. Note, however, that payoffs for both players are highest without rewetting only as long as $n < 5$. Once the L player has adopted conventional land use for five rounds or more ($n \geq 5$), their payoff under the sustainable land use would be higher than under the conventional land use. Thus, in later rounds, the L player would prefer to rewet, whereas the H player would still prefer not to rewet. L may thus offer side payments to H to promote H's cooperation and rewet.

The introduction of the uniform agglomeration payment (UA = 1) transforms the payoff matrix in Figure 1b into that of a coordination game. That is, there are two equilibria for payoff-maximising players: one where both cooperate (i.e., both adopt the sustainable land use) and another where both do not cooperate (i.e., both revert to conventional land use). Yet, both players are always better off under the sustainable land use as the uniform payment exceeds individual opportunity costs.[15] Thus, standard economic predictions based on payoff-maximising players would be that the payment treatment leads to the equilibrium where both players cooperate. However, the treatment involves a risk of coordination failure because the agglomeration payment is conditional on both players' cooperation. This means that risk-averse players might prefer to not cooperate rather than to risk coordination failure. These differences may induce differences in the actual impacts of the treatment depending on the subject pool as well as differences in the use of side payments to get to this equilibrium. Additionally, because rewetting is conditional on both players agreeing to rewet, there may be bargaining between group members over sharing the difference between the payment and the actual opportunity cost of the player, using side payments. This may delay cooperation. Indeed, since L player's farm profit diminishes in the conventional land use, this difference is increasing with number of periods in which L conducts conventional land use. H may exploit this fact to obtain side payments from L. Thus, it is clear that, despite a strong incentive for coordination, there are risks of either delaying or failing to coordinate. Comparing students and farm apprentices in such a context is therefore ideal to reveal the potential effect of specific features of these populations on individual and group decisions.

It is widely known that behaviour in such experiments does not coincide with standard predictions based on the assumption of payoff-maximisation. Players exhibit varying degrees of social preferences (Fehr & Fischbacher, 2003; Vollan & Ostrom, 2010). In our baseline scenario, H players have no economic incentive to cooperate, while L players do after some rounds. In the treatment scenario, because of the risk involved by the agglomeration payment and the fact that their payoff from the conventional land use does not diminish over time, H players still have a weaker incentive than L players to cooperate. It appears therefore intuitive that prosocial preferences may play a role for H's decision to cooperate. Our hypothesis is further backed by literature that highlights non-pecuniary reasons, including concerns for environmental improvement and social norms, as important motivators for farmers to participate in agri-environment schemes (e.g., Botazzi et al., 2018; Howley et al., 2015; Sheeder & Lynne, 2011).

3.2 | Contextualised framing and abstract framing

The unframed version of the experiment was presented in abstract and neutral terms. For example, a player's choices were labelled with 'Activity A' and 'Activity B' and there was no mention of an agri-

[15]In the absence of side payments, initial opportunity costs are 800–40 = 760, while the payment is 770 per player. Opportunity costs of the L player decrease as the number of rounds in conventional land use increase. Thus, the payment always exceeds opportunity costs.

cultural context. In the framed (contextualised) version of the experiment, the visual background of the web-based experiment provides a picture of cultivated organic soils within an agricultural landscape that was inspired by the case study region. It visually reflects the impact of players' decisions on their land. Second, it uses heavily loaded instructions, which were shared on paper and through an introductory video with references to organic soils and cultivation of the land. Third, instead of referring to players, participants are labelled as farmers and players' actions referred to as land use options. See the Appendix S1 for an example of the baseline instructions in an abstract and contextualised version (Section 1.2 and 1.3) as well as screenshots of the experimental software (Section 1.4).

3.3 | Sample populations and organisation

The contextualised version of the experiment was conducted with a generic sample of university students from the Swiss Federal Institute of Technology (ETHZ) and University of Zurich, and with a sample of farm apprentices recruited from a set of agriculture schools in Western Switzerland. University students at ETHZ were recruited and invited to participate in an experiment at the ETHZ Decision Science Laboratory, while we made use of a conventional PC lab in the agricultural schools. Note that farm apprentices are not only taking a university degree specialised in agriculture but also undergo substantial on-farm practical training. Furthermore, the vast majority of them have a farming family background and plan to work as farmers once finishing their studies. We do acknowledge that they are not professional farmers (yet), and have therefore not as much practical experience as real farmers would have. On the other hand, farm apprentices are also an interesting sample, reflecting more closely the preferences of the next generation of farmers, which may be directly affected by soil degradation.[16] The abstract version of the experiment was conducted with university students only.

We test the effect of framing only with a university student sample and not with farm apprentices. This is because most economic experiments in the literature are conducted with university students and with limited or no contextualisation. Thus, our main interest is to test the impact of contextualisation on the behaviour of university students. Experiments conducted with professionals and that are motivated by providing policy insights are usually contextualised to some degree. This also makes sense because the potential confounding of decision-making due to unobserved contextual cues is not a concern and actually rather desirable when conducting the experiment with policy addressees. Testing the impacts of abstract framing on professionals would be more interesting for theory testing, which is not our objective here. Moreover, it is generally difficult to motivate professionals to participate in very abstract experiments. In total, we obtained decisions of 80 university students in a framed version and 76 students in an unframed version of the experiment and 88 (respectively 78 in the baseline) of farm apprentices in the framed version. With university students, we conducted four sessions each in the contextualised and abstract framing.[17] Table 2 provides an overview of the experimental sessions across framings and populations.

To provide subjects with an incentive to take all decisions in the experiment seriously, we employed the following procedure. After all decisions had been made, the computer randomly determined one round from the baseline and one round from the treatment phase to be payoff relevant. In addition, one decision from the social value orientation test (which consists of multi-

[16]We initially planned to conduct our experiments with professional farmers. The very same schools involved in these experiments regularly organise thematic training days for professional farmers. However, because of substantial time constraints and opportunity costs of professional farmers, we were unable to set up experimental workshops that would have satisfied our sample size requirements.

[17]With respect to sample size, we acknowledge that we first conducted the sessions with farm apprentices in the framed version of the experiment for Ferré et al. (2018). We find on average that 60% of subjects choose the sustainable land use option in the policy treatment. Conducting a simple sample size calculation using a test of two independent proportions assuming an expected proportion of sustainable land use of 60% in the farm apprentices' sample and assuming that the use of generic university student sample increase/decrease the probability of choosing the sustainable land use option by 20%. This results in a required sample size 86 subjects per condition. Note that we did not have expectations about the size of the treatment effect as there is no other similar study on policy treatment effectiveness across different sample populations.

TABLE 2 Experiments[a].

	Framed	Unframed
Farm apprentices	Baseline (N = 78) + Policy (N = 88)	-
University students	Baseline (N = 80) + Policy (N = 80)	Baseline (N = 76) + Policy (N = 76)

Note: For the sample of university students, each subject in each session participates in a baseline and policy phase while subjects were rematched into new groups after the baseline phase. For the sample of farm apprentices, we also conducted the baseline as a separate experimental session. This means subjects in this session played the baseline twice (baseline + baseline) (N = 78) and compare this to a session in which apprentices played the baseline + the policy phase. Doing this was necessary to ensure that players' decisions were not affected by the complexity of the game. However, we found that behaviour did not differ significantly between the first 10 rounds of the baseline and the second 10 rounds of the baseline. For this reason, we also did not implement this procedure for the university students.

[a]Experiments were programmed on the computer and conducted either at the Decision Science Laboratory of ETHZ or at the respective agricultural schools in Switzerland. With farm apprentices, the recruitment procedure relied on in-class advertising and flyers on campus. It made use of conventional PC labs equipped with additional blinders to create a more anonymous decision environment.

ple mini-dictator games) was selected to be implemented and subjects were paid accordingly. An experimental session lasted about 90 minutes and subjects earned CHF 40 on average (including a CHF 10 show-up fee).

4 | RESULTS

Besides examining behaviour in the experiment, we are interested in two aspects of policy insights: direction of impact (compared to baseline) and magnitude of the impact (size of coefficient). We first consider the impact of sample population on these three outcomes (Section 4.1), and then consider the effects of framing on the same three outcomes (Section 4.2). To assess policy impacts, we focus on the outcomes of the experiment in terms of individual rates of sustainable land use. A higher rate translates into a better environmental policy outcome as more soil is being preserved. In Section 4.3, we check whether the policy effect differs after controlling for social preferences and a rich set of socio-demographic characteristics of participants.

4.1 | Sample populations effects

First, we study if sustainable land use levels differ between students and farm apprentices.[18] Thus, we compare outcomes of the two sample populations for the baseline and the policy treatment (UA), both under contextualised framing. Figure 2 shows rates of sustainable land use in percentages across all rounds. This percentage is calculated, for each round, based on the number of subjects choosing sustainable land use in the given round and the sample size (indicated in Table 2). We find that this rate differs significantly among students and apprentices, both in the baseline and in the policy treatment.[19] On average across the rounds, students choose sustainable

[18]In terminology of a typical social dilemma, we may also refer to sustainable land use as 'cooperation'. We acknowledge, however, that the baseline is calibrated as such that sustainable land use choices do not increase social welfare, and therefore in a strict sense might not be referred to as cooperative behaviour.

[19]Assuming profit-maximising players, and given there are joint gains from coordinating, the proportion of farm apprentices who adopt a sustainable land use in the presence of UA is lower than one would expect. There are several potential explanations for this finding. First, it is largely recognised in behavioural economics that players do not act purely based on payoff-maximisation: some players have non-standard motivations, and characteristics like trust, relation to risk, and time may be relevant for their behaviour. Second, farm apprentices are likely to have a higher preference for conventional land use than for the alternative land use, despite the economic incentive, due to a status-quo bias. This behaviour could be reinforced by the very high opportunity costs that characterise the case, which may necessitate even stronger incentives for players to change activity, especially under the agglomeration payment treatments, which involve a risk of coordination failure. Third, highlighting inequalities among farmers leads to bargaining issues. Thus, some farmers may tend to 'abuse' the difficult situation of others with short-term production potential by conditioning their cooperation on a side payment from their partner. Moreover, group members need time to coordinate on the equilibrium, which usually takes a few rounds.

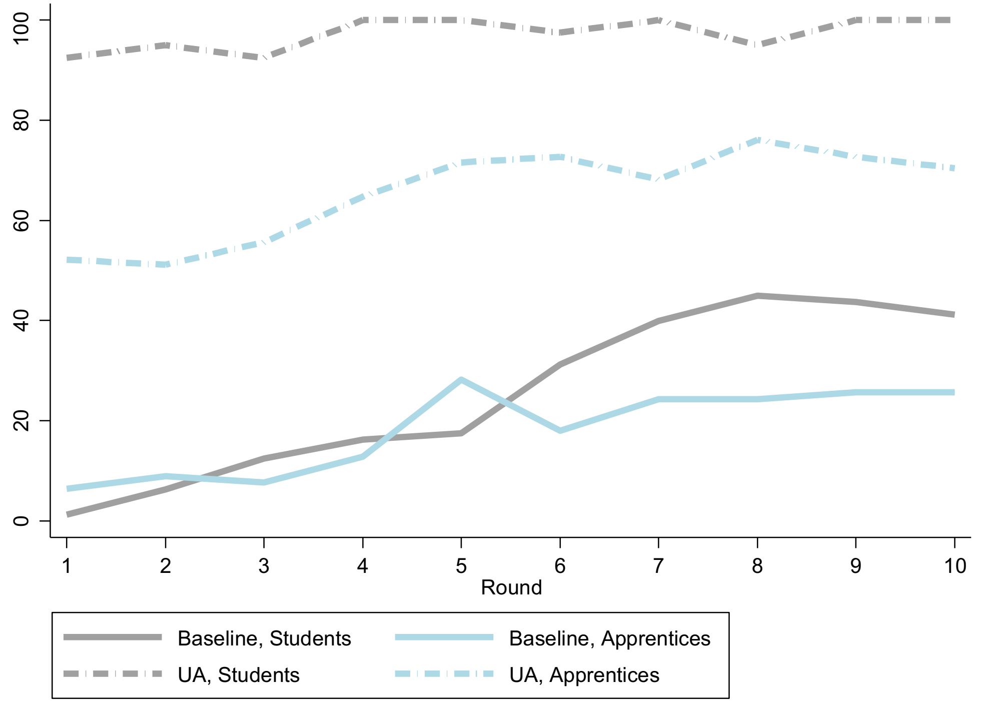

FIGURE 2 Rate of sustainable land use choices across populations in % (contextualised experiment). [Colour figure can be viewed at wileyonlinelibrary.com]

land use significantly more than farm apprentices. In the baseline phase, average choices across all rounds (calculated as the average percentage over the 10 rounds) amounts to 25.5% among students and to 18.2% in the farm apprentice sample (proportion test, p-value = 0.0004). Considering individual rounds, the average choice rate of sustainable land use differs significantly across the two populations from round 6 onwards.[20] Remember, in the dynamic context of our game, L's profit starts to decline from round 6 onward. In the treatment phase, sustainable choices are again significantly higher among students than among farm apprentices (97.3% and 65% respectively; proportion test, p-value = 0.00) and this holds for all rounds.

We proceed to analyse policy insights. With respect to the direction of the policy impact, the policy treatment leads to a significant increase in sustainable land use rates for both subject populations. The difference in 'sustainable' choices between the policy treatment and the baseline is 47.4% among farm apprentices and 71.75% among students, both of which are statistically significant (proportion tests: p-values < 0.001). This is not surprising because the policy treatment transforms the game structure into a coordination game. However, the magnitude of this effect differs significantly across sample populations (p-value < 0.05), with the policy having a stronger impact on students' behaviour than on farm apprentices.

Summarising, we have:

> Result 1: Treatment-dependent cooperation rates differ significantly across populations.
> Result 2: The population does not affect the direction but significantly affects the magnitude of the policy impact.

[20]Between rounds 6 and 10, sustainable choices amount to 40.25% for students and 23.6% for farm apprentices (p-value < 0.0005).

4.2 | Framing effect

We examine the impact of framing on sustainable land use rates among the two subject pools. Figure 3 illustrates sustainable choices across rounds. In the baseline, the average rate of sustainable land use is 25.5% in the contextualised design and 25.9% in the abstract design (proportion test, *p*-value = 0.88). Under the policy treatment, it is equal to 97.3% in the contextualised and 94.0% in the abstract design ($p = 0.31$). Thus, we find no significant differences of framing on average sustainable land use levels. We furthermore observe that the rate of sustainable land use choices follows a similar pattern over the course of the game regardless of the framing, and this holds in both the baseline and the policy treatment.

With respect to the direction of the policy impact, the policy treatment leads to a significant increase in sustainable land use rates compared to the baseline, under both types of framing. The difference in choices between the policy treatment and baseline is statistically different from zero and positive for both framings (71.75% under the contextualised and 68.05% under the abstract design; proportion tests: *p*-values <0.05). Moreover, the framing does not affect the magnitude of the policy impact: the difference in increase of cooperation (3.7%) is not statistically significant (proportion test; *p*-value = 0.68).

In summary, we have:

> Result 3: Framing does not affect treatment-dependent cooperation rates among university students.
> Result 4: Framing does not affect the direction nor the magnitude of the policy impact.

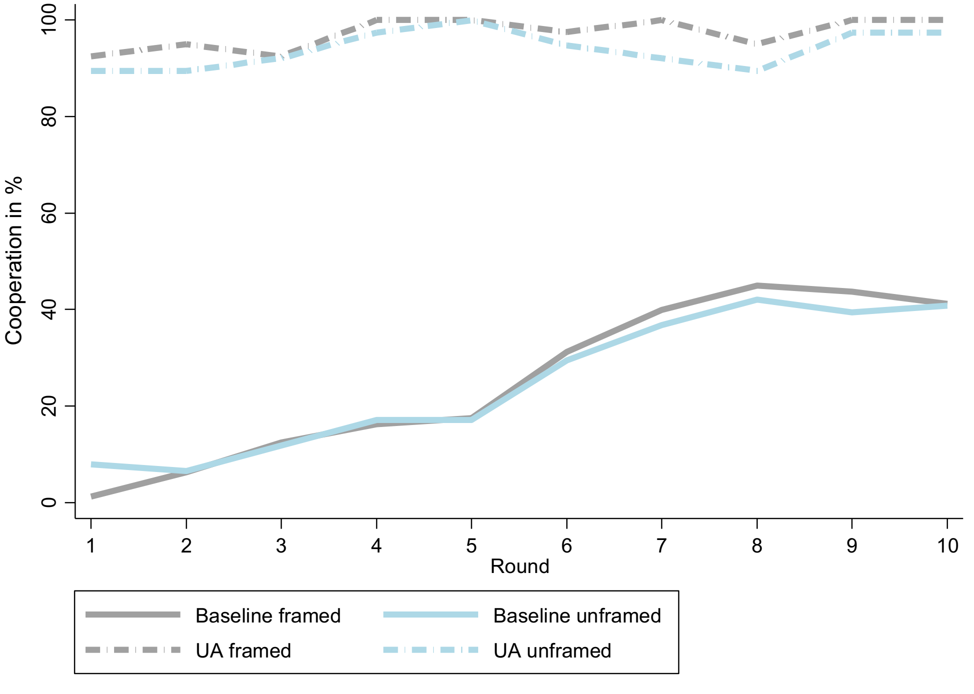

FIGURE 3 Rate of 'sustainable' land use choices in the contextualised framing and abstract framing (university students only) in %. [Colour figure can be viewed at wileyonlinelibrary.com]

4.3 | Effect of individual characteristics across populations and framings

We examine the impact of socio-demographic characteristics and individual preferences on behaviour in the experiment and analyse whether these impacts differ with the sample population or with the framing. There are two issues here. First, different sample populations may differ in terms of some characteristics or preferences that affect behaviour. For example, if risk or social preferences matter for behaviour in our experimental setup, a difference in such preferences between university and apprentices could play a role in explaining difference in behaviour between the populations. Second, the way in which characteristics and preferences affect behaviour may also differ when population or framing change: for example, environmental preferences may affect behaviour more in the contextualised than in the abstract setting.

Using our exit survey (see details in the Appendix S1), we surveyed all participants on a large number of characteristics, including standard demographics such as age and gender, a measure for environmental preferences (an index ranging from 0% to 100%; a higher score indicates a stronger concern for environmental issues such as climate change); risk preferences (a score ranging from 0 to 10; a higher score indicates greater willingness to take risks); time preferences (a score ranging from 0 to 10; a higher score indicates more patience); and reputation concerns (a score ranging from 0 to 10; a higher score indicates more concern for others' opinions). Moreover, the incentivised implementation of the SVO slider measure (Murphy et al., 2011) provides an individual and continuous measure of social preferences, that is, higher scores indicate stronger social preferences.[21] Table 3 summarises the main characteristics across samples.

We find that our sample of farm apprentices contained a significantly lower percentage of women and was somewhat younger than the sample of university students. In terms of environmental, risk and time preferences as well as concern for reputation, farm apprentices report to be, on average, less concerned with environmental issues, more risk loving and less patient, and care more about reputation than generic university students. Considering social preferences captured by the SVO angle, we find that the farm apprentice population is characterised by higher prosocial preferences than the student population (Mann–Whitney test, p-value = 0.00). In addition, the distribution of social preferences between the university student sample and farm apprentices differ significantly (Kolmogorov–Smirnov test, p-value = 0.00). For a detailed graphical distribution of the SVO score across population, see the Appendix S1, Section 3. Furthermore, the SVO angle allows categorising players into two distinct social preference types namely proself or individualistic (i.e., those with an angle between −12.04° and 22.45°) and prosocial (i.e., those with an angle ranging between 22.45° and 57.15°) (Murphy et al., 2011). We find that the percentage of prosocial subjects is significantly higher among the farm apprentices (64.5%) than among the student sample (52.6%) (chi^2 test: p-value = 0.03). These findings are in line with other studies cited in Section 2.1 that also found that students tend to be less prosocial than professionals.

We investigate whether the differences in preferences relate to differences in gender composition between the populations (for an overview on gender differences in preferences in the experimental economic literature, see Croson & Gneezy, 2009; Charness & Gneezy, 2012). Table S4 in the Appendix S1 shows the main individual characteristics and preferences split by gender. Considering males versus females separately generates the same findings as with the whole population, except for the environmental preference variable where we find that male students do not have a significantly higher concern for the environment than male apprentices (t-test, p-value = 0.20). Thus, the unequal representation of gender in the two populations does not seem to be a key explanatory factor to players' difference in preferences, and subsequently in behaviour

[21] The SVO slider measure elicits the extent to which individuals place importance on their own payoff in relation to the payoff of some other individual by asking subjects to make 15 distributional payoff decisions between oneself and another person. The test results in an individual measure of prosociality and yields an angle ranging between 16.26° and 61.39°. The higher the individual's SVO angle, the stronger their social preference.

TABLE 3 Socio-demographic and preference characteristics across populations.

Characteristics	Farm apprentices (*N* = 166)	Students (*N* = 156)	Test for difference in mean or %
Female (%)	16.1	54	***
Age (years)	19.7 (3.4)	21.3 (2.2)	***
Farm household (Yes, %)	50.6	-	
Plan to work in farm business (Yes, %)	83.1	-	
Environmental preference (0 = no concern; 100 = strong concern for the environment)	56.2 (15.3)	62.7 (13.2)	***
Risk preference (0 = avoid, 10 = willing to take risks)	6.2 (2.3)	4.8 (2.3)	***
Time preference (0 = very impatient, 10 = very patient)	4.8 (2.8)	5.7 (2.3)	**
Reputation (1 = do not care at all, 4 = care a lot)	1.8 (0.8)	1.6 (0.7)	**
SVO angle	25.1 (14.9)	22.3 (13.4)	*

Note: ***$p < 0.001$, **$p < 0.05$; *$p < 0.10$.

in the experiment. This suggests that the agricultural background might be a more important underlying reason for the differences between sample populations.

Next, we conduct a random-effect logistic panel regression on players' willingness to cooperate and test separately for interaction effects with respect to population use and framing of the experiment (Tables 4 and 5).[22] Because our focus is on population and framing effects, we do not discuss the impacts of individual explanatory variables in detail, unless there are significant effects of variables, which we found above to differ between sample populations, or unless the impact of the variables differs by population or framing (interaction terms).

We first consider the effect of the sample population in the baseline without policy (Table 4, first results column). Our results indicate that cooperation in the baseline is more likely when players are female, younger and more risk averse (odds ratio for risk-loving variable <1). Recall that the student sample was found to be more risk averse and with a higher proportion of females as compared to farm apprentices. This may therefore contribute to an explanation of why students cooperated more. Regarding the interaction effects between population and explanatory variables, we do not find any significant effect in the baseline. Thus, players' characteristics seem to impact behaviour in the same way for both populations.

Next, consider the policy treatment (UA) (Table 4, second results column). We find that cooperation in this treatment is significantly affected by players' age and environmental preferences. Cooperation is more likely when players are older and more concerned about the environment. Recall that apprentices were found to be less concerned about the environment and younger. Thus, this may be part of the explanation why they cooperate less in the policy treatment. Moreover, for the policy treatment, we find a significant interaction effect between players' concern for the environment and population. First, environmental concern seems to influence the behaviour of students more strongly than the behaviour of apprentices.[23] Recall that in this

[22]We choose to apply a random-effect logistic panel regression model because: (i) the dependent variable is a dummy variable (non-cooperation = 0, cooperation = 1), hence the 'logistic' regression; (ii) we account for the time (the round), hence the panel regression; and (iii) social characteristics vary across observations, hence the random effect.

[23]The coefficient on environmental concern alone refers to the effect for farm apprentices (reference level in the regression). The total effect of environmental concern on students is calculated by multiplying the two coefficients, that is, $1.032 \times 1.077 = 1.11$.

TABLE 4 Panel regression on player's willingness to choose 'sustainable' land use and population effect.

Variables	'Sustainable' land use baseline	'Sustainable' land use policy treatment
Round	1.407***	1.215***
	(0.0692)	(0.0784)
H Player	0.0588***	0.579**
	(0.0176)	(0.136)
Population (0 = farm apprentices, 1 = students)	0.00593*	0.404
	(0.0180)	(2.194)
Age	0.813*	1.086*
	(0.0921)	(0.0535)
Female	2.790*	0.756
	(1.562)	(0.424)
Environment concern	0.996	1.032*
	(0.0148)	(0.0172)
Risk loving	0.820*	0.959
	(0.0893)	(0.130)
Patience	1.010	1.057
	(0.0860)	(0.118)
Reputation	0.691	0.972
	(0.199)	(0.344)
SVO angle	1.003	0.987
	(0.0194)	(0.0170)
Farm household	0.534	0.955
	(0.342)	(0.547)
Farm business	0.932	0.839
	(0.579)	(1.117)
Age × Population	1.185	1.036
	(0.154)	(0.290)
Female × Population	1.872	0.204*
	(1.143)	(0.192)
Environment × Population	0.993	1.077**
	(0.0183)	(0.0391)
RISK Loving × Population	1.137	1.765***
	(0.138)	(0.361)
Patience × Population	1.084	0.837
	(0.112)	(0.183)
Reputation × Population	1.376	0.313*
	(0.493)	(0.190)
SVO Angle × Population	1.005	1.045
	(0.0219)	(0.0378)
Constant	22.01	0.0830
	(54.00)	(0.212)

(Continues)

TABLE 4 (Continued)

Variables	'Sustainable' land use baseline	'Sustainable' land use policy treatment
Observations	1560	1670
Number of groups	156	167

Note: We report odds ratios: ***$p<0.01$; **$p<0.05$; *$p<0.1$. To control for group membership, all regression models also include group dummies (not reported). Standard errors clustered at the group level are shown in parentheses.

comparison, all experiments were contextualised. Therefore, the strong impact of environmental concern on students' behaviour may be due to the environmental framing. Second, we find a significant interaction effect between players' willingness to take risk and population. Specifically, risk preferences affect students' behaviour more strongly than farm apprentices. This may be explained by the higher level of familiarity of farm apprentices with making land use decisions that involve risk. Third, we find a significant interaction effect between players' reputation and population. The care for reputation affects the decision of farm apprentices more than the decision of students. Reputation has been recognised as a potentially important determinant of decision-making in agriculture (Sattler & Nagel, 2010; Sereke et al., 2016).

In summary, we find:

> Result 4: The sample population affects behaviour of players in the contextualised experiment. In both baseline and policy treatment, differences between sample populations in socio-economic characteristics and preferences can help explain differences in behaviour. Moreover, under the policy treatment, cooperation of farm apprentices is affected more strongly by reputation concerns and less strongly by environmental and risk preferences than is the behaviour of students.

Next, we analyse the effect of framing and the interaction effects between framing and the characteristics of players (Table 5). Note again that here we only look at experiments conducted with university students. We find that framing has no overall effect on behaviour, which confirms our earlier results. Moreover, interaction terms between framing and the other explanatory variables are not significant in the baseline. In the policy treatment, we find that the impact of environmental concern and willingness to take risk is stronger in the contextualised than in the abstract experimental design. One potential explanation is that the contextualised design, by more closely capturing the reality of farm management and depletion of peat, is more likely to trigger such environmental and risk concerns in the decision-making process. In summary, we have:

> Result 5: Framing has no overall impact on cooperation. However, environmental and risk preferences have a somewhat stronger effect on behaviour in the contextualised than in the abstract experiment.

Placeholder TextPlaceholder Text[24]

Placeholder Text

5 | CONCLUSION

We examined the impact of population and framing effects in economic experiments designed to test the design of agri-environmental schemes. Our comparative analysis reveals that neither the type of subject nor the framing affects the direction of the policy impacts. However, we find

[24]This corresponds to 0.000000917 (0.00000785).

TABLE 5 Panel regression on player's willingness to choose 'sustainable' land use and framing effects.

Variables	'Sustainable' land use baseline	'Sustainable' land use policy treatment
Round	1.575***	1.249
	(0.0896)	(0.199)
H Player	0.0241***	1.482
	(0.00969)	(0.664)
Framing	0.345	9.17 e-07[a]
(0 = abstract;1 = contextualised)	(1.116)	(7.85 e-06)
Age	0.906	1.060
	(0.0607)	(0.171)
Female	1.064	1.609
	(0.446)	(2.042)
Environment concern	0.989	0.994
	(0.0191)	(0.0308)
Risk loving	0.781*	0.792
	(0.116)	(0.170)
Patience	1.181	0.783
	(0.142)	(0.169)
Reputation	1.131	0.408
	(0.377)	(0.366)
SVO angle	1.006	0.970
	(0.0144)	(0.0420)
Age × Framing	1.047	1.122
	(0.100)	(0.448)
Female × Framing	0.599	0.365
	(0.296)	(0.593)
Environment × Framing	0.999	1.151***
	(0.0224)	(0.0589)
Risk loving × Framing	1.203	2.250***
	(0.185)	(0.650)
Patience × Framing	0.914	1.127
	(0.120)	(0.380)
Reputation × Framing	0.872	0.456
	(0.356)	(0.514)
SVO angle × Framing	1.006	1.067
	(0.0181)	(0.0612)
Constant	0.945	3.265
	(2.465)	(20,230)
Observations	1550	1550
Number of groups	155	155

Note: We report odds ratios: ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$. The regressions also included group dummies (not 631 reported).[a]This corresponds to 0.000000917 (0.00000785).

a difference in the level of the policy impact when we compare university students with farm apprentices. This suggests that quantitative estimates of policy effects could be misleading when based on student samples only.

More specifically, we find that the rate of sustainable land use is significantly larger among students than farm apprentices, in both absence and presence of a payment policy. While we find—in line with other studies—that students show less prosocial preferences than non-students, this difference does not seem to drive the difference in behaviour in our experiment. Rather, differences in sample populations with respect to environmental and risk preferences, and reputation concerns seem to be reasons for the difference in 'sustainable' land use rates. Moreover, for the agri-environmental policy treatment, student behaviour is found to be more strongly driven by environmental and risk preferences, whereas farm apprentices' behaviour is more strongly affected by reputation concerns. The result on environmental preferences may be due to the contextualisation of the experiment in terms of sustainable land use. A similar difference in the impact of environmental and risk preferences on behaviour is also found between framings, with the impacts being stronger in the contextualised design.

Regarding framing, we found it to have no overall effect on our results, neither on 'sustainable' land use rates nor on policy directions or magnitudes of policy impact. One potential explanation may relate to the fact that the experiment was already highly 'loaded' through making the structure of decision-making very similar to the actual resource problem, so that framing could not have 'improved' the understanding of the problem by the player, and thus the results. It should be noted, however, that we tested the impact of framing only among university students, and the result may be different for farm apprentices. Moreover, one added-value of contextualisation resides in the increase in credibility of related policy recommendations to policy-makers. Thus, the experimental setting tested in this study, which closely captures the resource management problem and the actual soil degradation dynamics, may well be seen as preferable by policy-makers.

Our results suggest that the importance of sample population and possibly framing depends on the type of policy insight sought. Abstract laboratory experiments with students appear reasonable for providing an indication of the broad impact of a policy. However, if the interest is in magnitude of policy effects or drivers of behaviour, a careful choice of sample population and possibly framing appears more important. Our study also demonstrates that the use of farm apprentices can be valuable for providing agri-environmental policy insights. Accessing such a population and implementing the experiment with them is relatively easy: agricultural schools often have a computer room, which can be set up similar to a university's laboratory, and farm apprentices have similar hourly wages as students and can therefore be paid in exchange for their participation using the same payoff rate. Yet, from our experience, conducting experiments with farm apprentices is easier if the experiment is implemented as part of the usual training hours, and if it includes an educational dimension and a connection to their farming background.

One of the limitations of our study is that we could not test the impact of framing on farm apprentices. This is partly due to logistical constraints in increasing the sample size of farm apprentices and partly to the difficulty of motivating farm apprentices to participate in a rather abstract experiment. Nevertheless, future research should aim at conducting such a comparison of contextualised and abstract experiments with farm apprentices or other professionals. Since professionals would relate more to the issue at hand, framing may well have a stronger impact on their behaviour. A second limitation relates to the fact that only one policy treatment is being tested in this study. Testing and comparing the relative outcomes of different policy treatment across populations and across framings could help to shed light on the relative preferability of alternative policy options, and provide indication on whether using student samples is an appropriate practice to provide a ranking of different policy options. Finally, although it seems reasonable that farm apprentices think and behave similarly to farmers, we acknowledge the fact that farm apprentices are not experienced farmers. For example, apprentices are younger than average farmers and may be more open to innovation. Because it is often easier to conduct experiments with farm apprentices, future research should compare the behaviour of farm apprentices to the one of farmers. This would likely provide further insights on behaviour and on the potential uptake of the policy.

ACKNOWLEDGEMENTS

We acknowledge the support from the Swiss National Science Foundation in the frame of the project 'Sustainable management of organic soils in Switzerland' (project number 406840-143145) as part of a National Research Program (NRP 68), and the Alexander von Humboldt Foundation for complementary funding in the framework of the Alexander von Humboldt Professorship endowed by the German Federal Ministry of Education and Research. We are very grateful to the Swiss agricultural schools that participated in this study, and in particular to the professors for their collaboration, and the students for their time to participate. We thank Orencio Robaina Martinez de Salinas for programming the computer application supporting the experiment, the Descil ETH laboratory for enabling the organisation of the experiments. We also want to thank Adrian Muller for support and feedback in this project, and anonymous reviewers for their constructive comments on an earlier draft of this paper.

REFERENCES

Akers, J.F. & Yasué, M. (2019) Motivational crowding in payments for ecosystem service schemes: a global systematic review. *Conservation and Society*, 17(4), 377–389.

Alekseev, A., Charness, G. & Gneezy, U. (2017) Experimental methods: when and why contextual instructions are important. *Journal of Economic Behavior and Organization*, 134, 48–59.

Alm, J., Bloomquist, K.M. & McKee, M. (2015) On the external validity of laboratory tax compliance experiments. *Economic Inquiry*, 53(2), 1170–1186.

Alpízar, F., Nordén, A., Pfaff, A. & Robalino, J. (2017) Unintended effects of targeting an environmental rebate. *Environmental and Resource Economics*, 67(1), 181–202.

Anderson, J., Burks, S.V., Carpenter, J., Götte, L., Maurer, M., Nosenzo, D. et al. (2013) Self-selection and variations in the laboratory measurement of other-regarding preferences across subject pools: evidence from one college student and two adult samples. *Experimental Economics*, 16, 170–189.

Andersson, K.P., Cook, N.J., Grillos, T., Lopez, M.C., Salk, C.F., Wright, G.D. et al. (2018) Experimental evidence on payments for forest commons conservation. *Nature Sustainability*, 1(3), 128–135.

Armsworth, P.R., Acs, S., Dallimer, M., Gaston, K.J., Hanley, N. & Wilson, P. (2012) The cost of policy simplification in conservation incentive programs. *Ecology Letters*, 15, 406–414.

Banerjee, S. (2017) Improving spatial coordination rates under the agglomeration bonus scheme: a laboratory experiment with a pecuniary and a non-pecuniary mechanism (NUDGE). *American Journal of Agricultural Economics*, 100(1), 172–197.

Banerjee, S., Cason, T.N., de Vries, F.P. & Hanley, N. (2017) Transaction costs, communication and spatial coordination in payment for ecosystem services schemes. *Journal of Environmental Economics and Management*, 83, 68–89.

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

Banerjee, S., Kwasnica, A.M. & Shortle, J.S. (2012) Agglomeration bonus in small and large local networks: a laboratory examination of spatial coordination. *Ecological Economics*, 84, 142–152.

Beck, A., Kerschbamer, R., Qiu, J. & Sutter, M. (2014) Car mechanics in the lab—investigating the behavior of real experts on experimental markets for credence goods. *Journal of Economic Behavior & Organization*, 108, 166–173.

Behaghel, L., Macours, K. & Subervie, J. (2019) How can randomised controlled trials help improve the design of the common agricultural policy? *European Review of Agricultural Economics*, 46(3), 473–493.

Bell, A., Parkhurst, G., Droppelmann, K. & Benton, T.G. (2016) Scaling up pro-environmental agricultural practice using agglomeration payments: proof of concept from an agent-based model. *Ecological Economics*, 126, 32–41.

Bellemare, C. & Kröger, S. (2007) On representative social capital. *European Economic Review*, 51(1), 183–202.

Belot, M., Duch, R. & Miller, L. (2015) A comprehensive comparison of students and non-students in classic experimental games. *Journal of Economic Behavior & Organization*, 113, 26–33.

Bernold, E., Gsottbauer, E., Ackermann, K.A. & Murphy, R.O. (2015) Social framing and cooperation: the roles and interaction of preferences and beliefs. Available at SSRN 2557927.

Botazzi, P., Wiik, E., Crespo, D. & Jones, J.P.G. (2018) Payment for environmental "self-service": exploring the links between Farmers' motivation and additionality in a conservation incentive programme in the Bolivian Andes. *Ecological Economics*, 150, 11–23.

Brick, K. & Visser, M. (2015) Risk preferences, technology adoption and insurance uptake: a framed experiment. *Journal of Economic Behavior & Organization*, 118, 383–396.

Carpenter, J. & Seki, E. (2011) Do social preferences increase productivity? Field experimental evidence from fishermen in Toyama Bay. *Economic Inquiry*, 49(2), 612–630.

Cason, T.N. & Gangadharan, L. (2003) Transactions costs in tradable permit markets: an experimental study of pollution market designs. *Journal of Regulatory Economics*, 23(2), 145–165.

Cason, T.N. & Gangadharan, L. (2004) Auction design for voluntary conservation programs. *American Journal of Agricultural Economics*, 86(5), 1211–1217.

Cason, T.N. & Gangadharan, L. (2006) Emissions variability in tradable permit markets with imperfect enforcement and banking. *Journal of Economic Behavior & Organization*, 61(2), 199–216.

Cason, T.N. & Gangadharan, L. (2011) Price discovery and intermediation in linked emissions trading markets: a laboratory study. *Ecological Economics*, 70(7), 1424–1433.

Cason, T.N. & Raymond, L. (2011) Framing effects in an emissions trading experiment with voluntary compliance. In: *Experiments on energy, the environment, and sustainability*. Bradford: Emerald Group Publishing Limited, pp. 77–114.

Cason, T.N. & Wu, S.Y. (2017) *Subject pools and deception in agricultural and resource economics experiments*. Working Paper, Purdue University.

Charness, G. & Gneezy, U. (2012) Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, 83(1), 50–58.

Cleave, B.L., Nikiforakis, N. & Slonim, R. (2013) Is there selection bias in laboratory experiments? The case of social and risk preferences. *Experimental Economics*, 16(3), 372–382.

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

Cronk, L. & Wasielewski, H. (2008) An unfamiliar social norm rapidly produces framing effects in an economic game. *Journal of Evolutionary Psychology*, 6443(4), 1789–2082.

Croson, R. & Gneezy, U. (2009) Gender differences in preferences. *Journal of Economic Literature*, 47(2), 448–474.

Deaton, A. & Cartwright, N. (2018) Understanding and misunderstanding randomized controlled trials. *Social Science Medicine*, 210, 2–21.

Dörschner, T. & Musshoff, O. (2015) How do incentive-based environmental policies affect environment protection initiatives of farmers? An experimental economic analysis using the example of species richness. *Ecological Economics*, 114, 90–103.

Durlauf, L.E. & Blume, S.N. (2009) *Behavioural and experimental economics*. London: Palgrave Macmillan.

Eckel, C.C. & Grossman, P.J. (2000) Volunteers and pseudo-volunteers: the effect of recruitment method in dictator experiments. *Experimental Economics*, 3(2), 107–120.

Ellingsen, T., Johannesson, M., Mollerstrom, J. & Munkhammar, S. (2011) Social framing effects: preferences or beliefs? *Games and Economic Behavior*, 76(1), 117–130.

Elliott, C.S. & Hayward, D.M. (1998) The expanding definition of framing and its particular impact on economic experimentation. *Journal of Socio-Economics*, 27(2), 229–243.

Elliott, C.S., Hayward, D.M. & Canon, S. (1998) Institutional framing: some experimental evidence. *Journal of Economic Behavior and Organization*, 35(4), 455–464.

Exadaktylos, F., Espín, A.M. & Brañas-Garza, P. (2013) Experimental subjects are not different. *Scientific Reports*, 3, 1213.

Falk, A., Meier, S. & Zehnder, C. (2013) Do lab experiments misrepresent social preferences? The case of self-selected student samples. *Journal of the European Economic Association*, 11(4), 839–852.

Fehr, E. & Fischbacher, U. (2003) The nature of human altruism. *Nature*, 425, 785–791.

Fehr, E. & List, J.A. (2004) The hidden costs and returns of incentives – trust and trustworthiness among CEOs. *Journal of the European Economics Association*, 2(5), 743–771.

Ferraro, P.J. (2009) Counterfactual thinking and impact evaluation in environmental policy. *New Directions for Evaluation*, 2009(122), 75–84.

Ferré, M., Engel, S. & Gsottbauer, E. (2018) Which agglomeration payment for a sustainable Management of Organic Soils in Switzerland? – an experiment accounting for Farmers' cost heterogeneity. *Ecological Economics*, 150, 24–33.

Fréchette, G.R. (2015) Laboratory experiments: professionals versus students. In: Fréchette, G.R. & Schotter, A. (Eds.) *Handbook of experimental economic methodology*. Oxford, UK: Oxford University Press, pp. 360–390.

Fréchette, G.R. (2016) Experimental economics across subject populations. In: Kagel, J.H. & Roth, A.E. (Eds.) *Handbook of experimental economics*, Vol. 2. Princeton, NJ: Princeton University Press, pp. 435–480.

Gächter, S., Orzen, H., Renner, H. & Starmer, C. (2009) Are experimental economists prone to framing effects? A natural field experiment. *Journal of Economic Behavior and Organization*, 70, 443–446.

Gatiso, T.T., Vollan, B., Vimal, R. & Kühl, H.S. (2018) If possible, incentivize individuals not groups: evidence from lab-in-the-field experiments on Forest conservation in rural Uganda. *Conservation Letters*, 11(1), e12387.

Gneezy, U. & Imas, A. (2017) Chapter 10 – Lab in the field: measuring preferences in the wild. In: Banerjee, A.V. & Duflo, E. (Eds.) *Handbook of economic field experiments*, Vol. 1. Amsterdam: North-Holland, pp. 439–464.

Handberg, Ø.N. & Angelsen, A. (2015) Experimental tests of tropical forest conservation measures. *Journal of Economic Behavior & Organization*, 118, 346–359.

Handberg, Ø.N. & Angelsen, A. (2019) Pay little, get little; pay more, get a little more: a framed forest experiment in Tanzania. *Ecological Economics*, 156, 454–467.

Harrison, G.W. & List, J.A. (2004) Field experiments. *Journal of Economic Literature*, 42(4), 1009–1055.

Hasson, R., Löfgren, A. & Visser, M. (2010) Climate change in a public goods game: investment decision in mitigation versus adaptation. *Ecological Economics*, 70(2), 331–338.

Heinz, N. & Koessler, A.K. (2021) Other-regarding preferences and pro-environmental behaviour: an interdisciplinary review of experimental studies. *Ecological Economics*, 184, 106987.

Homar, A.R. & Cvelbar, L.K. (2021) The effects of framing on environmental decisions: a systematic literature review. *Ecological Economics*, 183, 106950.

Hossain, T. & List, J.A. (2012) The Behavioralist visits the factory: increasing productivity using simple framing manipulations. *Management Science*, 58(12), 2151–2167.

Howley, P., Buckley, C., Donoghue, C.O. & Ryan, M. (2015) Explaining the economic 'irrationality' of farmers' land use behaviour: the role of productivist attitudes and non-pecuniary benefits. *Ecological Economics*, 109, 186–193.

Janssen, M.A., Goldstone, R.L., Menczer, F. & Ostrom, E. (2008) Effect of rule choice in dynamic interactive spatial commons. *International Journal of the Commons*, 2(2), 288–312.

Janssen, M.A., Holahan, R., Lee, A. & Ostrom, E. (2010) Lab experiments to study social-ecological systems. *Science*, 328, 613–617.

Janssen, M.A., Lee, A. & Waring, T.M. (2014) Experimental platforms for behavioral experiments on social-ecological systems. *Ecology and Society*, 19(4), 20.

Kaczan, D.J. & Swallow, B.M. (2019) Forest conservation policy and motivational crowding: experimental evidence from Tanzania. *Ecological Economics*, 156, 444–453.

Kessler, J.B. & Vesterlund, L. (2015) The external validity of laboratory experiments: qualitative rather than quantitative effects. In: *Handbook of Experimental Economic Methodology*. Oxford: Oxford University Press. Available from: https://doi.org/10.1093/acprof:oso/9780195328325.003.0020

Kesternich, M., Reif, C. & Rübbelke, D. (2017) Recent trends in behavioral environmental economics. *Environmental and Resource Economics*, 67(3), 403–411.

Kieninger, R., Gugerell, K., Biba, V., Auberger, I., Winter, S. & Penker, M. (2018) Motivation crowding and participation in Agri-environmental schemes – the case of the Austrian Öpul-Programme in vineyards. *European Countryside*, 10(3), 355–376.

Koneberg, C., Yaish, M. & Stocké, V. (2010) Norms and rationality in electoral participation and in the rescue of Jews in WWII: an application of the model of frame selection. *Rationality and Society*, 22, 3–36.

Le Coent, P., Préget, R. & Thoyer, S. (2014) Why pay for nothing? *An Experiment on a Conditional Subsidy Scheme in a Threshold Public Good Game: Economics Bulletin*, 34, 1976–1989.

Levin, I.P., Schneider, S.L., & Gaeth, G.J. (1998) All frames are not created equal: a typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, 76(2), 149–188.

Levitt, S.D. & List, J.A. (2007) What do Laboratory experiments measuring social preferences reveal about the real world ? *The Journal of Economic Perspectives*, 21, 153–174.

Loft, L., Gehrig, S., Le, D.N. & Rommel, J. (2019) Effectiveness and equity of payments for ecosystem services: real-effort experiments with Vietnamese land users. *Land Use Policy*, 86, 218–228.

Maart-Noelck, S.C. & Musshoff, O. (2013) 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.

Meier, S. (2006) Does framing matter for conditional cooperation? Evidence from a natural field experiment. *Contributions to Economic Analysis & Policy*, 5(2), 1–21.

Milinski, M., Semmann, D., Krambeck, H.-J. & Marotzke, J. (2006) Stabilizing the Earth's climate is not a losing game: supporting evidence from public goods experiments. *PNAS*, 103(11), 3994–3998.

Murphy, R.O., Ackermann, K.A. & Handgraaf, M.J.J. (2011) Measuring social value orientation. *Judgment and Decision Making*, 6(8), 771–781.

Omar, A.-U. & List, J.A. (2013) *On the generalizability of experimental results in economics: with a response to Camerer*. No. w19666. National Bureau of Economic Research.

Palm-Forster, L.H., Suter, J.F. & Messer, K.D. (2018) Experimental evidence on policy approaches that link agricultural subsidies to water quality outcomes. *American Journal of Agricultural Economics*, 101(1), 109–133.

Paluck, E.L. & Shafir, E. (2017) Chapter 6 – The psychology of construal in the design of field experiments. Revised version, prepared following the NBER Conference on Economics of Field Experiments, organized by Esther Duflo & Abhijit Banerjee. In: Banerjee, A.V. & Duflo, E. (Eds.) *Handbook of economic field experiments*, Vol. 1. North-Holland, pp. 245–268. Available from: https://doi.org/10.1016/bs.hefe.2016.12.001

Parkhurst, G.M. & Shogren, J.F. (2007) Spatial incentives to coordinate contiguous habitat. *Ecological Economics*, 64(2), 344–355.

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

Peth, D., Mußhoff, O., Funke, K. & Hirschauer, N. (2018) Nudging farmers to comply with water protection rules–experimental evidence from Germany. *Ecological Economics*, 152, 310–321.

Pevnitskaya, S. & Ryvkin, D. (2013) Environmental context and termination uncertainty in games with a dynamic public bad. *Environment and Development Economics*, 18(1), 27–49.

Reddy, S.M.W., Wardropper, C., Weigel, C., Masuda, Y., Harden, S., Ranjan, P. et al. (2020) Conservation behavior and effects of economic and environmental message frames. *Conservation Letters*, 13(2), e12750.

Rege, M. & Telle, K. (2004) The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics*, 88, 1625–1644.

Reutemann, T., Engel, S. & Pareja, E. (2016) How (not) to pay — field experimental evidence on the design of REDD + payments. *Ecological Economics*, 129, 220–229.

Roe, B.E. & Just, D.R. (2009) Internal and external validity in economics research: tradeoffs between experiments, field experiments, natural experiments, and field data. *American Journal of Agricultural Economics*, 91(5), 1266–1271.

Rolfe, J., Windle, J. & McCosker, J. (2009) Testing and implementing the use of multiple bidding rounds in conservation auctions: a case study application. *Canadian Journal of Agricultural Economics/Revue Canadienne d'Agroeconomie*, 57(3), 287–303.

Rommel, J., Hermann, D., Müller, M. & Mußhoff, O. (2017) Comprehension in risk elicitation experiments. *Applied Economics Letters*, 24(9), 627–634.

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

Salk, C., Lopez, M.C. & Wong, G. (2017) Simple incentives and group dependence for successful payments for ecosystem services programs: evidence from an experimental game in rural Lao PDR. *Conservation Letters*, 10(4), 414–421.

Sattler, C. & Nagel, U.J. (2010) Factors affecting farmers' acceptance of conservation measures—a case study from North-Eastern Germany. *Land Use Policy*, 27(1), 70–77.

Schilizzi, S. & Latacz-Lohmann, U. (2007) Assessing the performance of conservation auctions: an experimental study. *Land Economics*, 83(4), 497–515.

Schram, A. (2005) Artificiality: the tension between internal and external validity in economic experiments. *Journal of Economic Methodology*, 12(2), 225–237.

Sereke, F., Dobricki, M., Wilkes, J., Kaeser, A., Graves, A.R., Szerencsits, E. et al. (2016) Swiss farmers don't adopt agroforestry because they fear for their reputation. *Agroforestry Systems*, 90(3), 385–394.

Sheeder, R.J. & Lynne, G.D. (2011) Empathy-conditioned conservation. *Land Economics*, 87(3), 433–452.

Snowberg, E. & Leeat, Y. (2021) Testing the waters: behavior across participant pools. *American Economic Review*, 111(2), 687–719.

Suter, J.F. & Vossler, C.A. (2013) Towards an understanding of the performance of ambient tax mechanisms in the field: evidence from upstate New York dairy farmers. *American Journal of Agricultural Economics*, 96(1), 92–107.

Tavoni, A., Dannenberg, A., Kallis, G. & Löschel, A. (2011) Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *PNAS*, 108(29), 11825–11829.

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

Thoyer, S. & Préget, R. (2019) Enriching the CAP evaluation toolbox with experimental approaches: introduction to the special issue. *European Review of Agricultural Economics*, 46(3), 347–366.

Tversky, A. & Kahneman, D. (1981) The framing of decisions and the psychology of choice. *Science*, 211, 453–458.

Viceisza, A.C. (2016) Creating a lab in the field: economics experiments for policymaking. *Journal of Economic Surveys*, 30(5), 835–854.

Vollan, B. & Ostrom, E. (2010) Cooperation and the commons. *Science*, 330, 923–924.

Voors, M., Demont, M. & Bulte, E. (2016) New experiments in agriculture. *African Journal of Agricultural and Resource Economics*, 11(311–2016-5646), 1–7.

Waichman, I. & Ness, C. (2012) Farmers' performance and subject pool effect in decentralized bargaining markets. *Economics Letters*, 115(3), 366–368.

Warziniack, T., Shogren, J.F. & Parkhurst, G. (2007) Creating contiguous forest habitat: an experimental examination on incentives and communication. *Journal of Forest Economics*, 13(2–3), 191–207.

Wünscher, T., Engel, S. & Wunder, S. (2008) Spatial targeting of payments for environmental services: a tool for boosting conservation benefits. *Ecological Economics*, 65, 822–833.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ferré, M., Engel, S. & Gsottbauer, E. (2023) External validity of economic experiments on Agri-environmental scheme design. *Journal of Agricultural Economics*, 74, 661–685. Available from: https://doi.org/10.1111/1477-9552.12529