Free-riding on product quality in cooperatives: Lessons from an experiment

Olivier Bonroy, Alexis Garapin, Stephen F. Hamilton and Diogo M. Souza Monteiro*

Abstract

An important issue facing marketing cooperatives is that the overall quality of the product depends on the quality of farm products provided by individual members. We conduct an experiment to empirically investigate producer incentives to free-ride on quality among members of a marketing cooperative in a setting where the average quality provided by members of the cooperative results in a collective rent that is distributed back to members in a patronage dividend levied in proportion to the quantity produced. Hidden actions by cooperative members that impact quality are imperfectly monitored, but free-riding, when detected, results in exclusion from cooperative returns. The randomized payoff structure of our game results in a novel experimental design that nests public good games and multi-player assurance games. Our findings indicate that free-riding on product quality is deterred when: (i) cooperatives base patronage dividends on quality outcomes of smaller groups; (ii) payoffs from free-riding

*Olivier Bonroy is tenured research fellow at INRA (French National Institute for Agricultural Research) and at the GAEL, Université Grenoble Alpes, Grenoble, France. Alexis Garapin is professor in the GAEL at the Université Grenoble Alpes, Grenoble, France. Stephen F. Hamilton is professor in the Orfalea College of Business at California Polytechnic State University, San Luis Obispo. Diogo M. Souza Monteiro is senior lecturer in the School of Agriculture, Food and Rural Development, at the Newcastle University, UK. The authors gratefully acknowledge the financial support of the Université Grenoble Alpes. The authors are grateful for detailed comments and suggestions from the editor, three anonymous referees, Edward Cartwright, Paolo Crosetto, Jean-Loup Dupuis, Adelaide Fadhuile, Luca Panzone, Richard Sexton, Marie Claire Villeval and participants of various seminars and meetings, including AAEA Annual Meeting 2015, Dyson School of Applied Economics and Management at Cornell University 2016, JMA Meeting 2016, ASSET Annual Meeting 2016 and the Behavioral and Experimental Northeast Cluster (BENC) Seminar at Durham University Business School 2017. Responsibility for any errors and omissions is entirely ours. Correspondence to be sent to: alexis.garapin@univ-grenoble-alpes.fr.

are randomized by the possibility of exclusion from cooperative returns; and (iii) co-operatives distribute a larger share of returns to members through indirect payments such as capital pooling and cost sharing arrangements unrelated to product quality.

Key words: Free-riding, cooperative, experimental analysis, collective action game, probability of exclusion.

JEL codes: Q13, L66, C92, D82

Agricultural marketing cooperatives often promote premium product lines that emphasize high-quality farm products. This focus on product quality has long been an objective among marketing cooperatives in the wine industry, and is now emerging in other agricultural products sectors such as cheese and fruit and vegetable products.[1] As cooperative organizations continue to promote high-quality farm products, managing incentives to mitigate problems among agents becomes increasingly important, as the return to individual effort depends on the coordinated actions of all members of the cooperative to provide high-quality products. For example, rents returned to members of a winemaking cooperative depend on the quality of grapes contributed by individual vineyards.[2] Providing incentives for high-quality production is difficult when members face heterogeneous environmental conditions or have different levels of skill, and this is particularly true when actions of individual producers are unobserved (Mérel, Saitone, and Sexton, 2015).

It has been well-known since Hölmstrom (1979) that moral hazard results in suboptimal performance when collective rents depend on the hidden actions of individuals. The resulting incentive problem in cooperative organizations is poorly remedied by grading standards, because a minimum quality standard is only capable of controlling product quality at the lower tail of the quality distribution. In the case of a winemaking cooperative, for example, individual producers can bundle low- and high-quality grapes at the minimum quality threshold, while diverting high-quality grapes towards their own, branded labels (see e.g. Planas, 2016).

Compared with investor-owned firms, members of marketing cooperatives have several incentives to free-ride on product quality (see Pennerstorfer and Weiss, 2013). First, because aggregating production economizes on harvesting cost, members of a cooperative may not exert sufficient effort to segregate high-and low-quality products prior to grading. Second, when idiosyncratic shocks reduce product quality on some acres but not on others, farmers may choose to harvest all acreage to increase the share of output that qualifies for the patronage dividend. Because cooperative rents are determined jointly by the collective effort of all members, such actions harm performance of the cooperative by diluting the overall quality of the product mix. This article examines incentives for members of marketing cooperatives to contribute high quality products in settings where individual effort is unobserved.

We frame our analysis in the context of a multi-player, collective action experiment. Individuals choose whether or not to contribute costly private action in return for a share of a collective rent that depends on the joint contribution levels of all cooperative members. Specifically, collective rents are distributed back to members through a patronage dividend levied in proportion to the *quantity* of agricultural products contributed by each member, irrespective of differences in product *quality* among members. Individuals choose whether or not to incur additional cost to produce high-quality goods by undertaking an unobserved action; however, members of the cooperative that choose not to engage in the costly action face a random probability of detection, as would be the case for a cooperative that engages in random sampling. We follow the literature on social sanctions (Tirole, 1996) by modeling the penalty for detection as exclusion from the cooperative organization.

Our research relates to the experimental literature that considers exclusion in public goods settings (see e.g. De Geest, Stranlund, and Spraggon, 2017; Maier-Rigaud, Martinsson, and Staffiero, 2010; Cinyabuguma, Page, and Putterman, 2005); however, we depart from this literature by considering exclusion in the form of social sanctions as opposed to sanctions by peers.[3] Our work builds on the strand of the experimental literature following Swope (2002) and Croson et al. (2015) that considers exclusion through social sanctions, although the context of their work, which considers public good provision through voluntary contribution mechanism (VCM), is quite different. In Swope (2002), exclusion is applied to individuals who fail to meet a minimum contribution requirement, while in Croson et al. (2015) only the smallest contributor is excluded. Our research relates closely to this work in the sense that the probability of exclusion through social sanctions turns a pure public good game into a coordination game with multiple Nash equilibria. A novel aspect of our approach is that the coordination problem is driven by randomized exclusion of individuals who do not contribute.

The structure of our game introduces a novel elements of coordination between players that provides insights into individual behavior in cooperative settings. Specifically, our experimental framework nests several classes of multi-player collective action games according to the randomized returns to free-riding. It reduces to a public good game in treatments where payoffs are non-random, and to a multi-player assurance game (the so-called "stag hunt") under randomized payoffs from shirking. In groups with a "large" number of players, two Nash equilibria emerge: (*i*) one in which all players incur private costs to facilitate collective returns; and (*ii*) one in which all players free-ride. We believe the assurance game is particularly relevant to product quality outcomes in marketing cooperatives, because it maintains the challenge of facilitating group cooperation while introducing coordination issues among members.[4,5]

Our main findings are as follows. First, we extend the outcome found in experimental literature that unobserved effort declines with group size to settings with random probability of detection. Holding constant the marginal return from cooperating, previous studies have found cooperation rates to be lower in larger groups (see Kollock, 1998 for a review). Our analysis extends this result to settings with imperfect monitoring. When group size increases from 4 players to 8 players, collective rents decline as individuals increasingly shirk on effort to free-ride on the contributions of others. This finding has novel implications for eliciting effort to produce high-quality farm products in marketing cooperatives. For example, in the case of profit-sharing among members of a cooperative, greater effort among members can be facilitated by indexing compensation to the performance of individual divisions within a cooperative, rather than to aggregate performance across many divisions, for instance by segmenting patronage dividends in a wine cooperative according to the financial performance of individual varietals.

Second, we empirically investigate how randomized payoffs from shirking on quality facilitate unobserved effort by cooperative members. For our "large group" treatment with 8 players, we demonstrate that introducing an exclusion probability in the collective action game significantly induces individuals' propensity to engage in costly, unobserved effort. Whereas previous experiments on imperfect monitoring designs with sanctions by peers (Ambrus and Greiner, 2012; Xiao and Kunreuther, 2016) report mixed results on the ability of monitoring to deter shirking, our results provide striking evidence that designs based on social sanctions in a cooperative substantially deter free-riding.[6]

Third, to our knowledge, our study is the first to examine how variation in the magnitude of services provided by the cooperatives to individual members affects collective effort.[7] We show that services provided by a marketing cooperative that are independent of the product quality objective result in enhanced effort among members, even when the coordinated task generates relatively small collective returns. Providing services to assist growers is an important aspect of wine cooperatives, which often help growers promote desirable grape characteristics and adopt cultural practices that improve quality (Fares and Orozco, 2014, Goodhue et al., 2003). Our findings demonstrate that services provided to cooperative members that are independent of patronage refunds, such as capital pooling and cost sharing for technical expertise, offer an important and heretofore unrecognized mechanism to enhance the quality of goods sold in cooperatives.

The remainder of the article is structured as follows. In the next section we describe how our work relates with the existing literature. In section 3 we present the theoretical framework. In section 4, we describe the experimental design and procedure. In section 5 we present our experimental results, and in section 6 we conclude with final remarks.

Background

Our work contributes to the agricultural economics literature on product quality. A common objective of cooperative organizations is to promote producer welfare through collective action, particularly as relayed to product quality (see Saitone and Sexton, 2010). Winfree and McCluskey (2005), Fishman et al. (2015) and Castriota and Delmastro (2015) examine settings in which individual farmers sell products under a collective label or a collective brand. As in our framework, the collective rent an individual producer receives depends on the joint effort of all producers, typically in the form of a price premium based on average quality. We depart from this setting by considering individual contributions to be unobservable and by analyzing how incentives introduced by cooperatives serve to deter free riding behavior. An important literature exists on the free riding on quantity by cooperative's members, and a much smaller one exists on the free riding on quality.[8] Pennerstorfer and Weiss (2013) investigate free riding on quality (and quantity) by comparing incentives in cooperatives with investor-owned companies. As in our study, the cooperative aggregates the output of members into a composite product; however, an essential difference in our setting is that members' coordination on quality is endogenous.[9]

Our analysis is also related to the experimental literature that addresses cooperation (or contribution) in social dilemmas (e.g. prisoner's dilemma, public good or common pool resource games) using various monitoring schemes based on sanctions by peers. This literature largely investigates the topic under conditions of monitoring (see e.g. Fehr and Gächter, 2000; Maier-Rigaud, Martinsson, and Staffiero, 2010), although a number of papers have considered imperfect monitoring (Ambrus and Greiner, 2012; Xiao and Kunreuther, 2016).[10] The monitoring mechanisms considered include costly punishment (Fehr and Gächter, 2000), exclusion of the lowest contributor, either from the benefit (De Geest, Stranlund, and Spraggon, 2017) or from future interactions (Maier-Rigaud, Martinsson, and Staffiero, 2010) and noisy interactions on the opponent's moves or payoffs (see e.g. Bereby-Meyer and Roth, 2006). Our analysis departs from this literature by modeling the monitoring mechanism as a social sanction, which we believe is more relevant to understanding product quality contributions in marketing cooperatives than outcomes based on laboratory experiments that rely on sanctions by peers.

The effect of social sanctions has been investigated in the experimental literature on agency problems (see e.g. Nosenzo et al., 2016); however, this literature does not consider collective rents from costly choices among individuals. Thus, our research combines elements from two distinct strands of the literature. The effect of social sanctions in collective action games has received relatively little attention in the experimental literature (see Swope, 2002 and Croson et al., 2015), and the few studies that consider social sanctions do so under perfect monitoring. We depart from this line of inquiry by considering: (i) the effects of imperfect monitoring regimes characterized by randomized payoffs from shirking, and (ii) variations in the base payoff level through member services. These novel modeling elements provide a realistic setting for examining incentive problems for producing high-quality products in marketing cooperatives.

The theoretical framework

Consider a group of N members of a marketing cooperative who seek to gain collective rents from high quality products. Each member can produce at most one unit of the product and simultaneously chooses between one of two actions, a high-effort activity (action A) and a low-effort activity (action B). These actions have corresponding costs c_A and c_B, respectively, where we assume that $c_A > c_B$.

Together, the collective effort of the cooperative members produces an overall quality level, $\rho(N_A) = N_A/N$, where N_A denotes the number of players selecting action A. Overall quality is an increasing function of collective effort, but the effort level of each individual member is not revealed to the cooperative other than by random chance (as discussed below).

The cooperative's product is purchased by a group of M consumers ($M \geq N$). Each consumer chooses to purchase either one unit of the good supplied by the cooperative at price α, one unit of the standard good in a competitive market at price $\beta = c_B$, or no good at all. We assume that consumers evaluate the average quality $\rho(N_A)$ of the product as in Fulton and Giannakas, 2004 and Hamilton and Zilberman, 2006, and are willing to pay $V + \rho(N_A)a$ for products of such quality, where V refers to the utility derived from consuming a standard good and a refers to consumers' taste for quality, as in a Mussa and Rosen-type utility function. Consumer utility is given by:

$$(1) \quad \begin{cases} U = V + \rho(N_A)a - \alpha & \text{when the consumer buys the cooperative's product,} \\ U = V - \beta & \text{when the consumer buys the standard product,} \\ U = 0 & \text{otherwise.} \end{cases}$$

The cooperative sets α at the reservation price of consumers, $\beta + \rho(N_A)a$,[11] and distributes sales revenue to members through patronage dividends. Notice that each member receives $\rho(N_A)a$, i.e. the patronage dividend depending on the realized quality of the good produced by the cooperative, which, in turn, is determined by the collective effort of all members. Cooperative members also receive a base payoff $p = \beta + S$ independent of the collective returns from product quality, with S represent various services in addition to patronage dividends, for instance by pooling resources for capital equipment and sharing technical expertise.[12] We model the services provided by the cooperative as a base payoff, that is independent of the collective returns from product quality. Therefore, choosing action A generates both a base payoff, p, and an additional product quality rent, $\rho(N_A)a$,[13] which depends on the coordinated action of all members.[14] The collective rent is strictly increasing in the share of players engaged in the high effort activity.

Without loss of generality, we set $\beta = c_B = 0$, such as the base payoff is only given by the value of cooperative services ($p = S$). Moreover, to guarantee positive demand when all members choose action A (i.e. $N_A = N$) in a cooperative that prices at marginal cost c_A, we assume the mild regularity condition that the collective rent exceeds the cost of effort, $a > c_A$. If this was not true, consumer net utility from purchasing the high-quality good at marginal cost would be $V + a - c_A < V$, and a high-quality goods market could not exist.

Both the base payoff and collective rent are received by all cooperative members irrespective of their individual effort, except under circumstances of exclusion. We assume that members choosing action B that attempt to sell their product to a high-quality cooperative face a random probability of detection. If a member is "caught" engaging in action B, his output is excluded from the cooperative's production (market A) and his gain is then given by the sale of his output on the competitive market at price c_B.[15] Formally, our model is directly analogous to a collective action game for a club good, which is consistent with the interpretation of a marketing cooperative as solving a mechanism design problem for provision of a club good (see Hueth, 2014).

Notice, by design, that choosing action B involves identical payoffs to choosing action A with two exceptions: (*i*) action B involves no cost ($c_B = 0$), which makes shirking attractive; and (*ii*) action B, if detected, leads to exclusion from the group returns in the higher-yielding market for products produced by action A. A member randomly detected in action B also suffers a loss of services p provided by the cooperative.

The incentive structure of the model can be seen by considering the value of the collective rent. As more members choose action A, the payoff to each member of the cooperative increases. Specifically, each member choosing A receives the following return:

$$\Gamma_A(N_A) = p + \rho(N_A)\, a - c_A. \tag{2}$$

It follows immediately, absent exclusion, that the individual payoff for choosing action

B is larger for any level of collective effort, because the low-effort activity entails no cost. Absent a probability of exclusion, the payoff to a player choosing action B is

$$\Gamma_B(N_A) = p + \rho(N_A)\, a > \Gamma_A(N_A). \tag{3}$$

Notice that the payoff from both action A and action B decreases as more members engage in action B, as this behavior reduces the collective rent, $\rho(N_A)a$. Accordingly, the decision to choose A or B for each member depends on the decisions of the other $(N-1)$ members to contribute effort. To see this, consider a "late-arriving" member who can decide on an action after observing the actions selected by all other members.

Suppose a late-arriving member observes m other players selecting action A. If the member also chooses A, then number of members coordinating on action A is $m+1$ and the payoff from action A will be $\Gamma_A(m+1)$. Conversely, if the late-arriving member chooses action B, the payoff is given by $\Gamma_B(m)$. Therefore, the payoff maximizing choice of a late-arriving member is driven by the comparison between these two payoffs: A payoff-maximizer would opt for action A when $\Gamma_A(m+1) > \Gamma_B(m)$, and otherwise opt for action B.

To make this example concrete, consider a two-player game in which $m = 1$. In this case, the late-arriving player that observes his group mate playing action A faces the payoffs $\Gamma_A(2) = p + a - c_A$ and $\Gamma_B(1) = p + \frac{a}{2}$, so that the payoff maximizing choice is action A whenever $a > 2c_A$.

Our framework encompasses the essential coordination problem in collective action games. By construction, the welfare of the collective is always higher when all members choose action A, because the collective rent exceeds the cost of effort, $a > c_A$; however, this action may work against the self-interest of individuals. In the 2-player case described above, for example, a payoff maximizer would wish to defect to action B whenever $2c_A > a > c_A$, even though such behavior works against the social interest.

In general, $\Gamma_A(m+1) > \Gamma_B(m)$, $\forall m \in [0, N-1]$ is satisfied when the number of players is sufficiently small, $N < \frac{a}{c_A}$, such that choosing action A is the dominant strategy irrespective of what the other $(N-1)$ players are doing. Conversely, $\Gamma_A(m+1) < \Gamma_B(m)$, $\forall m \in [0, N-1]$ holds when the number of players is sufficiently large, $N > \frac{a}{c_A}$, in which case action B becomes the dominant strategy of all players. In the "large group" case, the framework reduces to a prisoners' dilemma game with a one-shot Nash equilibrium involving coordination on action B. Thus, our framework nests two important classes of economic games depending on the number of players, N, and the relative return to effort, $\frac{a}{c_A}$.

In practice, social sanctions and legal institutions often exist that limit individuals from engaging in actions that work against the social interest. For example, workers caught shirking on the job can be terminated and firms selling mislabeled products at premium prices may be liable to compensate consumers for harms. In the present setting, such sanctions among cooperative members exist when a cooperative member caught providing quality below a minimum standard has his output excluded from production. In general, enforcing against such behavior is problematic, because individuals that shirk are difficult to catch.

We denote the detection frequency for individuals engaged in the low-effort action with the random variable, θ. Following Tirole (1996), we interpret θ as the probability of exclusion from the collective rent.

When members engaged in action B are faced with a probability of exclusion, the payoff from action B is 0 with probability θ and $\Gamma_B(N_A) > 0$ with probability $1-\theta$.

The final decision of a player is driven by the comparison of the following payoffs:

$$\begin{cases} \Gamma_A(m+1) = p + \rho(m+1)\,a - c_A & \text{when the member chooses } A \\ \Gamma_B(m) = (1-\theta)\,[p + \rho(m)\,a] & \text{when the member chooses } B. \end{cases} \tag{4}$$

As in the case with non-randomized payoffs, $\Gamma_A(m+1) > \Gamma_B(m)$ and action A is the dominant strategy when $N < \frac{a}{c_A}$, and this remains true for any permissible values of m and θ. Under randomized payoffs, action A can remain the dominant strategy for "large" groups, $N > \frac{a}{c_A}$, but only in cases where the probability of exclusion is sufficiently large. Specifically, the framework reduces to a prisoners' dilemma game in which action B is the dominant strategy when $\theta < \theta_1(N) = \frac{c_A\left(N-\frac{a}{c_A}\right)}{(a+p)N-a}$, whereas contributing effort towards collective returns in action A is the dominant strategy when $\theta > \theta_2(N) = \frac{c_A\left(N-\frac{a}{c_A}\right)}{pN}$.

Randomized returns to choosing action B introduces a new and third class of game. When $N > \frac{a}{c_A}$ and $\theta \in [\theta_1(N), \theta_2(N)]$ the model has multiple equilibria: the socially optimal outcome and the prisoner's dilemma outcome.[16] The framework reduces to a multi-player assurance game (a "stag hunt") with two pure-strategy equilibria. Specifically, letting $\tilde{n}(N) \equiv \frac{N(c_A-\theta p)-a}{a\theta}$, the dominant strategy is action B for $N_A \in [0, \tilde{n}(N)]$, whereas the dominant strategy is action A for $N_A \in [\tilde{n}(N), N-1]$ (see Dixit, Skeath, and Reiley, 2009). In our symmetric payoff game, a Nash equilibrium exists for any value of $N_A > \tilde{n}(N)$ in which all players contribute, $N_A = N$. Conversely, for any value of $N_A < \tilde{n}(N)$, a Nash equilibrium exists in which all players choose action B, resulting in zero contributions to the collective good, $N_A = 0$.[17]

Our observations above on the framework can be summarized as follows:

Lemma 1 *When $N < \frac{a}{c_A}$ and $\theta \in [0,1]$ the one-shot Nash equilibrium involves all members selecting action A. When $N > \frac{a}{c_A}$, three equilibrium configurations emerge: (i) for $\forall\theta \in [0, \theta_1(N)]$ the one-shot Nash equilibrium involves all members selecting action B; (ii) for $\forall\theta \in [\theta_1(N), \theta_2(N)]$ multiple Nash equilibria emerge in a multi-player assurance game over actions A and B; and (iii) for $\forall\theta \in [\theta_2(N), 1]$ the one-shot Nash equilibrium involves all members choosing action A.*

It is possible to further elaborate on the payoff-maximizing equilibrium outcomes in the case of a multi-player assurance game. Given $N > \frac{a}{c_A}$, we have $\frac{\partial\theta_1(N)}{\partial p} = \frac{-Nc_A\left(N-\frac{a}{c_A}\right)}{((a+p)N-a)^2} < 0$ and $\frac{\partial\theta_2(N)}{\partial p} = \frac{-c_A\left(N-\frac{a}{c_A}\right)}{Np^2} < 0$, with $\mid \frac{\partial\theta_1(N)}{\partial p} \mid < \mid \frac{\partial\theta_2(N)}{\partial p} \mid$. Thus, we arrive at:

Lemma 2 *When $N > \frac{a}{c_A}$, increasing the base payoff, p: (i) increases the range of θ where the social optimum is achieved in the one-shot Nash equilibrium; (ii) reduces the range of θ where the framework reduces to a prisoners' dilemma; and (iii) reduces the range of θ that results in a multi-player assurance game.*

Experimental design

To empirically test the predicted outcomes of our model, we employ a between-subjects experimental design. In this section, we describe our treatments and derive equilibrium predictions of the model.

Experimental treatments and parameters

Our experiment presents subjects with various forms of the model detailed above. Subjects were told that they had to choose between action A or action B, and that all members of their group had to make a simultaneous choice over a problem identical to the one they were facing. Subjects were informed of the number of subjects in their group, but not of the identities of group members, and communication with other subjects during the sessions was not allowed. Subjects were informed that the composition of their group would remain fixed over all periods of the session. Each subject played the game for 20 periods and the number of periods to be played was common information.[18]

We control for subjects' attitudes towards risk using the Holt and Laury (2002) risk tests. We administer the test before and after the collective action game to control possible order effects.[19]

Subjects were given a payoff table and encouraged to use them before making their choices. This table showed the payoff for the period when choosing A or B, according to the number other members in their group deciding to play A or B. To build payoff tables for the experiment, we used the following parameter values: $p = c_A = 2$, $p = 2$, $c_A = 4$, $c_B = 0$, $a = 10$ such that:

* for $N = 4$ and $p = 2$, only one Nash equilibrium exists in which all players adopt action A;

* for $N = 8$ and $p = 2$, multiple equilibria exist with an equilibrium partition defined by $\theta_1 = 0.07$ and $\theta_2 = 0.375$.[20]

* for $N = 8$ and $p = 4$, multiple equilibria exist with an equilibrium partition defined by $\theta_1 = 0.059$ and $\theta_2 = 0.187$.[21]

To test the effects of the group size, the exclusion probability and the fixed payoff on equilibrium behavior, we vary the values of N, θ and p across treatments. Thus, a treatment is defined by a triplet [N; θ; p]. For example, the treatment [4; 0.25; 2] corresponds to experimental sessions conducted with groups of 4 subjects, with an exclusion probability of $\theta = 0.25$ and a fixed payoff $p = 2$.

We employ a between-subjects procedure. Because Lemma 1 suggests that both the number of subjects and the value of θ affect the optimal coordination strategy of players, we design 8 experimental treatments with p = 2 that varies the size N of the group, $N = [4; 8]$, and the exclusion probability, $\theta = [0; 0.05; 0.25; 0.45]$, with values that span all the equilibrium partitions of the $N = 8$ subject game. Because Lemma 2 suggests that increasing the fixed payoff p modifies the equilibrium partition for N sufficiently large, we design 2 more experimental treatments with $p = 4$ and $N = 8$ only: one with $\theta = 0.25$ and the other one with $\theta = 0.45$. We consider treatments [4; 0; 2] and [8; 0; 2] as "benchmark" treatments.

The payoff tables presented to the subjects (see examples in the appendix) display all the possible payoffs for every combination of own and others' choices.[22] In treatments in which the payoff to action B was randomized, the payoff table displayed that the choice of B could lead to a zero payoff with probability θ. For example, in treatments [8; 0.25; 2] or [8; 0.25; 4], subjects were told that if they chose B they had a 0.75 chance of receiving a positive payoff and a 0.25 chance of receiving no payoff at all.[23] To avoid introducing biases in subjects' risk preferences, the resulting expected payoff for action B was not calculated and displayed in the payoff table.

The subject's payoff from the experiment was comprised of the cumulative payoff over 20 periods. Each subject was credited each period with a payoff in "Ecus" corresponding to the frequency of action A choices among members of her group (including her own choice). The "Ecu" is the currency used for the game in the experiment, which was settled at the end of the experiment at an exchange rate of 1 Euro = 1,250 Ecus. After each period, each subject was informed of her decision, the frequency of A choices in her group (including her own choice), her own payoff, and her cumulative payoff since the beginning of the experiment. Payoffs in each period were independent of choices made in previous periods, although each subject could see the entire history of choices and payoffs in previous periods throughout the experiment.

Predictions

Based on both Lemmas 1 and 2 the equilibria predictions of the game played are:

Prediction 1 ***The effect of N***

An increase in the number of players reduces the frequency of selecting action A. Specifically: (i) For $N = 4$ a payoff-maximizing player would choose action A irrespective of the level of θ; and (ii) for $N = 8$ a payoff-maximizing player would choose action B (action A) whenever $\theta \leq \theta_1 = 0.07$ ($\theta \geq \theta_2 = 0.375$). For intermediate ranges of $\theta_1 < \theta = 0.25 < \theta_2$, either equilibrium configuration is possible.

For the $N = 4$ treatment, parameter values are selected such that self-interested individuals choose action A irrespective of the level of θ. For the $N = 8$ treatment, individual payoffs are expected to be higher under action B that action A when $\theta = 0$ and $\theta = 0.05$. Thus, we postulate that increasing the number of players reduces the incentive to choose action A.

Prediction 2 ***The effect of θ***

For intermediate values of θ, players are confronted with a coordination problem in the equilibrium choices that modifies their incentive to choose action A. For $N = 8$, $\theta = 0.25$ and $p = 2$ the collective action game reduces to a multi-player assurance game in which multiple Nash equilibrium exist.

For intermediate values of θ, the players face a coordination problem in attaining an equilibrium, as the game structure becomes a multi-player assurance game. Lemma 2 establishes that increasing p increases the region of θ where the social optimum is achieved. In order to test this result a treatment [8; 0.25; 4] is conducted by considering $p = 2$ and $c_A = 4$. For such a value of p the partition of equilibrium is defined by $\theta_1 = 0.059$ and $\theta_2 = 0.187$. Thus we have:

Prediction 3 ***The effect of p***

Increasing the base payoff p increases the proportion of players choosing A. For $N = 8$, $\theta = 0.25$ and $p = 4$ all players would participate in action A; however, for $N = 8$, $\theta = 0.25$ and $p = 2$ the coordination problem remains.

This prediction postulates that increasing the base payoff p from belonging to a group helps players avoid the coordination problem in attaining the socially optimal equilibrium, thereby reducing the incentive to free ride. Figure 1. describes this outcome. Increasing the base payoff facilitates coordination on action A, even though p is independent of the magnitude of the collective rent arising from action A, by squeezing together the boundaries defined by θ_1 and θ_2. Raising the base payment provided to players makes social sanctions more salient, thereby facilitating coordinated actions.

Experimental procedure

When subjects arrived in the laboratory, they received a personal code to preserve their anonymity and were randomly assigned to a computer station. The laboratory consisted of 16 working stations and each session involved anonymous matching of subjects into either four groups of 4 players or two groups of 8 players. Each session of the experiment corresponded to a single treatment and we adopt a between subjects design in which subjects participated in a singular treatment with fixed group composition. Subjects were informed that the composition of their group would remain fixed over all periods of the session and that no interaction between the group members would be permitted.

Each subject was provided with an envelope at their station containing a show-up fee of 5 euros. Before the actual experiment started, the experimenter read the instructions aloud to the subjects. In addition, subjects were able to read these instructions on their individual screen, which clearly identified that the instructions were identical for all participants.

To ensure that the subjects understood the instructions, they were asked to complete a questionnaire to assess their understanding of the game and the meaning of the variables, profit calculations, etc. The questionnaires were then marked by and along with the experimenter before the experiment started. Participants had complete information on their own and others' payoffs among members of their group and were provided with payoff tables corresponding to the ones depicted in the previous section. The subjects were informed that the game would be repeated exactly 20 times.

To ensure non-hypothetical decisions, the subjects were informed at the beginning of each session that they would anonymously be paid cash at the end of the experiment in an amount that depended on the decisions they made and on the decisions made by others. Subjects could earn between 8.4 and 28.7 Euros including the show up fee, their payoff for the collective action game,[24] and their payoff for the risk elicitation task.[25]

At the end of the experiment, subjects filled in a small questionnaire asking them basic demographic information. At this point, the subjects could see their total payoff from the experiment. Finally, they were called one by one in a separate room to receive privately their money in cash, after which they were free to leave the lab.[26] Each session lasted approximately 75 minutes, including time devoted to the subjects' payment.

To test predictions 1 (effect of N) and 2 (effect of θ), we conducted 16 experimental sessions from February 2014 to May 2014, with a total of 244 subjects (Female: 155, Male: 89). To test the prediction 3 (effect of p, with $p = 2c_A = 4$), which concerns only subjects in the $N = 8$ treatments, we conducted an additional set of 8 experimental sessions over March-April 2015 with a total of 128 subjects (Female: 89, Male: 39). In this new set of experiments, we performed treatments with $\theta = 0.25$ but also with $\theta = 0.45$.[27] Subjects were undergraduates from different universities (arts, sciences, social sciences, engineering schools) with no background in game theory.

The entries in Table 1 provide data about each session and the number of observations in our sample. Because we rely on fixed partner matching, each group can be considered as a statistically independent observation. The experiment was designed to provide at least 4 observations per treatment.

Results

In this section we present the results of our experiment. We start with an overview of our data, where we provide initial evidence supporting our predictions. We proceed with the results of non-parametric tests on our treatment groups and, finally, we present and discuss the results of our econometric estimations.

Descriptive results

Our theoretical model and our experimental procedure suggests three clear predictions. We expect to observe: (*i*) an increase in the level of free riding or shirking (i.e., choosing B) with an increase in the number of group members; (*ii*) greater cooperation (choosing A) with an increase in the exclusion probability; and (*iii*) a reduction of the level of free riding with larger base payoffs.

Figure 2 depicts the mean frequency of free riding behavior (i.e., action B choices) across periods for sessions with $N = 4$ and $N = 8$ players. As expected, the levels of free riding are "small" (below 10%) in the treatments with small groups, whereas for the large group treatments, the frequency of action B choices is slightly above 30%. This provides preliminary evidence for our first prediction on the effect of group size.

It should be noted that our experimental design does not enable us to disentangle changes in cooperative behavior from changes in group reputation effects among subjects in the experiment as group size increases. If subjects differ in their willingness to cooperate, then it is possible that subjects develop reputations for cooperation within groups that are easier to identify in smaller groups. Thus, an alternative explanation for the coordination problem when N gets larger is that small groups better facilitate reputation effects between subjects in the experiment.

Turning to our results on the probability of exclusion, Figure 3 shows the effects of θ on the levels of free riding for the sub-sample of our data with 8 players involved in the game and with the lower level of base payoff. Notice that when we introduce an exclusion probability, the level of free riding increases, but then decreases to levels identical to the case where $N = 4$ when there is a 45% probability of exclusion from the rent. It is clear by inspection that the mean free riding level is reduced almost in half as the exclusion probability increases to 45%. Figure 3. thus provides initial support for *Prediction 2* that purely self-interested players will select action A whenever $\theta \geq \theta_2 = 0.375$.

Finally, we inspect the effects of a higher base payoff on the level of free riding on the collective rent. This is reported in Figure 4, where we plot the effect of doubling on the base payoff, on mean free ride when there is a 25% probability of exclusion from the rent. The figure suggests that to an increase in base payoff corresponds a decrease in shirking, which is in line with our theoretical prediction and preliminary evidence supporting our *Prediction 3*

Treatment effects

We conduct non-parametric tests on our different treatments. Specifically, we examine whether there are significant differences between the group size and random payoff treatments using both the Kruskall-Wallis equality of population rank test and Mann-Whitney test. Comparing the choice of action B across the number of subjects and the different levels of θ, the Kruskall-Wallis equality of population rank test indicates significant differences at the 1% level between the population according the group size (Chi square (1) = 14.545 and *p-value*=0.0001), while for the levels of θ the difference is statistically significant at the 1% level (Chi square (3) = 14.451 and *p-value*=0.0024).

To compare outcomes across the treatments, we treat the outcome of each group as an independent variable and conduct the Mann-Whitney test for the average frequency of free riding. Recall that we refer to the treatments using the notations [N; θ; p], for instance [4; 0.25; 2] corresponds to the treatment [$N = 4; \theta = 0.25; p = 2$].

For the $N = 4$ player case, we fail to reject the null hypothesis that there is no difference in the frequency of free riding between the control treatment and the treatments in cases where payoffs to action B are randomized (*p-value*=0.132, [4; 0; 2] vs [4; 0.05; 2]; *p-value*=0.774, [4; 0; 2] vs [4; 0.25; 2]; *p-value*=0.773, [4; 0; 2] vs [4; 0.45; 2]).[28] A large majority of subjects play action A when there are 4 players in a group, which is consistent with *Prediction 1(i)* that subjects in "small" groups do not change their collective actions significantly following perturbations in the level of θ.

For the treatment where there were $N = 8$ subjects, the behavioral patterns differ consistently. Comparing across $\theta < \theta_1 = 0.07$ treatments (that is when $\theta = 0$ or $\theta = 0.05$), free riding behavior does not change significantly (*p-value*=0.771, [8; 0; 2] vs [8; 0.05; 2]); however, pooling across groups, the majority of free riding behavior occurs in these two treatments. The average frequency of choosing action B is 54.8% in treatment [8; 0; 2] and 51.8% in treatment [8; 0.05; 2], while the average frequency of choosing action B is 36.25% in treatment [8; 0.25; 2] and 27.08% in treatment [8; 0.45; 2]. This result provides evidence in support of *Prediction 1(ii)*. On the other hand, the null hypothesis of equal frequencies of free-riding across pairwise treatments is strongly rejected by the Mann-Whitney test (*p-value*=0.014, [8; 0; 2] vs [8; 0.25; 2]; *p-value*=0.009, [8; 0; 2] vs [8; 0.45; 2]; *p-value*=0.013, [8; 0.05; 2] vs [8; 0.25; 2]; *p-value*=0.009, [8; 0.05; 2] vs [8; 0.25; 2]). These results confirm that the extent of free-riding behavior decreases significantly for higher values of the exclusion probability. The null hypothesis of equal free-riding is also rejected for the high detection treatments (*p-value*=0.09, [8; 0.25; 2] vs [8; 0.45; 2]), albeit more weakly. For a high value of the exclusion probability (θ = 0.45), it is clear that a large majority of subjects tend to play action A. This provides some evidence in support of *Prediction 2* that players' incentive to free ride is affected by the coordination problem.

The Mann-Whitney test rejects the null hypothesis of no difference in the frequency of action B between the treatments [8; 0.25; 2] and [8; 0.25; 4] (*p-value*=0.03). Pooling across groups, the average frequency of choosing action B is 36.25% when $p = 2$, whereas it is 26.4% when $p = 4$, which provides preliminary evidence in support of *Prediction 3* that an increase in the value of the base payoff reduces the propensity of players to free-ride. On the other hand, the Mann-Whitney test fails to reject the null hypothesis of no difference in the frequency of action B between the treatment [8; 0.25; 4] and the treatment [8, 0.45;

2] (*p-value*=0.76). Pooling the groups, the average frequency of choosing B is 26.4% in [8; 0.25; 4] whereas it is 27.08% in [8; 0.45; 2]. This result suggests that an increase in p may have the same positive effect on players' coordination on action A than an increase in θ. Altogether, variation in the base payoff reveals that base income is an important determinant of collective behavior in settings with the possibility of exclusion.[29]

Econometric results

To further our empirical inquiry we make use of the pseudo-panel structure of our data and resort to standard panel data techniques to test our hypotheses. Note that the subjects in our experiment were randomly selected and allocated to each of our treatments. Specifically, each subjects was randomly allocated to each treatment and were only exposed to one treatment. Moreover all our sessions were conducted in the same lab and with identical conditions. Thus, we can assume that our observations are independent across treatments and sessions.[30] Our econometric strategy allows us to explore the independence of observations (Houser, 2008; Frechette, 2012). Because subjects in our experiment were asked to make dichotomous choices over a collective action, we resort to limited dependent variable panel data analysis in line with the approach in Rojas (2012). The fact that our subjects were only exposed to one treatment has two consequences: the first is that we cannot use fixed effects models because we have collinearity across periods and subjects; the second is that it is reasonable to assume the covariates are unrelated with subjects heterogeneity (Greene, 2008). For these reasons we obtain our results in this section using a random effects Probit model of the form:[31]

$$Pr(y_{it} = 1 \mid x_{it}) = \Phi(x_{it}\beta + \upsilon_i), \tag{5}$$

where y_{it} takes the unit value when the subject selects the choice B (i.e. to shirk), x_{it} is a vector of independent variables, υ_i is an iid error term, and Φ is the standard normal distribution. All our model estimations used robust standard errors clustered by group.[32] We interpret our dependent variable as the propensity to free ride or shirking on the collective rent. Variables in x_{it} include dummies accounting for the different treatments and a subject specific risk measurement variable.[33] The variable N accounts for the group size effects and takes unit value for the $N = 8$ player group. We introduce dummy variables to account for treatment effects with different exclusion probability, denoted $\theta 5$, $\theta 25$ and $\theta 45$, respectively, for a 5%, 25% and 45% probability of exclusion, and interpret our results in reference to the case of a zero probability of exclusion ($\theta = 0$).

Recall that we randomly assigned subjects to one of 3 treatments that vary in group size, in the exclusion probability, and in the base payoff of individuals. Each of our 372 subjects made choices over 20 periods of play, resulting in a total of 7,440 observations.

The effect of N on mean collective effort

We examine the effect of N on the mean effort level of individuals using all the observations in our panel data. Table 2. shows 3 models, the first examines the effect of N in isolation, then the second model examines the effect of a time trend and finally model 3 includes a linear and a quadratic time trend. The time trend variables are included to assess the impact of learning. As the table shows, there is not a significant difference in the coefficients of the mean free riding across the models. Model 3 provides the best fit to the data, suggesting that collective effort decreases at a decreasing rate in the number of rounds. Consistent with our first theoretical prediction, increasing the number of players in the group statistically significantly increases the propensity to shirk.

The effect of θ on mean collective effort

The second prediction of our model is that introducing a mechanism to identify and exclude shirking individuals from the collective rent reduces the propensity to shirk. As we show in Section 2, rent exclusion mechanisms are particularly effective when there is a larger number of participants in the group. Tables 3. and 4. respectively report the coefficients from our Probit model on shirking behavior and the marginal effects of covariates.

In Table 3., the first two models are estimated for the whole sample and include time trend variables and the risk aversion variable (model 2). Model 3 includes only observations where $N = 8$ and where the base payoff is at the lower end of the range.[34] For the estimation of models 4 and 5 we used data from all sessions that contain 8 members in a group. Note that in the base payoff treatments we double the value of p and we introduced a dummy variable taking the value of 1 when p takes a higher value. Models 2 through 5 include a variable to account for individual risk attitudes. This variable is statistically significant and has the expected sign, although its magnitude is rather low. Thus, controlling for risk attitudes does not substantially alter the coefficients on our treatment variables, an outcome consistent with findings in experiments conducted in strategic settings (Eckel and Wilson, 2004; Houser, Schunk, and Winter, 2010).[35]

Model 1 is similar to model 3 in Table 2., but now includes the exclusion probabilities. Recall that the exclusion probabilities are in terms of the baseline $\theta = 0$. The model clearly shows that an increase in the exclusion probability significantly decreases the propensity to shirk, consistent with *Prediction 2.* Moreover, notice that there is no significant difference in shirking behavior in the case of 0% and 5% probability of exclusion, an outcome that is consistent with *Prediction 1.* Turning to the marginal effects, table 4., shows that doubling the number of players increases the propensity to free ride by 32.6%, whereas a 25% and a 45% probability of rent exclusion, respectively, decreases the proportion of free riding by 15 and 24%.

Focusing on the sub-sample of our data with both eight players and $p = 2$, Model 3 shows rent exclusion is an effective mechanism to reduce the propensity to shirk. In fact, both probabilities of rent exclusion statistically and significantly reduce the rate of free riding when $N = 8$. Note that by examining the marginal effects on table 4. and as suggested in our models, the rent exclusion mechanism is particularly effective in incentivizing cooperation in large groups. The results reported in Model 3 provide strong evidence in support of *Prediction 2*, as an increase in the value of θ reduces shirking.

Taken together models 1, 2 and 3 in table 3. provide evidence in support of our second prediction. Across all the models, the θ treatment variables are strongly significant (with the exception of $\theta = 0.05$, which is consistent with *Prediction 1ii*).

The effect of p on mean collective effort

Monitoring and enforcement is a costly activity and, as our results show, there may a diminishing return to investments on punishing defecting agents. Thus, a combination of a carrot-and-stick policy might more effectively reduce the propensity to free ride as suggested in *Prediction 3* above. We examine this prediction in models 4 and 5 of Table 3.[36]. First, focusing on model 4, it is clear that including the treatments with a higher base payoff changes the magnitude of coefficients in the model and increase the values of the marginal effects. This suggests that the exclusion probabilities become more effective under higher payoffs (recall that the payment is forfeited when an individual is caught shirking).

Model 5 introduces the base payoff variable, which accounts for a higher fixed component of the rent from the collective reputation in the estimation. This covariate is statistically significantly and reduces the propensity to shirk as we hypothesized. Examining the marginal effects for models 4 and 5 in Table 4, it is clear that there seems to be a carrot-stick trade-off, as when we introduce the base payoff variable in the estimation the marginal effects of the probabilities of rent exclusion decrease to levels similar to those in model 3, while the increase in the fixed part of the payment to players reduces free riding by more that 12%. These results provide evidence supporting our *Prediction 3*.

Concluding Remarks and Management Implications

In this article we develop a framework to examine free-riding behavior on the quality of farm production in cooperative organizations. We construct and implement a multi-player collective action game in a controlled laboratory setting to analyze players' incentives to contribute collective effort towards group returns. We consider an imperfect monitoring design based on small detection frequencies that result in the potential for social sanctions. The discriminating features of our approach are that the collective rent depends on the coordinated investment of all players and the payoff from free-riding is randomized by an exogenous probability of exclusion. The laboratory setting results in a rich design framework that encompasses both a public good game and a multi-player assurance game. The games have different equilibrium outcomes, which allows us to systematically explore alternative policies to obtain the social optimum.

Our main testable hypotheses are as follows: (i) when the number of players is small, a pure public good game arises in which purely self-interested players incur costly effort to contribute to collective rents; (ii) when there is a large number of players and no exclusion probability, the game becomes a prisoners' dilemma in which the dominant strategy is to shirk on effort that generates collective rents; (iii) when there is a large number of players and a positive exclusion probability, the game reduces to a multi-player assurance game with multiple equilibria in which the Nash equilibrium involves coordinated effort for larger values of the exclusion probability; and (iv) when there is a large number of players and a positive exclusion probability, increasing the payoff individuals received independent of the collective rent leads to greater coordination.

To test our hypotheses, we develop a novel experimental protocol that nests a public good game when payoffs from free-riding are certain and a multi-player assurance game when a positive probability of exclusion results in randomized payoffs from free-riding.

Subjects were put in groups of 4 or 8 players and chose whether to incur a hidden cost to increase a collective rent. All the players in the group had to make their choice simultaneously and the magnitude of the collective rent depended on the coordination of effort across individuals. The groups of 4 and 8 players were further divided into four treatments corresponding to different probabilities of exclusion.

Our results indicate significant differences in the propensity to shirk on unobserved cost in larger groups of 8 players than in small groups of 4 players. In treatments with 8 players, we find collective effort significantly increases above a threshold level of the exclusion probability. Thus, the threat of being excluded from the collective rent significantly impacts behavior in the manner suggested by Tirole (1996)'s theory of collective reputation.

We also find that the contribution of unobserved costs by members towards collective rents is facilitated by the presence of a base payoff that is independent of returns from the coordinated outcome. Indeed, holding constant the size of the collective rent, we demonstrate that individuals who receive a larger base payoff contribute significantly higher effort to cooperative returns.

Our findings have a number of important management implications for cooperatives marketing strategies. First, our framework suggests returns to decentralizing management tasks. We find cooperative effort levels to be higher in smaller groups than in larger ones, which implies it may be possible to attain higher overall effort levels within a cooperative organization by sub-dividing production into a number of smaller divisions, for instance regional labels, where collective rents determined through profit-sharing arrangements can be organized among smaller groups. For instance, a wine cooperative may market several premium wines based on region-of-production and varietal to limit the number of members involved in each premium wine. Our findings suggest that such a strategy would enhance the quality contributed by all cooperative members.

Second, when members face a positive probability of being excluded from cooperative returns when shirking is detected, we find that even a "small" exclusion probability significantly increases coordination among members. This outcome emphasizes the essential role of social sanctions within marketing cooperatives in facilitating collective rents, an outcome that is broadly consistent with the theory of collective reputation.

Finally, marketing cooperatives may be able to successfully facilitate the production of high-quality agricultural products when members receive financial services that are unrelated to quality outcomes. Our findings suggest that the collective quality of cooperative-produced goods can be raised by providing members with a greater level of baseline services, which has significant policy implications when the return to providing baseline services to members is greater than the return from engaging in unproductive monitoring effort.

Notes

[1]Branding activities performed by marketing cooperatives differ by sector. In European Union, cooperatives develop branded products in the dairy and wine sectors and, to a lesser extent, in the fruit, vegetables and olive oil sectors (Hanisch and Rommel, 2003).

[2]Winemaking cooperatives typically purchase grapes from their members at market prices, and then distribute surplus proceeds from wine sales back to members in proportion to patronage (Hanisch and Rommel, 2003).

[3]By social sanctions, we consider sanctions imposed by the group as a result of monitoring and enforcement outcomes, as opposed to sactions by peers that involve individual responses to defection in repeated games.

[4]Such games have not received much attention the literature, apart from notable contributions in studying the Prisoner's Dilemma Game (Kollock, 1998).

[5]An important consideration for any laboratory setting is external validity of our work. From a methodological viewpoint, Kessler and Vesterlund (2015) argues that the relevant criteria for laboratory studies is whether the qualitative results of the study are externally valid (i.e., whether the sign of the effect would be the same in the lab and in the field). Field experiments (in our case, experiments with members of a marketing cooperative) are likely to produce qualitatively similar outcomes as lab experiments (Levitt and List, 2007).

[6]Ambrus and Greiner (2012) report that when noisy information on the subject's contribution is announced publicly, the sanction mechanism is efficient only when it is "strong" (i.e. when the damage is three times higher than under the baseline punishment). Xiao and Kunreuther (2016) find that a "restricted" regime of punishment (cooperators can punish defectors) is less efficient when (all) the outcomes of the game are uncertain.

[7]There are numerous examples of services provided to members of marketing cooperatives. For example, marketing cooperatives often perform first-stage processing services,

such as ginning cotton or hulling nuts, while others vertically integrate the processing and marketing functions of the supply-chain (USDA, 2000).

[8]An important exception is Saitone and Sexton (2009), who consider the effect of members' remuneration on free riding on quantity of high-quality products.

[9]Fares and Orozco (2014) examine the effect of contracts on product quality among wine cooperatives using a tournament mechanism and find that tournaments where producers can be promoted or demoted from high quality contracts provides stronger incentives than monitoring to control the moral hazard problem.

[10]In this literature a monitoring mechanism is "perfect" when the subjects' actions are perfectly observable and when the subjects' punishments are perfectly enforced.

[11]We assume here that the cooperative sells its products directly to final consumers. In cases where the cooperative sells its product to retailers, α is the portion of $\rho(N_A)a$ that the cooperative acquires from its bargaining power.

[12]For example, technicians in wine cooperatives engage in different tasks in collaboration with each member. These technicians have an important role of providing advice to wine growers on how to reduce production cost. Cooperative also provide other financial advantages, including regional subsidies for cooperatives located in EU countries (see Hanisch and Rommel, 2003). Such services may be financed by others activities of the cooperative, or by the membership fees.

[13]There are several interpretations of $\rho(N_A)a$ and p. Fu, Subramanian, and Venkateswaran (2015) define $\rho(N_A)a$ as the "performance-sensitive" component of the employee's compensation, $p + \rho(N_A)a$, where a refers to the "pay-performance sensitivity", while p is the "performance-invariant" component of the compensation. Alternatively, following Tirole (1996), p and $\rho(N_A)a$ may be interpreted, respectively, as the idiosyncratic values earned through past and present reputation.

[14]To reduce the analysis to the cases where the payoff to action A can never be negative,

we assume $p \geq c_A$.

[15]In practice, the member's production is excluded from the "premium-quality" activity of the cooperative. The member loses the services provided by the cooperative specifically for this activity and the cooperative sells the member's production at market price for a standard quality product (generally a competitive price). This is particularly the case in cooperatives that implement any pooling arrangement between the revenue of the "premium-quality" activity and the revenue of the the "standard-quality" activity.

[16]Note that $0 < \theta_1(N) < \theta_2(N) < 1$ for $N > \frac{a}{c_A}$.

[17]Technically, there is also a third Nash equilibria for $N_A = \tilde{n}$ in which some players contribute and some players free ride. This situation can be an equilibrium only if $\tilde{n}$ is exactly right, but it is strongly unstable (see Dixit, Skeath, and Reiley, 2009).

[18]This matching procedure corresponds to the so-called $partner$ procedure. Most of the recent experimental literature related to our study (e.g. Maier-Rigaud, Martinsson, and Staffiero, 2010, Ambrus and Greiner, 2012, Croson et al., 2015) relies on essentially the same matching procedure.

[19]This test consists of a menu of 10 paired lottery choices designed to make inferences about risk preferences under various payment conditions.

[20]Part a) of figure 1. illustrates the partition of equilibrium for N=8 and p=2.

[21]Part b) of figure1. illustrates the partition of equilibrium for N=8 and p=4.

[22]Payoffs were displayed only when the play of the game started. During the reading of the instructions, the payoffs in the tables were symbolized by letters.

[23]In all cases, we avoid the use of wording such as "probability of being monitored" or "exclusion probability".

[24]The payment to subjests was the cumulative sum of payoffs across all periods of the game, which is a conventional payment scheme used in multitask experiments. An alternative payment scheme would involve randomly picking one or more periods of the game

for subjects' payoffs, which has the advantage of avoiding "portfolio" effects in subjects' strategies. We use the cumulative payoff scheme because it is simpler to understand and more closely matches payoffs through repeated interaction among members of marketing cooperatives. Charness, Gneezy, and Halladay (2016) investigate the effect of "pay one vs pay all" payment schemes on subjects' actions in fourteen published laboratory experiments and report qualitatively similar outcomes for the two, alternative schemes in 10 out of 14 cases.

[25]In order to get their payoff for the Holt and Laury's test, a computer program enabled subjects to throw a 10-sided die twice: the first time to determine the relevant lottery, and the second time to determine the payoff for the chosen action. This procedure of payment was carried out at the end of the experiment, to ensure that the subjects' behaviors in the game were not influenced by their earnings in the risk test. Payoffs to lotteries were labeled in Euros and were identical to those of Holt and Laury (2002).

[26]Examples of subjects' payoff tables are available in the appendix.

[27]Prediction 3 states that an increase in the base payoff p should reduce the propensity to free ride. It is especially interesting to test this prediction for intermediate levels of the exclusion probability (θ = 0.25), because an increase in the base payoff is capable of solving the coordination problem in this range.

[28]The null hypothesis cannot be rejected either when we run the test between the other pairwise treatments.

[29]The Mann-Whitney test rejects the null hypothesis of no difference in the frequency of action B between the treatments [8; 0,45; 2] and [8; 0,45; 4] (*p-value*=0.07), albeit rather weakly. Pooling the groups, the average frequency of choosing B is 27.08% when $p = 2$, whereas it is 17.9% when $p = 4$.

[30]Recall that for each treatment we had several sessions, subjects were randomly selected and allocated across treatments and sessions.

[31]While it is common practice to start the estimation with a linear probability model, given the length of the article we opted to show those results in a supplemental web appendix.

[32]Note that in each session we had groups of 4 our 8 subjects clustered together. Given that in our experimental design each subject only was exposed to one treatment, there could be subjective specific variation so we also estimated the models clustering by individual. The later results are shown in the web appendix.

[33]The risk variable is a quantitative variable that corresponds to the number of safe choices made by the subjects in the Holt and Laury (2002)'s multiple prize list test.

[34]For this model we use data from session with 160 participants in total, each making 20 choices.

[35]Consistent with other studies using this test, our subjects are overwhelmingly risk averse (73.7%). Only 11% of the subjects are risk neutral and 5% are risk loving. The lack of correlation between risk attitudes and the decision to free ride on collective rents is somewhat surprising.

[36]To test the robustness of our estimation estimated these models using time trends and lags of the dependent variable. These are reported in a supplemental web appendix.

Tables and Figures

Table 1. Experimental sessions and group observations

Dates	Treatment $[N; \theta; p]$	Nb of Sessions	Nb of subjects	Nb of observations
Feb. 16, 2014	[4; 0; 2]	1	16	4
Feb. 18, 2014	[8; 0; 2]	2	32	4
Feb. 18, 2014	[4; 0.45; 2]	1	16	4
Feb. 19, 2014	[8; 0.45; 2]	1	16	2
Feb. 20, 2014	[8; 0.45; 2]	1	16	2
Feb. 24, 2014	[4; 0.05; 2]	1	12	3
Feb. 24, 2014	[8; 0.05; 2]	1	16	2
Feb. 26, 2014	[8; 0.05; 2]	1	16	2
Feb. 27, 2014	[4; 0.25; 2]	1	16	4
Feb. 27, 2014	[8; 0.25; 2]	2	32	4
May 19, 2014	[4; 0; 2]	1	12	3
May 19, 2014	[8; 0.45; 2]	1	16	2
May 20, 2014	[4; 0.05; 2]	1	12	3
May 21, 2014	[8; 0.25; 2]	1	16	2
March 23, 2015	[8; 0.25; 4]	1	16	2
March 26, 2015	[8; 0.25; 4]	3	48	6
March 27, 2015	[8; 0.25; 4]	1	16	2
March 31, 2015	[8; 0.45; 4]	1	16	2
April 1, 2015	[8; 0.45; 4]	1	16	2
April 2, 2015	[8; 0.45; 4]	1	16	2
	Total	24	372	57

Table 2. Random effects Probit estimation for the choice of *B* for the pooled data with robust standard errors clustered by session

	Model 1	Model 2	Model 3
	Coefficient (Standard Error)	Coefficient (Standard Error)	Coefficient (Standard Error)
N	1.014^{***} (0.211)	1.014^{***} (0.211)	1.014^{***} (0.212)
$period$		0.001 (0.005)	0.054^{***} (0.018)
$period^2$			-0.002^{***} (0.001)
cons	-1.659^{***} (0.194)	-1.668^{***} (0.215)	-1.863^{***} (0.227)
Log likelihood	-3559.65	-3559.61	-3550.96
ρ	0.454	0.454	0.455
$Wald\ \chi^2$	23.00^{***}	26.8^{***}	32.72^{***}
obs.	7440	7440	7440

Note: ***Corresponds to the 1% significance level, ** to the 5% and * to the 10%.

Table 3. Random effects Probit estimation for the choice of *B* with robust standard errors clustered by groups in each session

	Model 1	Model 2	Model 3	Model 4	Model 5
	Coefficient (Std. Err.)	Coefficient (Std. Err.)	Coefficient (Std. Err)	Coefficient (Std. Err)	Coefficient (Std. Err)
N	1.327^{***} (0.188)	1.322^{***} (0.188)			
$\theta 5$	0.149 (0.198)	0.155 (0.201)	-0.09 (0.107)	-0.094 (0.11)	-0.094 (0.107)
$\theta 25$	-0.608^{***} (0.191)	-0611^{***} (0.191)	-0.60^{***} (0.137)	-0.860^{***} (0.13)	-0.578^{***} (0.149)
$\theta 45$	-0.942^{***} (0.196)	-0.952^{***} (0.198)	-0.85^{***} (0.163)	-1.134^{***} (0.16)	-0.915^{***} (0.149)
$period$	0.054^{***} (0.018)	0.054^{***} (0.017)	0.089^{***} (0.018)	0.079^{***} (0.0172)	0.079^{***} (0.172)
$period^2$	-0.004^{***} (0.0008)	-0.002^{***} (0.0008)	-0.002^{***} (0.0008)	-0.003^{***} (0.0008)	-0.003^{***} (0.0007)
Risk		-0.0023^{***} (0.0007)	-0.0006 (0.0007)	-0.002^{**} (0.0008)	-0.002^{**} (0.0007)
Base Payoff					-0.450^{***} (0.123)
cons	-1.603^{***} (0.276)	-1.612^{***} (0.278)	-0.255^{*} (0.86)	-0.236^{**} (0.106)	-0.236^{**} (0.106)
Log likelihood	-3520.48	-3517.82	-1899.23	-3028.98	-3021.37
ρ	0.39	0.395	0.287	0.374	0.360
$Wald\ \chi^2$	109.39^{***}	147.48^{***}	55.29^{***}	115.41^{***}	144.34^{***}
obs.	7440	7440	3200	5760	5760

Note: ***Corresponds to the 1% significance level, ** to the 5% and * to the 10%.

Table 4. Marginal effects of covariates for Models 2 (7440 obs.) and 3 (3200) 4 and 5 (5760 obs.)

	Model 2	Model 3	Model 4	Model 5
	Marg. Effect (Std. Err)	Marg. Effect (Std. Err)	Marg. Effect (Std. Err)	Marg. Effect (Std. Err)
N	0.326 (0.043)			
$\theta 5$	0.038 (0.049)	-0.029 (0.035)	-0.026 (0.03)	-0.026 (0.030)
$\theta 25$	-0.15 (0.047)	-0.19 (0.043)	-0.239 (0.033)	-0.161 (0.037)
$\theta 45$	-0.235 (0.049)	-0.277 (0.050)	-0.315 (0.040)	-0.256 (0.39)
Base Payoff				-0.126 (0.034)

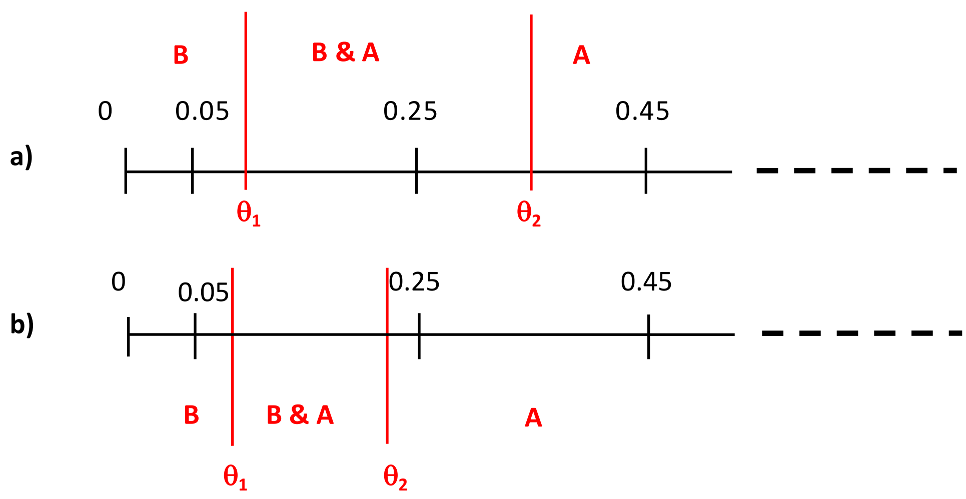

Figure 1. Partition of equilibrium for *N*=8 with a) *p*=2, and b) *p*=4

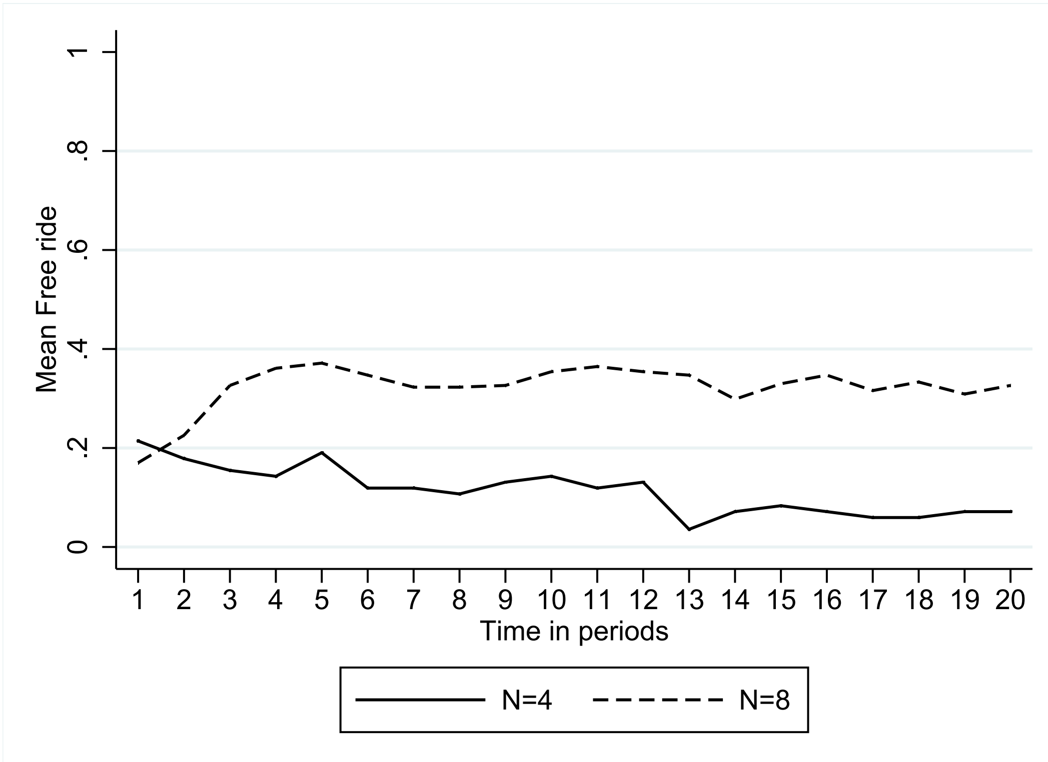

Figure 2. Mean free ride when *N*=4 and *N*=8 (7440 obs.)

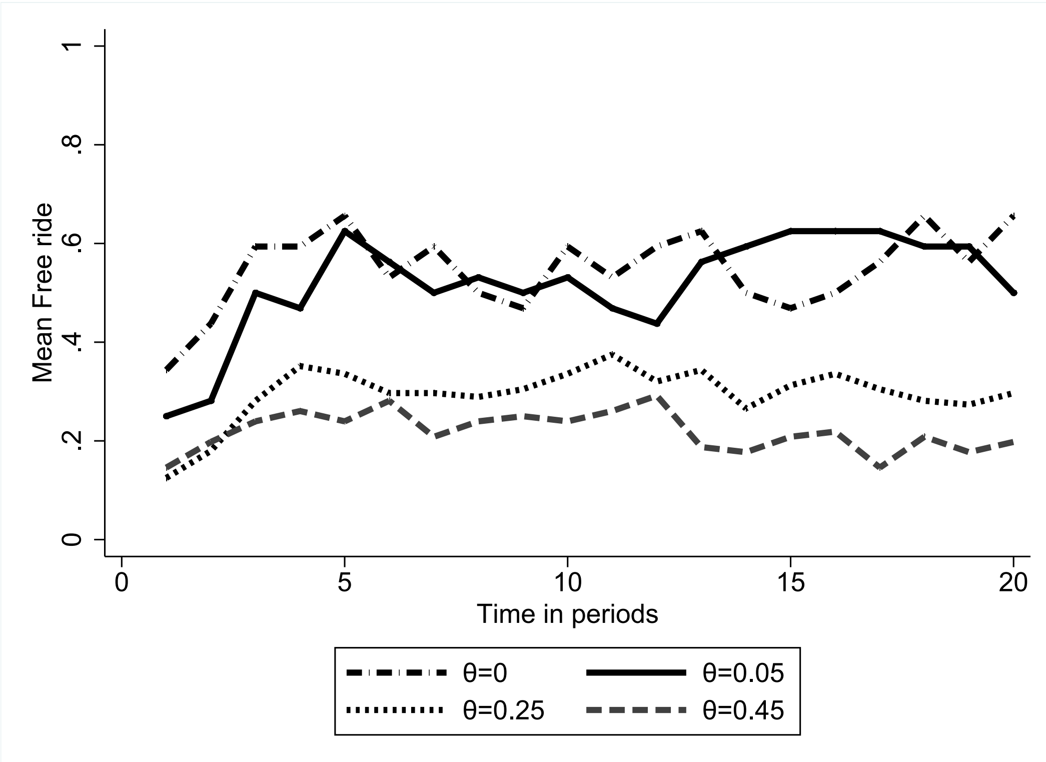

Figure 3. Average free ride when *N*=8 according the treatments and BasePayoff=2 (3200 obs.)

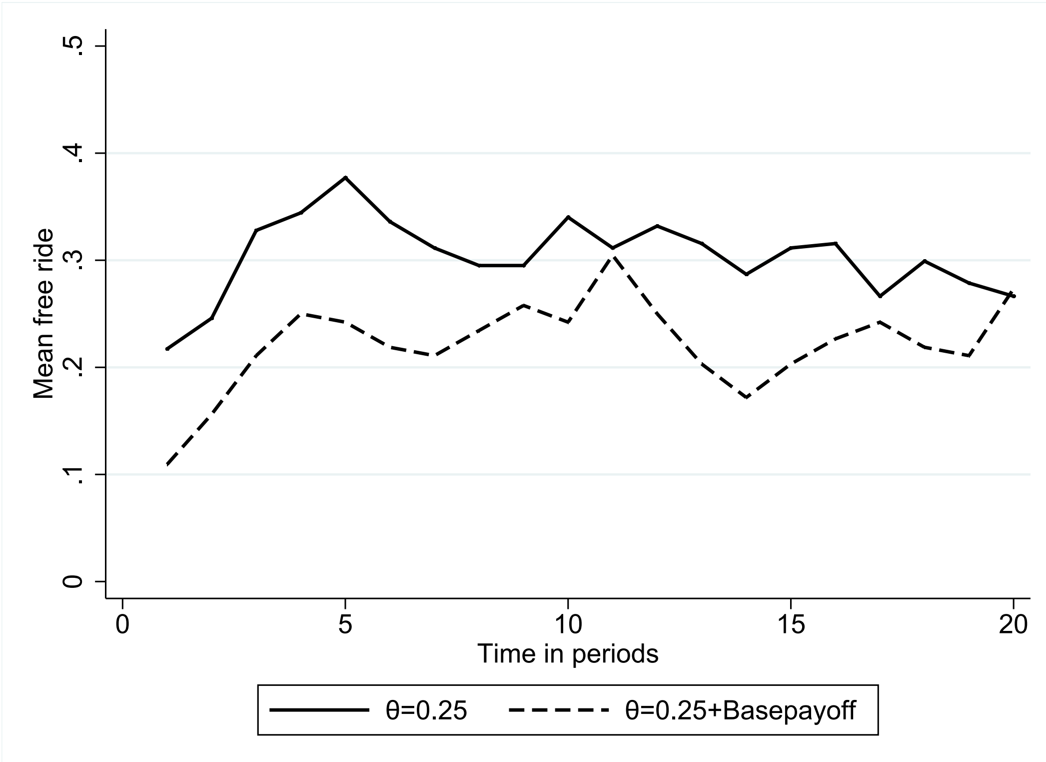

Figure 4. Average free ride when *N*=8, *θ*=0.25 and lower (960 obs) or higner levels of BasePayoff (1600 obs.)

References

Ambrus, A., and B. Greiner. 2012. Imperfect Public Monitoring with Costly Punishment: An Experimental Study. *American Economic Review* 102:3317–32.

Bereby-Meyer, Y., and A.E. Roth. 2006. The Speed of Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation. *American Economic Review* 96:1029–1042.

Castriota, S., and M. Delmastro. 2015. The Economics of Collective Reputation: Evidence from the Wine Industry. *American Journal of Agricultural Economics* 97:469–489.

Charness, G., U. Gneezy, and B. Halladay. 2016. Experimental Methods: Pay One or Pay All. *Journal of Economic Behavior and Organization* 131:141 – 150.

Cinyabuguma, M., T. Page, and L. Putterman. 2005. Cooperation Under the Threat of Expulsion in a Public Goods Experiment. *Journal of Public Economics* 89:1421–1435.

Croson, R., E. Fatas, T. Neugebauer, and A.J. Morales. 2015. Excludability: A Laboratory Study on Forced Ranking in Team Production. *Journal of Economic Behavior and Organization* 114:13–26.

De Geest, L.R., J.K. Stranlund, and J.M. Spraggon. 2017. Deterring Poaching of a Common Pool Resource. *Journal of Economic Behavior and Organization* 141:254 – 276.

Dixit, A., S. Skeath, and D. Reiley. 2009. *Games of Strategy*. W.W. Norton.

Eckel, C.C., and R.K. Wilson. 2004. Is Trust a Risky Decision? *Journal of Economic Behavior and Organization* 55:447–465.

Fares, M., and L. Orozco. 2014. Tournament Mechanism in Wine-Grape Contracts: Evidence from a French Wine Cooperative. *Journal of Wine Economics* 9:320–345.

Fehr, E., and S. Gächter. 2000. Cooperation and Punishment in Public Goods Experiments. *American Economic Review* 90:980–994.

Fishman, A., I. Finkelshtain, A. Simhon, and N. Yacouel. 2015. The Economics of Collective Erands. *Journal of Industrial Economics, forthcoming*, pp. .

Frechette, G. 2012. Session-effects in the Laboratory. *Experimental Economics* 15:485–498.

Fu, R., A. Subramanian, and A. Venkateswaran. 2015. Project Characteristics, Incentives, and Team Production. *Management Science* Articles in Advance.

Fulton, M., and K. Giannakas. 2004. Inserting GM Products into the Food Chain: The Market and Welfare Effects of Different Labeling and Regulatory Regimes. *American Journal of Agricultural Economics* 86:42–60.

Goodhue, R.E., D.M. Heien, H. Lee, and D.A. Sumner. 2003. Contracts and Quality in the California Winegrape Industry. *Review of Industrial Organization* 23:267–282.

Greene, W.H. 2008. *Econometric Analysis*, 6th ed. Upper Saddle River, NJ: Prentice Hall.

Hamilton, S.F., and D. Zilberman. 2006. Green Markets, Eco-Certification, and Equilibrium Fraud. *Journal of Environmental Economics and Management* 52(3):627–644.

Hanisch, M., and J. Rommel. 2003. *Support for Farmers Cooperatives; Case Study Report*. Wageningen University, Wageningen, The Netherlands.

Hölmstrom, B. 1979. Moral Hazard and Observability. *Bell Journal of Economics* 10:74–91.

Holt, C.A., and S.K. Laury. 2002. Risk Aversion and Incentive Effects. *American Economic Review* 92:1644–1655.

Houser, D., D. Schunk, and J. Winter. 2010. Distinguishing Trust from Risk: An Anatomy of the Investment Game. *Journal of Economic Behavior and Organization* 74:72–81.

Houser, D.E. 2008. Experiments and Econometrics. In S. N. Durlauf and L. E. Blume, eds. *The New Palgrave Dictionary of Economics*. Basingstoke: Palgrave Macmillan.

Hueth, B. 2014. Missing Markets and the cooperative firm. Workshop on Producers' Organizations in Agricultural Markets, September 4-5, 2014, Toulouse, Toulouse School of Economics.

Kessler, J.B., and L. Vesterlund. 2015. The External Validity of Laboratory Experiments: The Misleading Emphasis on Quantitative Effects. In G. R. Frechette and A. Schotter, eds. *Handbook of Experimental Economics Methodology*. Oxford Scholarship Online.

Kollock, P. 1998. Social Dilemmas: The Anatomy of Cooperation. *Annual Review of Sociology* 24:183–214.

Levitt, S.D., and J.A. List. 2007. Viewpoint: On the Generalizability of Lab Behaviour to the Field. *Canadian Journal of Economics* 40:347–370.

Maier-Rigaud, F.P., P. Martinsson, and G. Staffiero. 2010. Ostracism and the Provision of a Public Good: Experimental Evidence. *Journal of Economic Behavior and Organization* 73:387–395.

Mérel, P., T.L. Saitone, and R.J. Sexton. 2015. Cooperative Stability under Stochastic Quality and Farmer Heterogeneity. *European Review of Agricultural Economics* 42:765–795.

Nosenzo, D., T. Offerman, M. Sefton, and A. van der Veen. 2016. Discretionary Sanctions and Rewards in the Repeated Inspection Game. *Management Science* 62:502–517.

Pennerstorfer, D., and C.R. Weiss. 2013. Product Quality in the Agri-Food Chain: Do Cooperatives Offer High-Quality Wine? *European Review of Agricultural Economics* 40:143–162.

Planas, J. 2016. The Emergence of Winemaking Cooperatives in Catalonia. *Business History* 58:264–282.

Rojas, C. 2012. The Role of Demand Information and Monitoring in Tacit Collusion. *RAND Journal of Economics* 43:78–109.

Saitone, T.L., and R.J. Sexton. 2009. Optimal Cooperative Pooling in a Quality-Differentiated Market. *American Journal of Agricultural Economics* 91:1224–1232.

—. 2010. Product Differentiation and Quality in Food Markets: Industrial Organization Implications. *Annual Review of Resource Economics* 2:341–368.

Swope, K.J. 2002. An Experimental Investigation of Excludable Public Goods. *Experimental Economics* 5:209–222.

Tirole, J. 1996. A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality). *Review of Economic Studies* 63:1–22.

USDA. 2000. *Agricultural Marketing Cooperatives*. Cooperative Information Report 45, Section 15, USDA Publishing.

Winfree, J.A., and J.J. McCluskey. 2005. Collective Reputation and Quality. *American Journal of Agricultural Economics* 87:206–213.

Xiao, E., and H. Kunreuther. 2016. Punishment and Cooperation in Stochastic Social Dilemmas. *Journal of Conflict Resolution* 60:670–693.

Appendix: Examples of subjects' payoff tables

Matrix 1. Subjects' payoff matrix in the treatment [4; 0; 2]

In my group, number of	In my group, number of	My payoff for	My payoff for
κ	ϕ	κ	ϕ
4	0	1000	
3	1	750	950
2	2	500	700
1	3	250	450
0	4		200

Note: In the experiment, actions A and B were labelled κ and ϕ, respectively. To make their choice in each period, subjects had simply to click on screen buttons "κ" and "ϕ" placed below this matrix.

Matrix 2. Subjects' payoff matrix in the treatment [4; 0.45; 2]

In my group, number of	In my group, number of	My payoff for κ	My payoff for ϕ	
κ	ϕ	with 100/100	with 55/100	with 45/100
4	0	1000		
3	1	750	950	0
2	2	500	700	0
1	3	250	450	0
0	4		200	0

Matrix 3. Subjects' payoff matrix in the treatment [8; 0; 2]

In my group, number of	In my group, number of	My payoff for	My payoff for
κ	ϕ	κ	ϕ
8	0	1000	
7	1	875	1075
6	2	750	950
5	3	625	825
4	4	500	700
3	5	375	575
2	6	250	450
1	7	125	325
0	8		200

Matrix 4. Subjects' payoff matrix in the treatment [8; 0.45; 2]

In my group, number of	In my group, number of	My payoff for κ	My payoff for ϕ	
κ	ϕ	with 100/100	with 55/100	with 45/100
8	0	1000		
7	1	875	1075	0
6	2	750	950	0
5	3	625	825	0
4	4	500	700	0
3	5	375	575	0
2	6	250	450	0
1	7	125	325	0
0	8		200	0

Matrix 5. Subjects' payoff matrix in the treatment [8; 0.25; 4]

In my group, number of	In my group, number of	My payoff for κ	My payoff for ϕ	
κ	ϕ	with 100/100	with 75/100	with 25/100
8	0	1200		
7	1	1075	1275	0
6	2	950	1150	0
5	3	825	1025	0
4	4	700	900	0
3	5	575	775	0
2	6	450	650	0
1	7	325	525	0
0	8		400	0