



HAL
open science

When nudges backfire: evidence from a randomised field experiment to boost biological pest control

Sylvain Chabé-Ferret, Philippe Le Coënt, Caroline Lefebvre, Raphaële Préget, Francois Salanié, Julie Subervie, Sophie Thoyer

► To cite this version:

Sylvain Chabé-Ferret, Philippe Le Coënt, Caroline Lefebvre, Raphaële Préget, Francois Salanié, et al..
When nudges backfire: evidence from a randomised field experiment to boost biological pest control.
Annals of Economics and Statistics, 2024, 156, pp.9-30. 10.2307/48804180 . hal-04900081

HAL Id: hal-04900081

<https://brgm.hal.science/hal-04900081v1>

Submitted on 20 Jan 2025

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

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



Distributed under a Creative Commons Attribution 4.0 International License

WHEN NUDGES BACKFIRE: EVIDENCE FROM A RANDOMISED FIELD EXPERIMENT
TO BOOST BIOLOGICAL PEST CONTROL

SYLVAIN CHABÉ-FERRET^a, PHILIPPE LE COËNT^b, CAROLINE LEFEBVRE^c,
RAPHAËLE PRÉGET^d, FRANCOIS SALANIÉ^e, JULIE SUBERVIE^f AND SOPHIE
THOYER^g

Nudges are increasingly used to alter the behaviour of economic agents as an alternative to monetary incentives. However, little is known as to whether nudges can backfire, that is, how and when they may generate effects opposite to those they intend to achieve. We provide the first field evidence of a nudge that is designed to encourage pro-environmental behaviour, which instead backfires. We randomly allocate a social comparison nudge inviting wine-growers to adopt biological pest control as an alternative to chemical pesticide use. We find that our nudge decreases by half the adoption of biological pest control among the largest vineyards, where the bulk of adoption occurs. We show that this result can be rationalised in an economic model where wine-growers and wine-grower cooperative managers bargain over future rents generated by the adoption of biological pest control. This study highlights the importance of experimenting on a small scale with nudges aimed at encouraging adoption of virtuous behaviours in order to detect unexpected adverse effects, particularly in contexts where negotiations on the sharing of the costs of adoption are likely to occur.

JEL Codes: D90, Q25, Q58.

Keywords: Nudges, Behavioural Economics, Pesticides, Government Policy.

We thank the Cave Coopérative VPE—the experiment outlined in this paper would never have taken place without them. We acknowledge financial support from IDEX project AMEP (Advancing methods for evaluating environmental/health policies), ANR (project PENSEE # ANR-16-CE32-0011, Programme d’Investissement d’Avenir # ANR-17-EURE-0010 and CASD (Centre d’accès sécurisé aux données)). We also acknowledge financial support from the project ‘Facilitate public Action to exit from peSTicides (FAST)’ with the reference 20-PCPA-0005. We thank Jennifer Alix-Garcia, Nicolas Treich, Erez Yoeli, and the participants of the first PENSEE workshop in Toulouse, the World Congress of Environmental and Resources Economists in Gothenburg, and the Congress of the European Association of Agricultural Economics in Prague for their useful comments and suggestions. Any errors are our own.

^aCorresponding author. TSE, INRAE, U. Toulouse Capitole, Toulouse, France. Institute for Advanced Studies, Toulouse, France. sylvain.chabe-ferret@tse-fr.eu

^bBRGM, U. Montpellier, Montpellier, France; G-eau, UMR 183, INRAE, CIRAD, IRD, AgroParisTech, Institut Agro Montpellier, BRGM, Montpellier, France. p.lecoent@brgm.fr

^cCave Coopérative Les Vignerons du Pays d’Ensérune. caroline.lefebvre@labonatoli.fr

^dCEE-M, U. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France.
raphaele.preget@inrae.fr

^eTSE, INRAE, U. Toulouse Capitole, Toulouse, France. francois.salanie@inrae.fr

^fCEE-M, U. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France.
julie.subervie@inrae.fr

^gCEE-M, U. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France.
sophie.thoyer@inrae.fr

1. INTRODUCTION

Nudges aim at modifying agents' behaviour by tweaking their context of choice without resorting to monetary incentives (Thaler and Sunstein, 2009). Nudges are increasingly used by policymakers worldwide: for many years now, many countries, and first and foremost the UK, Germany and the USA, have started exploring the application of behavioral insights to their policies and programs, through so-called "nudge units". Social comparison nudges, where agents receive information about peer behaviour, are extensively used to encourage environmental conservation (Croson and Treich, 2014; Schubert, 2017) and have been shown to decrease electricity use (Allcott and Rogers, 2014; Delmas et al., 2013) and water consumption (Chabe-Ferret et al., 2019; Ferraro et al., 2011). A meta-analysis of more than 440 estimates shows that while nudges promote behavior change overall, their impact varies considerably from one behavioral domain to another, with effect sizes up to 2.5 times greater for food choices than for other decisions (Mertens et al., 2022).

An important outstanding question about the effectiveness of nudges, however, is whether they can backfire and prompt recipients to adopt harmful behaviour, and if so, in what context and why. A few empirical studies have reported that nudges can fail or backfire (Sunstein, 2017; Thunstrom et al., 2018). In particular, social comparison nudges have been shown to decrease savings behavior (Beshears et al., 2015), organ donation (Behavioral Insights Team, 2013), claims of social benefits (Bhargava and Manoli, 2015), and the purchase of healthy food (Diaz-Beltran et al., 2023; Stuber et al., 2022).¹ In this paper, we provide evidence of a social comparison nudge backfiring in the context of environmental conservation, an important area of application of social comparison nudges, along with a novel explanation for why it did.

We conducted a Randomized Controlled Trial (RCT) with a French wine-growers cooperative to test whether social comparison nudges can improve farmers' adoption of biological pest control as an alternative to chemical pesticide use. A randomly selected group of wine-growers received a flyer and text messages containing a social comparison message informing them of the widespread adoption of biological pest control in the neighbouring cooperative. Wine-growers in the control group received the same mailing and text messages, but without the social comparison component. The alternative to pesticide use that our nudge seeks to promote is mating disruption, an efficient and expanding method of biological pest control used in European vineyards against grape berry moths (Delbac et al., 2013; Hoffmann and Thiery, 2010). Mating disruption uses pheromone dispensers to disrupt the mating of pests and is able to drastically reduce their offspring. Our nudge aimed at informing wine-growers that farmers similar to them had successfully adopted biological pest control on a large scale, hinting that expected costs were probably over-estimated and could be revised downwards.

We find that our nudge decreased the adoption of biological pest control by half amongst the largest farms (for whom adoption reached 40% in the control group). We show that this result can be rationalised in an economic model where cooperative managers and wine-growers bargain over future rents generated by the adoption of this new technology. In our model, the nudge signals to wine-growers that cooperative managers value the adoption

¹Publication bias and file drawer effects may explain why reports of nudges backfiring are rare in the published literature (DellaVigna and Linos, 2022; Maier et al., 2022; Nemati and Penn, 2020).

of biological pest control more than the wine-growers initially believed. Consequently, wine-growers withhold adoption in order to extract a larger cut of the gains.

Several pieces of information from the field tend to support this model. First, in interviews carried out with the wine-growers after the results of the experiment were known, two primary explanations were put forward. The first was based on a psychological reactance to social comparisons. The other, more in line with our theoretical model, suggests that the nudge backfired because it was interpreted as implying that the wine-growers would receive no direct reward from the cooperative for adopting biological pest control. Second, treated wine-growers ended up adopting the practice three years after the experiment at the same rate as wine-growers from the control group, in line with the idea that the nudge triggered a long-lasting bargaining procedure. Our results thus bring a note of caution regarding the increasing number of nudges implemented within organisations or involving professional economic agents (Chabe-Ferret et al., 2019; Earnhart and Ferraro, 2021; Gosnell et al., 2016). Nudges may indeed backfire in economic contexts in which agents feel they can bargain to increase their share of the gains brought about by adopting the nudged behaviour.

Conservation nudges that rely on social comparison have already been shown to have boomerang effects, where the most virtuous agents slack off when they receive information about the average behaviour amongst their peers.² Nevertheless, the magnitude of this type of boomerang effect, as reported in the literature, is generally not great enough to dominate conservation effects in the least virtuous agents (Ayres et al., 2013; Chabe-Ferret et al., 2023; Fischer, 2008; Schultz et al., 2007). The only exception is when the advertised behaviour is so rare and extreme that the social comparison encourages most agents to slack off (Richter et al., 2018). In the present study, the social comparison nudge backfires despite most members of the comparison group having adopted the advertised technique.

Several mechanisms have been put forward to explain why nudges might backfire. Psychological phenomena such as reactance (Brehm, 1966; Clee and Wicklund, 1980; Osman, 2020), discouragement (Beshears et al., 2015), and seeing the behaviour of the reference group as unattractive (Bhargava and Manoli, 2015), have been cited as possible explanations. More recently, Bolton et al. (2020) propose a model to explain why making actions observable might backfire, based on social image motives. We contribute to this ongoing investigation by studying the effect of sending a nudge when a principal and an agent rationally play a strategic bargaining game.

The rest of the paper is structured as follows. Section 2 presents the context of the experiment. Section 3 describes the experimental design. Section 4 presents the results. Section 5 introduces the theoretical model. Section 6 concludes.

²In the case of nudges having unexpected effects, we can distinguish between “boomerang effects” and “backfiring”. Boomerang effects occur when agents, whose behavior is already more virtuous than the announced norm, react by adopting behavior closer to the announced norm, that is to say by becoming less virtuous. Backfiring occurs when agents, whose behavior is initially worse than the announced norm, react to a social comparison nudge by moving further away from the norm. Chabe-Ferret et al. (2019) find that their social comparison nudge has a boomerang effect: it increases water consumption among farmers who did not use any initially.

2. CONTEXT OF THE EXPERIMENT

In many French regions, pesticide use is a major source of water pollution. This problem is particularly acute in the wine-producing regions. In order to control the development of the grapevine moth (*Lobesia botrana*), whose larvae inflict great losses in grape production and quality, French wine-growers make use of a toxic chemical insecticide that is believed to have harmful effects on environmental ecosystems and human health. In the South of France, pesticides used on vineyards have been detected in more than 92% of rivers.

Mating disruption is an alternative to pesticides that uses dispensers to saturate the air above and between the grapevines with female sexual pheromones used by female moths to call males for mating. The ubiquitous presence of pheromones leads to the disorientation of the males and suppresses the calling behaviour of the females. The overall effect is to reduce the number of offspring produced by the pests and thus curb the damage to the grapes. The technique is primarily used against two pests: the tortricid moth species *Lobesia botrana* and *Eupoecilia ambiguella*, commonly called the European grapevine and grape berry moths, respectively. Compared to the spraying of insecticides, which cannot be done after heavy rain, the control of grape berry moths is less dependent on weather conditions during the vegetation period. Moreover, it is easier to manage than insecticide treatments, because it does not require the monitoring of oviposition periods (Hoffmann and Thiery, 2010).

Although this technique has been shown to be effective in a number of contexts, adoption rates in France are still very low: the proportion of the French vineyard area protected by biological pest control is currently only 3%, whilst it reaches 65% of the vineyards in Germany and 43% in Switzerland. There are at least three reasons for this. First, in the absence of subsidies, mating disruption is more expensive than insecticide use. French wine-growers apply an average of two pesticide treatments per year to control the moth,³ which cost approximately 35 euros per hectare each. The cost of the pheromone diffusers is around 110 euros per hectare, plus the time which must be devoted by the farmer to setting-up the diffusers in the vineyards, which can add another 200 euros per hectare. Second, to be effective, mating disruption requires plots or contiguous blocks of plots covering at least 10 hectares. In many parts of Europe, it is unusual to find a single farm with such large blocks of vineyard, which often requires that wine-growers spatially coordinate their efforts in order to simultaneously adopt the biocontrol technique on adjacent plots. Third, for many years, the use of chemical pesticides has been the norm in the French wine-growing sector, involving two chemical insect treatments per year on average. Habits and routines are hard to change. French farmers also doubt the efficacy of biological pest control.

We conducted an RCT in collaboration with a wine-growers cooperative, “Les Vignerons du Pays d’Enserune” (hereafter the VPE cooperative). One of the cooperative’s goals is to reduce pesticide use in order to improve the reputation of the wines and respond better to consumer demand, prevent criticism from neighbours increasingly concerned about pesticide-related health issues, and contribute to water quality and biodiversity improvement. Sustainable practices are a crucial marketing point that the cooperative seeks

³According to the statistics of the French Ministry of Agriculture, accessible here: <http://agreste.agriculture.gouv.fr/IMG/ods/pratiquesviticulture2015T2bsva.ods>.

to leverage. The VPE cooperative has 650 members accounting for 3,200 hectares of vineyards in the South of France's Occitanie region. At the time we initiated this partnership for the implementation of an RCT, the cooperative's technical staff had repeatedly invited wine-growers to information meetings on mating disruption, but attendance remained low and the alternative practice was only used on a very limited area.

Several features of the context support the idea that a social comparison nudge can have a positive impact on the adoption of mating disruption. First, monetary concerns are no longer a major reason for non-adoption. Indeed, as part of the Common Agricultural Policy of the European Union, agri-environmental measures are available to wine-growers with an annual payment of about 310 euros per hectare in exchange for replacing pesticide use with biocontrol techniques. Second, social comparison nudges can help alleviate the two main psychological costs detrimental to the adoption of a new technology: (i) the psychological costs of deviating from the norm of pesticide use; (ii) overestimation of coordination costs between neighbours for setting up the alternative practice. Our social comparison nudge therefore intended to reduce these two psychological costs by informing wine-growers that the alternative technique had already been largely adopted by neighbouring farmers. Our nudge aimed to convince farmers that the dominant norm on pesticide use was evolving, that turning to biological control was becoming a more readily accepted technique, and that many wine-growers had successfully achieved coordination, hinting that coordination costs might not be as high as wine-growers anticipated.

3. EXPERIMENTAL DESIGN

Treatment and outcomes

Our nudge includes two social comparisons: information on the total area at the provincial level (Herauld *département*), which is already protected with mating disruption (5,500 ha), and information on the high rate of adoption of biological pest control in the neighbouring wine-growers cooperative of Puicheric.

In October 2016, all the wine-growers of the VPE cooperative with a valid postal mailing address (a total of 532 wine-growers) received a letter from the cooperative advisory staff inviting them to a technical information meeting on biological pest control. Of these 532 wine-growers, about half of them (the treatment group) received an additional flyer with the following information: "*Using mating disruption to fight the grapevine moth works! Already 5,500 ha protected in the Herauld province. In 2017, the Puicheric cooperative will use it on its entire vineyard area.*" The other half (the control group) did not receive the flyer, only the invitation letter. The first impact of the nudge was then measured by the attendance rate of the technical meeting held on November 7.

In November 2016, less than a month later, all wine-growers with a valid mobile phone number (a total of 413 of the initial 532 wine-growers) received a text message: "*If you'd like to know more about grapevine moth control with mating disruption techniques, reply "OK" to this number, and the technician will call you back shortly.*" Wine-growers in the treatment group received an additional sentence in their text message: "*Already 5,500 ha protected in the Herauld province and soon the entire area of the Puicheric vineyards.*" A second measure of the impact of the nudge was thus whether a wine-grower replied "OK" to the technician. Note that all the wine-growers who received the additional sentence in the text message had also received the same message via the flyer.

Finally, eight months after the end of the experiment, in July 2017, we measured the adoption of mating disruption in order to check whether our social comparison nudge had

encouraged wine-growers to go beyond the simple decision to attend information meetings and indeed induced them to change their pest control technique. We also measured adoption of biocontrol three years after the experiment, in July 2020.

Our nudge displays five important features. First, we used the comparison with the Puicheric cooperative to show that adoption of biological pest control by similar wine-growers is possible. The VPE and Puicheric cooperatives, besides being geographically close, are both affiliated with the Union Foncalieu, which markets the wine produced by the two cooperatives. Puicheric is a smaller cooperative (with about 100 members), but it produces similar grape varieties and has the same objectives in terms of wine quality, since just like the VPE, it produces mainly wine covered by a Protected Geographical Indication (PGI) targeted for the export market. Second, we sent messages to the control group to isolate the impact of the social comparison nudge from the impact of the cooperative's standard promotion of biological control. Both treatment and control wine-growers thus had their attention directed towards biological control. The only difference between the two groups was the social comparison message. Third, we combined both postal mailings and text messages to increase the chances that wine-growers would pay attention to the social comparison message. Previous experiments with electricity and water consumers relied exclusively on postal mail messages, and their effectiveness may have been curbed by the limited attention that consumers pay to their electricity or water bills. Fourth, there was only a short time gap between the nudge and some of the measured outcomes (two weeks elapsed between the date the information meeting invitation was sent and the date the meeting was held, and wine-growers could reply instantly to the text message for information on biological pest control). This short time frame is an important feature of the experimental protocol, since it has been shown that the effect of nudges may fade over time. This timing thus allows us to capture the effect of the nudge when it is expected to be the greatest. Fifth, a number of studies have already shown that French wine-growers tend to be sensitive to the behaviour and opinions of their peers when deciding whether or not to adopt a technology (Kuhfuss et al., 2016; Le Coent et al., 2021). A nudge leveraging social comparison thus seemed likely to trigger changes in farmers' technology adoption choices.

Randomisation, stratification and inference

We tested our nudge using a stratified RCT. The strata were defined based on two variables: the area of the vineyard (3 classes: class 1 = missing value, class 2 for area ≤ 7 ha, and class 3 for area > 7 ha) and the geographical area (4 classes coded from 1 to 4 corresponding to the different zones that the cooperative covers). We chose these variables and classes with the help of the cooperative's technicians in order to reflect the diversity of the vineyards. The threshold of 7 hectares represents the limit generally agreed between "professional" winegrowers and farmers for whom viticulture constitutes a secondary activity. The four geographical zones are groupings of municipalities which joined the cooperative at different periods. Each zone is considered to have quite homogeneous agricultural practices, associated with specific pedoclimatic and local conditions. We ended up with 11 strata (one of the 12 strata was empty), each containing at least four wine-growers. The treatment was randomly allocated at the individual level within each stratum.

Estimation and inference in stratified designs have been recently clarified by Bugni et al. (2018, 2019). Following their suggestion, we implemented the *strata fixed effects*

estimator, by estimating the following equation using Ordinary Least Squares:

$$Y_i = \sum_{s \in S} \alpha_s 1[s(i) = s] + \beta R_i + \epsilon_i,$$

where R_i takes value one when farmer i has been randomized into the treatment and zero otherwise, $1[s(i) = s]$ are strata fixed effects ($1[A]$ takes value one when A is true and zero otherwise, and s is a mapping associating each farmer i to her strata s). β is consistent for the average treatment effect under randomization (Chabé-Ferret, 2023). We used heteroskedasticity-robust standard errors to estimate the precision of the strata fixed effects estimator. Bugni et al. (2018) show in their Theorem 4.3 that this estimator has correct coverage when the proportion of treated units in each strata is 0.5, as in our case. In order to check for robustness, we also estimated the average treatment effect using a simple *with/without* regression without any controls and a *fully saturated* model, also studied by Bugni et al. (2018, 2019). We found very similar results. We moreover computed the corrected standard errors suggested by Bugni et al. (2018, 2019) for the case when the proportion of treated units in each strata is not equal to 0.5, along with a randomization inference test. All results remained unchanged.

Since the experiment is stratified, not clustered, standard errors do not have to be adjusted for clustering, even if our outcomes might be spatially autocorrelated (Abadie et al., 2023; Barrios et al., 2012). For the sake of robustness, however, we computed cluster-robust standard errors for all of our estimators, by clustering at the *commune* level, the smallest administrative level in France, roughly equivalent to a US census block. We found that the cluster-robust standard errors are very similar to the basic heteroskedasticity-robust ones, without changing any of our results neither quantitatively nor qualitatively.

Table I presents descriptive statistics for the main sample, by strata and treatment status. Pre-treatment variables are adequately balanced in each strata, apart from age which is unbalanced across the smallest strata (1, 6, 8 and 10), but these differences wash out when averaging over the whole sample (even when focusing on small and large farms separately). Note also that this imbalance does not characterize the strata with large farms, which is where our most important result is found. We also have a slightly larger proportion of treated farms adopting mating disruption before the experiment (especially in strata 3). Again, this difference is not statistically significant when averaged across the whole sample and across the sample of large farms. Strata 9 contains only one observation and will not be used in the strata fixed effects regression.

Qualitative survey

During the summer of 2020, we returned to the field and conducted a qualitative survey with a sample of wine-growers who had been involved in the experiment. The interviews were conducted via telephone and recorded (with the agreement of the participants). We surveyed only large vineyards with a registered telephone number. We aimed to interview at least one wine-grower in each of nine strata, defined by whether or not they had received the nudge, whether they had adopted mating disruption, and whether they belonged to one of the three largest municipalities. We randomly ordered wine-growers within these strata and stopped collecting data once we had interviewed one wine-grower in each strata. In the end, we interviewed nine wine-growers. The survey was mostly composed of closed questions: we first asked the wine-growers about their background, then about their opin-

ion on our nudge, and finally about its expected impact. We then revealed the actual results of the RCT experiment to each interviewee (none of the survey participants had the means to know the results of the RCT in advance) and asked their opinion about what mechanisms might explain the results.

In the survey, four possible explanations were put forward as to why the nudge backfired: (i) the mention that 5,500 ha were already equipped with biocontrol devices in the region was disheartening to the wine-growers; (ii) mentioning that almost all wine-growers in the Puicheric cooperative had already adopted biocontrol techniques was experienced as unfair and/or irrelevant by the wine-growers; (iii) the fact that almost all of the growers in the Puicheric cooperative had already adopted biocontrol techniques was seen as a signal that the cooperative expected wine-growers to do the same without additional incentives; (iv) the fact that almost all wine-growers in the Puicheric cooperative had already adopted biocontrol techniques was interpreted as a sign that additional efforts to decrease chemical pesticide use and improve water quality locally were not necessary.

4. RESULTS

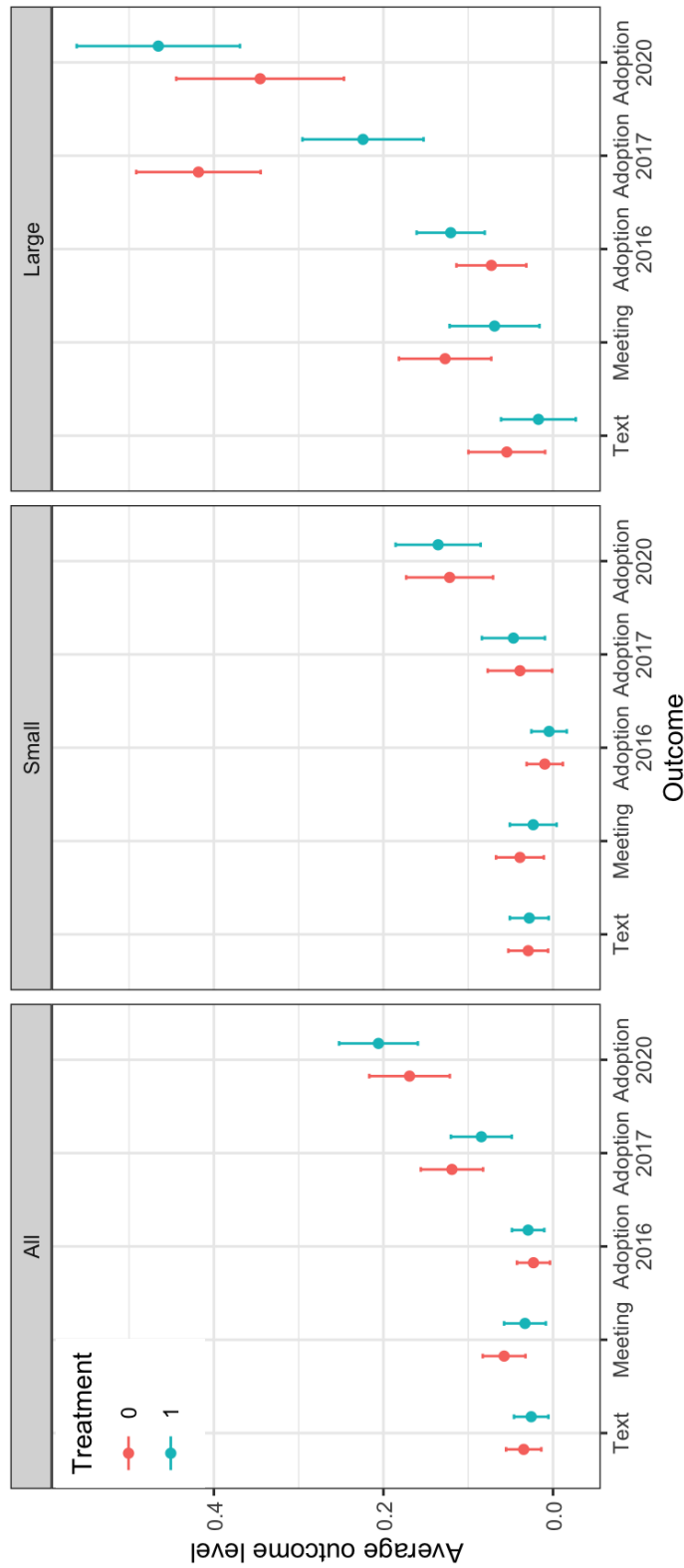
Figure 1 and Table II present the results of the analysis on the sample of 532 wine-growers included in the experiment. Figure 1 shows that the proportion of wine-growers who adopted biocontrol techniques increased over time in all groups, from 2% in 2016 to 20% in 2020. Large farms (area >7 ha) seem to lead the adoption dynamics, with an adoption rate increasing from around 10% in 2016 to over 40% in the control group in 2017.

TABLE I
DESCRIPTIVE STATISTICS BY STRATA

Strata	Treatment	Large farm	Area	Cell	Age	Adoption 2016	N
1	0	0	.	0.44	69.86	0.00	9
1	1	0	.	0.45	50.80	0.00	11
2	0	0	2.31	0.80	57.47	0.01	75
2	1	0	1.99	0.80	60.92	0.00	76
3	0	1	19.10	0.94	49.38	0.06	16
3	1	1	17.57	0.89	51.94	0.17	18
4	0	0	.	0.80	57.00	0.00	5
4	1	0	.	0.71	56.29	0.00	7
5	0	0	1.85	0.71	56.90	0.00	31
5	1	0	1.89	0.72	59.77	0.00	32
6	0	1	20.45	1.00	52.75	0.00	13
6	1	1	18.09	1.00	49.85	0.00	13
7	0	0	.	0.50	77.20	0.00	6
7	1	0	.	0.50	83.17	0.00	6
8	0	0	0.96	0.80	81.60	0.00	10
8	1	0	1.78	0.75	55.00	0.08	12
9	0	1	9.19	1.00	48.00	0.00	1
10	0	0	.	0.29	69.40	0.00	7
10	1	0	.	0.38	81.75	0.00	8
11	0	0	2.18	0.76	66.43	0.02	62
11	1	0	2.38	0.76	61.57	0.00	62
12	0	1	16.36	0.96	50.68	0.12	25
12	1	1	19.41	0.93	47.88	0.15	27

Notes: This table presents descriptive statistics by strata and treatment status. *Treatment* takes value 1 when a farm is assigned to receiving our nudge and 0 otherwise. *Large farm* takes value 1 when farm size is larger than 7 ha and 0 otherwise. When farm size is missing, *Large farm* takes value 0. *Cell* gives the proportion of winegrowers in the strata and treatment category that have a valid cell-phone number in our database. *Area* measures the area of the farm (in ha). Observations with missing information on farm size are assigned to a separate strata, identified by a dot (“.”). *Adoption 2016* measures the proportion of winegrowers who had adopted mating disruption in 2016, prior to our experiment. *N* is the number of observation by strata and treatment status.

Figure 1: Results of the field experiment



Notes: This figure presents the main results of our experiment. *All* refers to all farms in the sample (532 wine-growers). *Small* refers to the 419 farms with either a missing area or an area lower than 7 ha. *Large* refers to the 113 farms with a size larger than 7 ha. *Treatment* takes value 1 when a farm is assigned to receiving our nudge and 0 otherwise. *Text* measures the proportion of wine-growers who have replied to the text message sent by the caseworker asking them whether they were interested in receiving more information about mating disruption. *Meeting* measures the proportion of wine-growers who have participated in a meeting presenting mating disruption, which took place after our experiment. *Adoption 2016* measures the proportion of wine-growers who had adopted mating disruption in 2016, prior to our experiment. *Adoption 2017* measures the proportion of wine-growers who adopted mating disruption in 2017, after our experiment. *Adoption 2020* measures the proportion of wine-growers who adopted mating disruption in 2020, three years after our experiment. Error bars are 95% confidence intervals.

Amongst large vineyards, the adoption rate of biological pest control in 2017 (the year following the sending of the nudge) is around 20% in the treated group and around 40% in the control group, thus revealing an almost 20 percentage point (p.p.) or 50% decline of adoption due to the nudge. Table II shows that the estimated decrease is actually 18 p.p. amongst all large wine-growers and 20 p.p. in the sub-sample of large wine-growers equipped with a mobile phone.⁴ In 2020, there is no statistically significant difference in adoption rates between the treated and control groups: the adoption rate in the treated group eventually caught up with that of the control group. We do not find statistically significant differences between the treated and control groups, whatever the outcome considered, either in the overall sample or in the stratum of small vineyards. This seems consistent with the fact that adoption rates are lagging amongst small wine-growers. All results of robustness checks are provided in Table III. They all support our identification strategy.

Figure 2 plots the number of wine-growers interviewed who accepted one or more of the explanations put forward for why our nudge backfired. Only three of the nine wine-growers thought that the mention of 5,500 ha already using mating disruption would explain the result via a discouragement effect. Similarly, only two of them thought that wine-growers could have interpreted the mention that almost all wine-growers in the Puicheric cooperative had already adopted mating disruption as a sign that no further efforts were needed to decrease pesticide use in the area. The two leading explanations from six of the nine wine-growers were: (i) wine-growers reacted negatively to the mention of the Puicheric cooperative in feeling unfairly compared with the *do-gooders*; (ii) wine-growers felt that mentioning Puicheric meant that the VPE cooperative had no intention of giving any additional incentives to the farmers.

These two explanations are obviously not mutually exclusive and we have no way of choosing between them with our data. The first explanation, which is related to reactance, has already been studied in the literature. In what follows, we therefore further explore the second explanation with a model in which a nudge can trigger a bargaining game between the wine-grower and the management of the cooperative.

⁴The analysis on the group of 413 wine-growers with a valid mailing address and mobile phone number yields very similar results and are available upon request.

TABLE II
AVERAGE TREATMENT EFFECTS, BY FARM SIZE AND SAMPLE COMPOSITION

Outcome	Sample	Farm size		
		All	Small	Large
Meeting	All	-0.02 (0.02)	-0.02 (0.02)	-0.04 (0.05)
Text	All	-0.01 (0.01)	-0.00 (0.02)	-0.04 (0.03)
Adoption 2016	All	0.01 (0.01)	-0.01 (0.01)	0.04 (0.06)
Adoption 2017	All	-0.03 (0.02)	0.01 (0.02)	-0.18** (0.09)
Adoption 2020	All	0.04 (0.03)	0.01 (0.03)	0.13 (0.09)
Meeting		-0.03 (0.02)	-0.02 (0.02)	-0.04 (0.06)
Text		-0.01 (0.02)	-0.00 (0.02)	-0.04 (0.04)
Adoption 2016	Mobile	0.01 (0.02)	-0.01 (0.01)	0.05 (0.06)
Adoption 2017	Mobile	-0.03 (0.03)	0.02 (0.03)	-0.20** (0.09)
Adoption 2020	Mobile	0.06 (0.04)	0.03 (0.04)	0.14 (0.10)

Notes: This table presents the estimated average treatment effect of our experiment, by farm size and sample composition. *All* refers to all farms in the sample. *Small* refers to farms with either a missing area or an area of less than 7 ha. *Large* refers to farms larger than 7 ha.

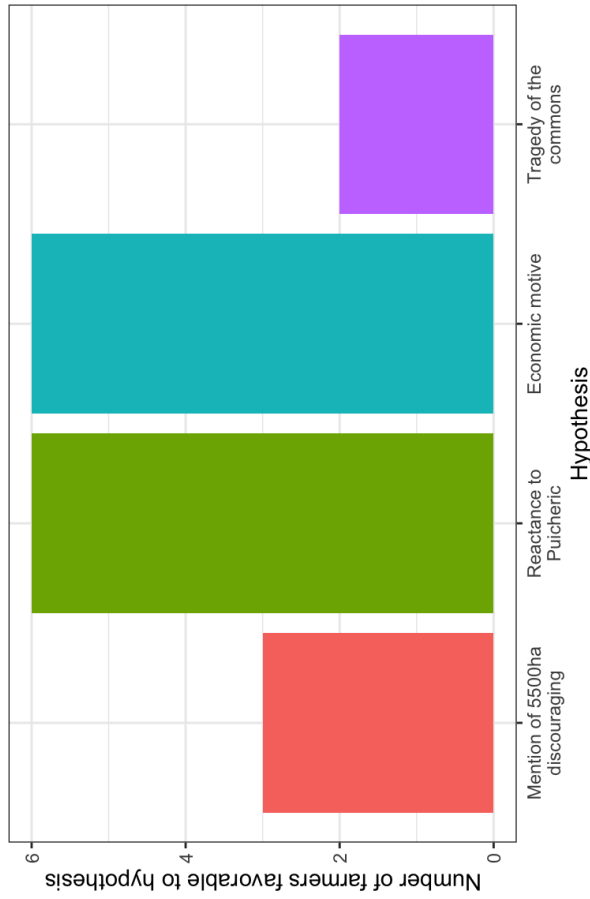
Meeting refers to the meeting presenting the information on mating disruption, which took place after our experiment. *Text* refers to the text sent by the caseworker asking wine-growers whether they were interested in receiving more information about mating disruption. *Adoption 2016* measures the proportion of wine-growers who had adopted mating disruption in 2016, prior to our experiment. *Adoption 2017* measures the proportion of wine-growers who adopted mating disruption in 2017, after our experiment. *Adoption 2020* measures the proportion of wine-growers who adopted mating disruption in 2020, three years after our experiment. Estimates are obtained using OLS regressions with strata fixed effects. Standard errors are in parentheses. * denotes 1% significance level, ** 5% level, *** 10% level.

TABLE III
EFFECTS OF THE TREATMENT BY ESTIMATOR, SAMPLE AND FARM SIZE

Outcome	Sample	Statistics	Farm size								
			All			Small			Large		
			SAT	SFE	WW	SAT	SFE	WW	SAT	SFE	WW
Meeting	All	Coef	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.04	-0.04	-0.06
Meeting	All	SeCoefBasic	0.02	0.02	0.02	0.02	0.02	0.02	0.05	0.06	0.06
Meeting	All	SeCoefBCS	0.02	0.02	0.02	0.02	0.02	0.02	0.06	0.06	0.06
Meeting	All	SeCoefCluster	0.02	0.02	0.02	0.02	0.02	0.02	0.05	0.05	0.05
Meeting	All	p.valueBCS	0.19	0.19	0.13	0.24	0.24	0.24	0.30	0.30	0.19
Text	All	Coef	-0.01	-0.01	-0.01	-0.00	-0.00	-0.00	-0.08	-0.04	-0.04
Text	All	SeCoefBasic	0.02	0.01	0.01	0.02	0.02	0.02	0.08	0.04	0.04
Text	All	SeCoefBCS	0.02	0.01	0.02	0.02	0.02	0.02	0.08	0.04	0.04
Text	All	SeCoefCluster	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03
Text	All	p.valueBCS	0.44	0.44	0.44	0.76	0.76	0.76	0.11	0.11	0.11
Adoption 2016	All	Coef	0.01	0.01	0.01	-0.01	-0.01	-0.01	0.06	0.04	0.05
Adoption 2016	All	SeCoefBasic	0.02	0.01	0.01	0.01	0.01	0.01	0.07	0.06	0.06
Adoption 2016	All	SeCoefBCS	0.02	0.01	0.01	0.01	0.01	0.01	0.07	0.06	0.06
Adoption 2016	All	SeCoefCluster	0.01	0.01	0.01	0.01	0.01	0.01	0.05	0.05	0.05
Adoption 2016	All	p.valueBCS	0.78	0.78	0.78	0.25	0.25	0.25	0.52	0.52	0.52
Adoption 2017	All	Coef	-0.03	-0.03	-0.03	0.01	0.01	0.01	-0.18**	-0.18**	-0.19**
Adoption 2017	All	SeCoefBasic	0.03	0.02	0.03	0.02	0.02	0.02	0.09	0.09	0.09
Adoption 2017	All	SeCoefBCS	0.03	0.02	0.03	0.02	0.02	0.02	0.09	0.09	0.09
Adoption 2017	All	SeCoefCluster	0.02	0.02	0.02	0.02	0.02	0.02	0.09	0.09	0.09
Adoption 2017	All	p.valueBCS	0.18	0.18	0.14	0.83	0.83	0.83	0.04	0.04	0.03
Adoption 2020	All	Coef	0.04	0.04	0.04	0.01	0.01	0.01	0.13	0.13	0.12
Adoption 2020	All	SeCoefBasic	0.03	0.03	0.03	0.03	0.03	0.03	0.09	0.09	0.09
Adoption 2020	All	SeCoefBCS	0.03	0.03	0.03	0.03	0.03	0.03	0.09	0.09	0.09
Adoption 2020	All	SeCoefCluster	0.03	0.03	0.04	0.03	0.03	0.03	0.07	0.07	0.08
Adoption 2020	All	p.valueBCS	0.25	0.25	0.25	0.66	0.66	0.66	0.12	0.12	0.18
Meeting	Cell	Coef	-0.03	-0.03	-0.03	-0.02	-0.02	-0.02	-0.04	-0.04	-0.06
Meeting	Cell	SeCoefBasic	0.02	0.02	0.02	0.03	0.02	0.02	0.06	0.06	0.06
Meeting	Cell	SeCoefBCS	0.02	0.02	0.02	0.03	0.02	0.02	0.06	0.06	0.06
Meeting	Cell	SeCoefCluster	0.02	0.02	0.02	0.02	0.02	0.02	0.05	0.05	0.06
Meeting	Cell	p.valueBCS	0.25	0.25	0.19	0.35	0.41	0.41	0.49	0.49	0.33
Text	Cell	Coef	-0.01	-0.01	-0.01	-0.00	-0.00	-0.00	-0.08	-0.04	-0.04
Text	Cell	SeCoefBasic	0.02	0.02	0.02	0.02	0.02	0.02	0.08	0.04	0.04
Text	Cell	SeCoefBCS	0.02	0.02	0.02	0.02	0.02	0.02	0.08	0.04	0.04
Text	Cell	SeCoefCluster	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03
Text	Cell	p.valueBCS	0.53	0.47	0.63	0.88	0.79	0.83	0.34	0.13	0.34
Adoption 2016	Cell	Coef	0.01	0.01	0.01	-0.01	-0.01	-0.01	0.07	0.05	0.05
Adoption 2016	Cell	SeCoefBasic	0.02	0.02	0.02	0.01	0.01	0.01	0.08	0.06	0.06
Adoption 2016	Cell	SeCoefBCS	0.02	0.02	0.02	0.01	0.01	0.01	0.08	0.06	0.06
Adoption 2016	Cell	SeCoefCluster	0.01	0.01	0.01	0.01	0.01	0.01	0.05	0.05	0.05
Adoption 2016	Cell	p.valueBCS	0.57	0.60	0.56	0.24	0.29	0.24	0.39	0.39	0.35
Adoption 2017	Cell	Coef	-0.04	-0.03	-0.04	0.03	0.02	0.02	-0.20**	-0.20**	-0.21**
Adoption 2017	Cell	SeCoefBasic	0.03	0.03	0.03	0.03	0.03	0.03	0.09	0.09	0.09
Adoption 2017	Cell	SeCoefBCS	0.03	0.03	0.03	0.03	0.03	0.03	0.09	0.09	0.09
Adoption 2017	Cell	SeCoefCluster	0.03	0.03	0.03	0.02	0.02	0.02	0.10	0.10	0.09
Adoption 2017	Cell	p.valueBCS	0.25	0.31	0.25	0.33	0.35	0.35	0.02	0.03	0.02
Adoption 2020	Cell	Coef	0.06	0.06	0.05	0.03	0.03	0.03	0.14	0.14	0.12
Adoption 2020	Cell	SeCoefBasic	0.04	0.04	0.04	0.04	0.04	0.04	0.10	0.10	0.10
Adoption 2020	Cell	SeCoefBCS	0.04	0.04	0.04	0.04	0.04	0.04	0.10	0.10	0.10
Adoption 2020	Cell	SeCoefCluster	0.03	0.03	0.03	0.03	0.03	0.03	0.08	0.08	0.08
Adoption 2020	Cell	p.valueBCS	0.16	0.15	0.17	0.53	0.49	0.49	0.15	0.12	0.15

Notes: This table presents the estimated average treatment effect of our experiment, by estimator, farm size and sample composition. All refers to all farms in the sample. Small refers to farms with either a missing area or an area of less than 7 ha. Large refers to farms larger than 7 ha. Meeting refers to the presence at the meeting introducing mating disruption which took place after our experiment. Text refers to responding “yes” to the text sent by the caseworker asking winegrowers whether they were interested in receiving more information about mating disruption. Adoption 2016 measures the proportion of winegrowers who had adopted mating disruption in 2016, prior to our experiment. Adoption 2017 measures the proportion of winegrowers who adopted mating disruption in 2017, after our experiment. Adoption 2020 measures the proportion of winegrowers who adopted mating disruption in 2020, three years after our experiment. WW refers to the with/without estimator. SFE refers to the strata fixed effects estimator. SAT refers to the saturated estimator. Coef refers to the treatment effect estimate. SeCoefBasic refers to the heteroskedasticity-robust standard error. SeCoefCluster refers to the standard error clustered at the commune level. SeCoefBCS refers to the estimate of the standard error with Bugni et al. (2019)’s correction. p.valueBCS refers to the randomization-inference-based p-value of the null of no treatment effect.

Figure 2: Explanations chosen by interviewed wine-growers as to why the nudge backfired



Notes: This figure presents the number of wine-growers (amongst the nine who took part in our post-experimental survey) who agreed with a given response as likely to explain why our nudge backfired. *Mention of 5,500 ha discouraging* refers to wine-growers believing that their peers found the mention of 5,500 ha already practicing mating disruption in the region as discouraging. *Reactance to Puicheric* refers to wine-growers believing that their peers reacted negatively to the mention of the fact that almost all wine-growers in the Puicheric cooperative had already adopted mating disruption. *Economic motive* refers to wine-growers believing that their peers interpreted the mention of the fact that almost all wine-growers in the Puicheric cooperative had already adopted mating disruption as a sign that the cooperative expected them to adopt mating disruption without providing any additional benefits. *Tragedy of the commons* refers to wine-growers believing that their peers interpreted the mention of the fact that almost all wine-growers in the Puicheric cooperative had already adopted mating disruption as a sign that additional efforts to decrease pesticide use were not necessary.

5. WHEN NUDGES BACKFIRE: A SIGNALLING GAME

In this section, we offer an economic explanation for why nudges may backfire, using the framework of a signalling game played by two rational agents, namely, the manager of the cooperative, who proposes to adopt biological pest control, and the individual wine-grower, who decides whether to accept this proposal. In a nutshell, more communication efforts exerted by the manager (via a nudge rather than a simple email/letter) may lead the wine-grower to suspect that the manager has private information on the collective benefits that could be gained from the wine-grower's adoption of biological pest control and intends not to share these benefits with him. The wine-grower thus reacts by delaying adoption, in the hope that the manager may later offer a better deal. This reasoning might explain why our nudge led to a smaller adoption rate of biological pest control in the treated group than in the control group.

Let us begin by describing the agents' payoffs. The bargaining game between the two agents takes place under asymmetric information on both sides. We can safely assume that the wine-grower (hereafter he) has private information about his payoff u from adoption. u aggregates the private benefits minus the private costs of biological pest control, plus perhaps some "warm glow" value for behaving in an environmentally-friendly manner. Note that u may be positive or negative.

Similarly, we can consider that the manager of the cooperative (hereafter she) better knows her own benefits $v > 0$ from the adoption of biological pest control, as this value may include gains in reputation or in profits from increased sales or increased prices of a more ecological wine. Additionally, the success of the adoption of biological pest control may depend on the surface area of the vineyards on which it is used; thus, it is in the manager's interest to enrol large farms. The larger the wine-grower's vineyard, the higher v .

We now propose a simple timing sequence for this game. In the first stage, the manager privately learns the value of v , and then chooses a design x for her proposal; x may be a simple email or may include more sophisticated nudges. Design x is thus chosen in some set X , with some cost $c(x)$. In the second stage, the wine-grower privately learns his type u , observes the proposal x , and decides whether or not to accept it. If he accepts the proposal, adoption takes place and the payoffs are u and $v - c(x)$.

If the game was to stop here, it would be easily solved: in all Perfect Bayesian Equilibria, there would be adoption if, and only if, it were in the wine-grower's interest ($u > 0$), and if the cooperative's top management chose the least costly design, that minimises $c(x)$. Note, however, that the cooperative has control over its budget and can implement transfers to its members. A rejection at the second stage may thus be interpreted as an element of a bargaining strategy: the wine-grower expects the manager to make a more favourable offer. To take this possibility into account, we enrich the game as follows: if the wine-grower rejects the proposal x made at the second stage, with some probability $\delta < 1$ we go to a third stage in which, in addition, the manager offers the wine-grower a transfer, $t \geq 0$. This transfer may occur, for example, inside the cooperative, in the form of a better price paid for the grapes harvested under biological pest control, or of a cost subsidy for the installation of diffusers. Once more, the wine-grower may accept or reject this new proposal, and then the game ends. Notice that δ may also be interpreted as a discount factor due to a delay between the second and the third stages. The game thus allows for both immediate and delayed adoption of biocontrol.

For technical reasons, we assume that v can take only two values $v_2 > v_1 \geq 0$ with

probabilities m_2 and m_1 ($m_1 + m_2 = 1$). The distribution of u is characterised by a cumulative distribution function F and a density function f , strictly positive on the set of real numbers. We assume that the ratio $(1 - F(u))/f(u)$ is decreasing with respect to u .⁵ A key assumption is that these values u and v are privately known and independent.

We are looking for Perfect Bayesian Equilibria (PBE) of this game, i.e., strategies that are best-responses to each other, and beliefs revised at each stage from prior beliefs (m_1, m_2) , using Bayes' rule whenever possible. An initial striking result is that all such equilibria are pooling (a formal proof is given in the Appendix):

PROPOSITION 1 *Whatever her private value v for the adoption of biological pest control, in equilibrium, the manager of the cooperative chooses the same design x .*

The intuition for this result can be sketched by proceeding by contradiction. Suppose that in equilibrium, a manager with a high private value chooses a more costly design. By observing this costly design, the wine-grower is thus able to infer that the manager's payoff is high. Accordingly, the wine-grower knows that if the last stage takes place, the manager will offer a more generous proposal. This in turn makes the wine-grower less willing to accept the initial proposal. Overall, by transmitting the information that v is high, the manager only delays adoption and makes adoption more costly to her budget, since adoption is more often accompanied by a transfer. This is the contradiction we needed. We finally show that the game admits a simple equilibrium:

PROPOSITION 2 *There exists a Pareto-dominating equilibrium, in which the manager always chooses the least costly communication design (i.e., without resorting to any nudge), whatever her private value for biocontrol adoption. Deviations to a costlier design are unprofitable, because they are interpreted by the wine-grower as proving that the manager has a high private value.*

Overall, as in many bargaining games under asymmetric information, it is better not to signal a strong interest for the object to be awarded. In the proposed equilibrium, adoption takes place immediately if the wine-grower's payoff is above a positive threshold u^* , because delaying is more costly for wine-growers with higher payoffs. Adoption occurs after a delay for wine-growers with $u \in [-t^*, u^*]$, where t^* is the transfer proposed in equilibrium at the last stage. When the manager deviates and offers a nudge, there is less immediate adoption, while the effects on delayed and/or on total adoption are ambiguous.⁶

These results may help us understand why our nudge backfired. Including a nudge in the design of the proposal was interpreted by wine-growers as a deviation from the pooling equilibrium and as evidence that the manager's payoff from adoption was high. The consequence is that large wine-growers delayed adoption, as observed in Figure 1, in the hope that they would extract some additional benefits from the manager later on. We did

⁵This "hazard rate" property is standard in game theory and is satisfied by many distributions, including normal, log-normal and uniform distributions. Moreover, for our purposes it needs only hold when u is negative.

⁶It should be noted that our data does not allow us to show a statistically significant effect of the nudge at the end of the bargaining period, only that the adverse effect eventually disappeared. This does not necessarily mean that there was not an ultimately positive impact of the nudge, but more likely that we lack the statistical power to highlight it. Indeed, our estimate of the (positive) effect of the nudge in 2020 among large farmers is very close to being significant at 10% (pvalue of 0.12).

not observe the last stage of the game, but Figure 1 also shows that delayed adoption took place after the initial proposal, and this is consistent with our model's predictions. Finally, the observation that large farms delay adoption more often may be due to the fact that larger vineyards are more indispensable to the success of the project, and thus are more tempted to exploit their stronger bargaining position.

6. CONCLUSION

We carried out a field experiment to test the effectiveness of a nudge initially designed to encourage the adoption of biological pest control in French viticulture. We bring an additional brick to the edifice of the literature that highlights the unexpected and possibly harmful effects of nudges, by focusing here on a still little-explored field, that of environmental and natural resources conservation. To better understand why our nudge backfired, we carried out a qualitative survey and further examined one of the two paths suggested by the answers in the survey, namely, that of a bargaining game between wine-growers and their cooperative's top management. Our model shows that a nudge can lead to a situation in which some wine-growers knowingly delay the adoption of the technique encouraged by the nudge if they interpret the nudge as a signal that the cooperative values the technique more than they initially believed. In such a case, they may withhold adoption to force the cooperative to share some of the gains of adoption with them.

This backfire effect may, however, only be short-lived. Indeed, if the wine-grower and the cooperative's top management are able to come to an agreement for sharing the benefits of the nudged behaviour, adoption can resume, as shown in both our model and in the results of our experiment. Nevertheless, the delay imposed by the nudge backfiring may be sizable. In our experiment, the nudge decreased adoption of biocontrol by 50%, from a baseline adoption rate of 40%, and adoption in the treated group caught up with that in the control group only three years after the experiment.

Our study suggests that when a nudge indirectly reveals information about the benefits expected by the principal who nudges the agents, the nudge can backfire and reduce, or at least delay, the adoption of the practices expected from the agents. Enriching the nudge with information on benefit-sharing rules between the principal and agents might reduce this backfiring effect and even restore the effectiveness of the nudge. This hypothesis could be tested in a subsequent study. Another, more general, lesson is that even nudging policies should be evaluated before implementation at a larger scale. Nudges that fail can bring adverse consequences, like slowing down the agro-ecological transition of the wine cooperatives in our case, when they are ill-designed and used without precaution.

APPENDIX

Proof of Proposition 1: The manager's (hereafter she, or M) strategy is (x, t) , where $x(v)$ is the design of the proposal chosen by a manager with type v , and $t(v, x)$ is the transfer the manager offers at the last stage if the design is x , and the wine-grower has rejected the proposal. The wine-grower's (hereafter he, or W) strategy is (α, β) , where $\alpha(u, x) \in \{0, 1\}$ specifies whether W accepts or rejects the proposal when his type is u and the design is x , whilst $\beta(u, x, t) \in \{0, 1\}$ specifies whether W accepts or rejects the transfer t at the last stage of the game. Moreover, W revises his prior beliefs (m_1, m_2) into $(\mu(x), 1 - \mu(x))$ upon observing x . A Perfect Bayesian Equilibrium requires that these beliefs are obtained from the Bayes rule whenever possible, and that each player plays their best response, given these beliefs and the other player's strategy.

Assume such an equilibrium exists. Then β is easily determined: in an equilibrium, it must be that W accepts the transfer if and only if $t + u \geq 0$ (whether this inequality is strict or not does not matter, since the distribution of the wine-grower's types is continuous.) We can also determine α easily. Indeed, W accepts the proposal by comparing his payoff u to the expected payoff $Z(u, x)$ associated to waiting, in the hope that with probability δ a sufficiently high transfer will be offered:

$$(1) \quad Z(u, x) \equiv \delta E[\max(t(\tilde{v}, x) + u, 0)|x].$$

Note that the expectation is taken with respect to the revised distribution $\mu(x)$ of v , since the wine-grower has already observed x . Moreover, $\delta < 1$ implies that the difference $u - Z(u, x)$ is strictly increasing in u ; and since Z is at least zero, when $u < 0$ this difference is strictly negative. Therefore, there exists a unique threshold value $u^*(x) \geq 0$ such that W rejects the offer if $u < u^*(x)$, and accepts it otherwise. Finally, since $u^*(x) \geq 0$ and transfers are at least zero, using (1) the equality $u^*(x) = Z(u^*(x), x)$ becomes

$$(2) \quad u^*(x)(1 - \delta) = \delta E[t(\tilde{v}, x)|x].$$

Let us now turn to the transfer $t(v, x)$ chosen by M at the last stage. Since M knows that the remaining types are those with $u \leq u^*(x)$, when her payoff is v she chooses t to maximise:

$$b(t, u^*, v) \equiv (F(u^*) - F(-t))(v - t).$$

The "hazard rate" assumption ensures that this problem has a unique solution.⁷ We also learn that this solution is a function of x only through $u^*(x)$. Therefore, the equilibrium strategy can be written $t(v, x) = T(v, u^*(x))$, for some function T . Finally, since the derivative of b with respect to t is increasing with respect to v (the cross-derivative is $f(-t) > 0$), and decreasing with respect to u^* (the cross-derivative is $-f(u^*) < 0$), the solution $T(v, u^*)$ must be increasing with respect to v , and decreasing with respect to u^* .

⁷Indeed, when solving, we can focus on the interval $-u^* < t < v$, on which the manager gets a positive payoff. The derivative of this payoff with respect to t has the same sign as $-(F(u^*) - F(-t))/f(-t) + v - t$. The derivative of this expression is $-2 - (F(u^*) - F(-t))f'(-t)/f^2(-t)$. It is negative if $f' \geq 0$, and otherwise, it is less than $-2 + (F(-t) - 1)f'(-t)/f^2(-t)$, which is itself negative by the hazard rate assumption. This shows that the derivative of the payoff is positive, then negative, so that the maximum is unique.

We can now turn to the first stage, at which the manager chooses x to maximise:

$$-c(x) + (1 - F(u^*(x)))v + \delta \max_t b(t, u^*(x), v).$$

Thanks to the envelope theorem, the derivative of this expression with respect to $u^*(x)$ is:

$$(3) \quad -f(u^*)v + \delta f(u^*)(v - T(v, u^*)),$$

which is negative since $\delta < 1$, and decreases in v since $\delta < 1$ and T increases with v . Therefore, a manager with a higher v must choose a design associated to a lower acceptance threshold $u^*(x)$.

Suppose now that there exists a separating equilibrium in which v_1 chooses a design x_1 , and $v_2 > v_1$ chooses a design $x_2 \neq x_1$. From what we just said, it must be that $u^*(x_1) \geq u^*(x_2)$. Since $x_1 \neq x_2$, the wine-grower learns v by observing x , so that his revised beliefs are characterised by $\mu(x_1) = 1$ and $\mu(x_2) = 0$. We then apply (2) to get

$$u^*(x_1)(1-\delta) = \delta t(v_1, x_1) = \delta T(v_1, u^*(x_1)) \geq u^*(x_2)(1-\delta) = \delta t(v_2, x_2) = \delta T(v_2, u^*(x_2)).$$

But we have already shown that T increases with v and decreases with u^* . Thus, we have reached a contradiction: a separating equilibrium cannot exist.⁸

Proof of Proposition 2: We now know that all equilibria are pooling, in the sense that all types of the cooperative choose the same design x^* . Let us build one such equilibrium. On the equilibrium path, W does not learn anything from the proposal x^* , and the threshold $u^*(x^*)$ can be computed from (2):

$$u^*(x^*)(1 - \delta) = \delta \left[m_1 T(v_1, u^*(x^*)) + m_2 T(v_2, u^*(x^*)) \right].$$

In equilibrium, when her type is v M thus gets the payoff:

$$B^*(v) = -c(x^*) + (1 - F(u^*(x^*)))v + \delta \max_t b(t, u^*(x^*), v).$$

If M deviates to a different design $x \neq x^*$, we impose that the wine-grower interprets this deviation as signalling that the manager is of type v_2 . We obtain a new threshold $u^*(x)$, and (2) implies:

$$u^*(x)(1 - \delta) = \delta T(v_2, u^*(x)).$$

Because T is increasing in v and decreasing in u^* , we obtain $u^*(x) \geq u^*(x^*)$: the deviation thus makes immediate adoption less probable. Moreover, the manager's payoff becomes:

$$B(x, v) = -c(x) + (1 - F(u^*(x)))v + \delta \max_t b(t, u^*(x), v).$$

⁸Note that in the case $\delta = 0$, this inequality decreases to $u^*(x_1) = u^*(x_2) = 0$. Then one can only conclude that the manager must choose a design that minimises the cost $c(x)$, in accordance with the remark made in the definition of the timing of the game.

As already observed (see (3)), this expression is decreasing with respect to u^* , so that:

$$B(x, v) \leq -c(x) + (1 - F(u^*(x^*)))v + \delta \max_t b(t, u^*(x^*), v) = -c(x) + c(x^*) + B^*(v),$$

and therefore the deviation to x is unprofitable as soon as x^* minimises the cost c , whatever the type v of the manager. This concludes the proof.

REFERENCES

- ABADIE, A., S. ATHEY, G. W. IMBENS, AND J. M. WOOLDRIDGE (2023): “When Should You Adjust Standard Errors for Clustering?,” *The Quarterly Journal of Economics*, 138(1), 1–35.
- ALLCOTT, H., AND T. ROGERS (2014): “The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation,” *American Economic Review*, 104(10), 3003–3037.
- AYRES, I., S. RASEMAN, AND A. SHIH (2013): “Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage,” *The Journal of Law, Economics, and Organization*, 29(5), 992–1022.
- BARRIOS, T., R. DIAMOND, G. W. IMBENS, AND M. KOLESÁR (2012): “Clustering, Spatial Correlations, and Randomization Inference,” *Journal of the American Statistical Association*, 107(498), 578–591.
- BEHAVIORAL INSIGHTS TEAM (2013): “Applying Behavioural Insights to Organ Donation: preliminary results from a randomised controlled trial,” Discussion paper, UK Cabinet Office.
- BESHEARS, J., J. J. CHOI, D. LAIBSON, B. C. MADRIAN, AND K. L. MILKMAN (2015): “The Effect of Providing Peer Information on Retirement Savings Decisions,” *The Journal of Finance*, 70(3), 1161–1201.
- BHARGAVA, S., AND D. MANOLI (2015): “Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment,” *American Economic Review*, 105(11), 3489–3529.
- BOLTON, G., E. DIMANT, AND U. SCHMIDT (2020): “When a Nudge Backfires: Combining (Im)Plausible Deniability with Social and Economic Incentives to Promote Behavioral Change,” SSRN Scholarly Paper 3531419, Social Science Research Network, Rochester, NY.
- BREHM, J. W. (1966): *A theory of psychological reactance*, A theory of psychological reactance. Academic Press, Oxford, England, Pages: x, 135.
- BUGNI, F. A., I. A. CANAY, AND A. M. SHAIKH (2018): “Inference Under Covariate-Adaptive Randomization,” *Journal of the American Statistical Association*, 113(524), 1784–1796.
- (2019): “Inference under covariate-adaptive randomization with multiple treatments,” *Quantitative Economics*, 10(4), 1747–1785.
- CHABÉ-FERRET, S. (2023): *Statistical Tools for Causal Inference*. <https://chabefer.github.io/STCI/>.
- CHABÉ-FERRET, S., P. LE COENT, V. DAVID-LEGLEYE, AND V. DELANNOY (2023): “Non-Monetary Incentives to Increase Enrollment in Payments for Environmental Services,” *European Review of Agricultural Economics*, 50(4), 1401–1427.
- CHABÉ-FERRET, S., P. LE COENT, A. REYNAUD, J. SUBERVIE, AND D. LEPERCQ (2019): “Can we nudge farmers into saving water? Evidence from a randomised experiment,” *European Review of Agricultural Economics*, 46(3), 393–416.
- CLEE, M. A., AND R. A. WICKLUND (1980): “Consumer Behavior and Psychological Reactance,” *Journal of Consumer Research*, 6(4), 389–405.
- CROSON, R., AND N. TREICH (2014): “Behavioral Environmental Economics: Promises and Challenges,” *Environmental and Resource Economics*, 58(3), 335–351.
- DELBAC, L., M. DUPONT, P. MOREAU, AND D. THIERY (2013): “Surveying the leaf arthropod community in Medoc vineyards under mating disruption against the European grapevine moth,” *Integrated Protection and Production in Viticulture*, 85, 223–229.
- DELLAVIGNA, S., AND E. LINOS (2022): “RCTs to Scale: Comprehensive Evidence From Two Nudge Units,” *Econometrica*, 90(1), 81–116.
- DELMAS, M. A., M. FISCHLEIN, AND O. I. ASENSIO (2013): “Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012,” *Energy Policy*, 61, 729–739.
- DIAZ-BELTRAN, M., B. ALMANZA, K. BYRD, C. BEHNKE, AND D. NELSON (2023): “Visual Cues and Optimal Defaults in Fast-Food Combo Meals Benefit Health-Concerned Consumers-A Randomized Scenario-Based Experiment,” *Journal of the Academy of Nutrition and Dietetics*, 123(1), 52–64.e1.
- EARNHART, D., AND P. J. FERRARO (2021): “The Effect of Peer Comparisons on Polluters: A Randomized Field Experiment among Wastewater Dischargers,” *Environmental and Resource Economics*, 79(4), 627–652.
- FERRARO, P. J., J. J. MIRANDA, AND M. K. PRICE (2011): “The Persistence of Treatment Effects with Norm-Based Policy Instruments: Evidence from a Randomized Environmental Policy Experiment,” *American Economic Review*, 101(3), 318–322.
- FISCHER, C. (2008): “Feedback on household electricity consumption: a tool for saving energy?,” *Energy*

- Efficiency*, 1(1), 79–104.
- GOSNELL, G. K., J. A. LIST, AND R. METCALFE (2016): “A New Approach to an Age-Old Problem: Solving Externalities by Incenting Workers Directly,” Working Paper 22316, National Bureau of Economic Research.
- HOFFMANN, C., AND D. THIERY (2010): “Mating disruption for the control of grape berry moths – Bottlenecks and conditions of adoption in different European vine growing regions.,” Grapevine case study number 3., Inra.
- KUHFUSS, L., R. PRÉGET, S. THOYER, N. HANLEY, P. L. COENT, AND M. DÉSolÉ (2016): “Nudges, Social Norms, and Permanence in Agri-environmental Schemes,” *Land Economics*, 92(4), 641–655.
- LE COENT, P., R. PRÉGET, AND S. THOYER (2021): “Farmers Follow the Herd: A Theoretical Model on Social Norms and Payments for Environmental Services,” *Environmental and Resource Economics*, 78(2), 287–306.
- MAIER, M., F. BARTOS, T. D. STANLEY, D. R. SHANKS, A. J. L. HARRIS, AND E.-J. WAGENMAKERS (2022): “No evidence for nudging after adjusting for publication bias,” *Proceedings of the National Academy of Sciences*, 119(31), e2200300119, Publisher: Proceedings of the National Academy of Sciences.
- MERTENS, S., M. HERBERZ, U. J. J. HAHNEL, AND T. BROSCHE (2022): “The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains,” *Proceedings of the National Academy of Sciences*, 119(1), e2107346118.
- NEMATI, M., AND J. PENN (2020): “The impact of information-based interventions on conservation behavior: A meta-analysis,” *Resource and Energy Economics*, 62, 101201.
- OSMAN, M. (2020): “Backfiring, reactance, boomerang, spillovers, and rebound effects: Can we learn anything from examples where nudges do the opposite of what they intended?,” .
- RICHTER, I., J. THOGERSEN, AND C. A. KLOCKNER (2018): “A Social Norms Intervention Going Wrong: Boomerang Effects from Descriptive Norms Information,” *Sustainability*, 10(8), 2848.
- SCHUBERT, C. (2017): “Green nudges: Do they work? Are they ethical?,” *Ecological Economics*, 132, 329–342.
- SCHULTZ, P. W., J. M. NOLAN, R. B. CIALDINI, N. J. GOLDSTEIN, AND V. GRISKEVICIUS (2007): “The Constructive, Destructive, and Reconstructive Power of Social Norms,” *Psychological Science*, 18(5), 429.
- STUBER, J. M., J. LAKERVELD, L. W. KIEVITSBOSCH, J. D. MACKENBACH, AND J. W. J. BEULENS (2022): “Nudging customers towards healthier food and beverage purchases in a real-life online supermarket: a multi-arm randomized controlled trial,” *BMC Medicine*, 20(1), 10.
- SUNSTEIN, C. R. (2017): “Nudges that fail,” *Behavioural Public Policy*, 1(1), 4–25.
- THALER, R. H., AND C. R. SUNSTEIN (2009): *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Penguin Books, New York, updated edn.
- THUNSTROM, L., B. GILBERT, AND C. J. RITTEN (2018): “Nudges that hurt those already hurting - distributional and unintended effects of salience nudges,” *Journal of Economic Behavior & Organization*, 153, 267–282.