Voting and Contributing While the Group is Watching

Emeric Henry, Charles Louis-Sidois

Voting and contributing when the group is watching

Emeric Henry and Charles Louis-Sidois*

October 26, 2015

Abstract

Members of groups and organizations often have to decide on rules that regulate their contributions to common tasks. They typically differ in their propensity to contribute and often care about the image they project, in particular want to be perceived by other group members as being high contributors. In such environments we study the interaction between the way members vote on rules and their subsequent contribution decisions. We show that multiple norms can emerge. We draw surprising policy implications, on the effect of group size, of supermajority rules and of the observability of actions.

1 Introduction

In May 2009, the elected members of the French Assemblee Nationale (French Parliament) voted a law imposing sanctions for those among them not attending weekly meetings of committees. Even without sanctions, some officials would have attended, driven by an individual sense of duty or a concern for the image the group has of them. Sanctions nevertheless significantly increased attendance: from an average of 7 meetings attended per year, it jumped to 19 meetings. Moreover there appeared to be a systematic link between the way parliamentaries voted on sanctions and the way they behaved after that.

Similarly, most members of groups and organizations (firms, NGOs, academic departments...) choose the rules that govern their interactions, in particular those regulating tasks with group externalities, such as attending meetings, writing reports or participating in team work. One important driver of contributions to common tasks are image concerns, i.e caring about how the group perceives you. For instance

*Sciences Po Paris, 28 rue des Saint Peres, 75007 Paris (France)

Ariely et. al. (2009) show that efforts made to contribute to a good cause are much higher when individuals are observed by others.

In such contexts, we study how group members vote on rules and subsequently behave. There is a close interaction between voting behavior and contribution decisions. A sanction, by increasing overall contributions, will decrease the honor that can be derived from individually contributing. Thus voting affects the calculus of reputation. In turn, when votes are public, concern for image will affect voting behavior. Uncovering the subtle interactions inherent in these environments is the object of this paper.

Specifically, we study a model that includes two stages involving the same group of players. In the second stage players simultaneously choose whether to contribute to a public good. In the first stage, the same players vote on a given sanction s to be imposed in the second stage to non contributors. Group members are heterogenous in their intrinsic propensity to contribute, what we call their type, but all members want the group to perceive them as having a high type. These inferences made by the group are based on observed actions. We consider several types of situations. In the first, the vote is secret and the action observed by others is the contribution decision. In the second, only the vote is public and the inferences of others are based exclusively on voting behavior. Finally we consider the case where both are public.

In the public good contribution stage, three categories of members emerge. Those with a high type, called always-participants, contribute regardless of whether the sanction was voted or not. Those with a low type, called never-participants, never contribute. Finally those with intermediate values, called swing-participants, contribute if and only if the sanction was voted.

Consider first the case where the vote is secret. In the voting phase, the members of these three groups will have different incentives. We first show the surprising result that it can be a dominant strategy for always participants to vote against the sanction, even though the sanction will never apply to them. Indeed, even though these members gain from a sanction because other group members increase their contribution and thus impose a positive externality, they however lose in reputation as taking the action is no longer so rare that it signals a high intrinsic value. If the sensitivity to reputation is sufficiently high, these members will vote against the sanction.

If the externality gain is big enough they will however vote in favor. Furthermore we show that, in spite of the very different motivations of the three categories of members, always participants always have higher incentives to vote in favor than the swing participants who themselves have higher incentives than the never participants. We thus find that the equilibrium in pure and symmetric strategies is of the cutoff form where members vote for the sanction if and only if their type is above a cutoff value V^*. We show that there can exist several equilibria, that can be interpreted as corresponding to different norms of behavior: some with a high V^*, i.e groups tending to vote against sanctions because the gain in additional public good is expected to be low and those with a low V^* more likely to vote for sanctions as they expect higher benefits.

Technically, the multiplicity is linked to the information aggregated when voters consider the case where their vote is pivotal. The gain from a sanction comes from the additional public good provided in the second stage. The only types that change their actions depending on the sanction are the swing participants. Thus, to evaluate the benefit of the sanction, individuals need to estimate the number of swing participants in the group. We find that the information aggregated on the proportion of swing participants in the group can generate the multiplicity of equilibria mentioned above.

We derive policy implications, restricting our attention to cases where there is a unique stable equilibrium. We show that the effect of increasing sanctions on the outcome of the vote can go in both directions depending on parameters. This reflects a tradeoff similar to the one mentioned above. Higher sanctions will mean a higher expected externality gain, but at the same time a reduced payoff from reputation since the act becomes more banal. We also show that the supermajority required for approval of the law has a surprising effect on the vote outcome. Increasing the required supermajority can decrease the voting cutoff and even overall make approval more likely even though more votes in favor are required for approval. Finally we show that the manager of the organization wanting to maximize contributions often has an incentive to publicly disclose contributions in the second phase: it always increases contributions since members care about their image, but also often increases the probability that the sanction is adopted in the first stage.

We then turn to the case where the individual votes are public, but the contributions are kept secret. There is then a reputation payoff attached to the vote (regardless of whether the member is pivotal or not) that has to be traded off against the probability of being pivotal. We find that the voting equilibrium is again characterized by a cutoff, since the members with a higher type have more to gain from the sanction. We show also that as the size of the group increases, sanctions will tend to be approved regardless of their size: the probability of being pivotal goes to zero and the only motivation that remains is to vote in the way sanctioned by the social norm.

We finally examine the case where both votes and contributions are observable. We show that the equilibrium is no longer necessarily of the cutoff type, given that the members now have two ways of signalling their type. It is possible that low types vote against and don't contribute, intermediate types vote for and don't contribute while high types vote against and contribute.

Bénabou and Tirole (2011) is the paper closest to ours. They examine a public good problem, very similar to the second stage of our model, and show how the calculus of honor and stigma can be derived. Their key focus is on how an informed principal can optimally set incentives. They also examine the case where the choice of the principal reveals information about the societies values. We take a different stance and examine instances where the rules are endogenously determined by the group, not set by a principal. Such situations are of high practical relevance and furthermore lead to subtle theoretical interactions. We show how the law can shape the calculus of honor and how in turn social norms can influence voting. We derive surprising predictions on the effect of different institutional features.

There is in fact quite a large experimental literature examining the difference between exogenously and endogenously set sanctions on future behavior. Part of the literature (Galbiati and Vertova 2008 and Galbiati et al. 2013) examines the case where the designer who decides on the sanction is informed, contrary to our setting, which relates to the expressive nature of the law that motivates Bénabou and Tirole (2011). Tyran and Feld (2006) consider an experimental setting closer to our model and show that if the group votes for the sanction (rather than have a sanction exogenously imposed), it is followed by higher contributions. We return to that evidence in the conclusion.

Our paper is also closely related to the literature on strategic voting and information aggregation in voting, where the initial motivation was to revisit the Condorcet jury theorem when including strategic concerns (Austen-Smith and Banks 1996, Feddersen and Pesendorfer 1996, 1997, 1998). To the best of our knowledge, in all the papers in this literature, the benefits of the law submitted to a vote is exogenously given (but not publicly observed). In our public good setting, the benefit of the sanction is endogenously determined by how voters react to it. This leads to a multiplicity of equilibria not present in the rest of the literature.[1]

There is also a growing literature on aggregation of information in committees which is related to our analysis of public voting in section 5. In many of these papers, voters experience a given disutility if their vote indicates they had bad information. For instance in Midjord et al. (2013), privately informed agents vote on an approval decision and get a negative reputation payoff (of fixed value) if the outcome is to approve and the state was in fact bad. They show that this will lead the committee to be overly cautious and in the limit when the size is large, will lead to sure rejection. This relates to our result in the case of public voting, where we show that for large organizations sanctions are always adopted, except that in our case the reputation payoff is endogenously determined. Ottaviani and Sorensen (2001) include a similar concern for appearing well informed in a model where voters express themselves sequentially.

Godefroy and Perez-Richet (2013), consider a sequence of two votes in a committee with privately informed voters, the first to select the issue to be submitted to a vote and a second to vote on approval of this issue versus status quo. They focus on the effect of the supermajority requirement in the selection stage on voting behavior and show that a more conservative rule implies more conservative voting behavior. Similarly, in section 4.2, we show that in our very different setting, the supermajority requirement in the first stage can have surprising implications on voting behavior due to information aggregation properties, but the mechanics of information aggregation are quite different. In fact, in our setting we find that a more stringent rule can make voters more inclined to vote for the sanction. Even more surprisingly, a more conservative rule can make overall approval more likely.

One of the key elements of the model is the fact that individuals care about the image others have of them, what has been coined image concerns. Many papers, in the lab or in the field, establish empirically the importance of these image concerns. Ariely et. al. (2009) for instance compare effort levels in treatments that varied in three dimensions: subjects were either observed or unobserved, received monetary incentives or not and contributed either to a “good cause” (Red Cross) or a “bad one” (NRA). They find that being observed increased effort levels only when subjects did not receive monetary incentives or when they volunteered for a good cause.

[1] One exception is Callander (2008) who examines a model where privately informed voter want to elect the best candidate but also want to vote for the winner. The source of multiplicity is very different: if the rest of the group is more likely to vote against their signal, it also makes sense to do so as there is a desire to be on the winning side.

Della Vigna et al. (2012) show that notifying residents in advance of the exact time of sollicitation in a door to door fundraiser significantly decreases the share of households opening doors, one possible interpretation being that image concerned individuals attempt to avoid the pressure.[2]

There is also extensive laboratary experiment. Andreoni and Petrie (2004) find that contributions in a public goods game increased when the players were not anonymous. Dana et al. (2006) offer participants a costly option to opt out of a dictator game and show that giving in the dictator game is in part motivated by participants not wanting to appear selfish. In the same spirit, other contributions (Rege and Telle 2004, Sanek and Sheremeta 2013), find that providing options for the participants to overcome their moral dilemmas significantly lowers transfers.

The rest of the paper is organized as follows. In section 2 we present the model, in section 3 we derive results for the case of a secret vote but an observable contribution, and discuss robustness in the case of bonuses or complementarities in public good production. We draw implications in section 4. In sections 5 and 6, we examine instances where individual votes are publicly revealed. Finally we discuss empirical implications in section 7 and conclude.

2 Model

We consider a two stage game involving $2N+1$ players. In the first stage, a rule (or law in certain contexts) is submitted to a vote. The rule specifies a sanction $s > 0$ that will be imposed in case of free riding in the public good stage that will follow. Note that the case of a bonus for contributing would lead to very similar results (discussed in section 3.3.1). However for notational simplicity we focus for most of the results on a positive sanction.

All players cast their vote simultaneously. The voting decision of individual i is denoted $b_i \in \{0,1\}$ (where b stands for ballot). If strictly more than K voters vote in favor, the law is adopted. For most of the paper we will consider the case of majority rule, i.e $K = N$. The outcome of the vote is publicly revealed and the players then simultaneously decide, in a second stage of the game, on a public good contribution a_i, where $a_i \in \{0,1\}$.

[2] Henry and Sonntag (2015) propose an experimental game to measure individual sensitivity to image concerns and show that image concerned individuals often try to avoid situation where their reputation is at risk. Note that there are also instances where image concerns don't seem to play a role, such as in Gill et al (2015).

For a given approved sanction s and a given vector of contributions to the public good $a = \{a_1, a_2, ..., a_{2N+1}\}$, the utility of player i is given by:

$$U_i = (v_i - c)a_i - s(1 - a_i) + e\frac{\sum_{j \neq i} a_j}{2N} + \mu E[v_i|y_i]$$

Individual i gets an intrinsic benefit of contributing to the public good denoted v_i, which will characterize the type of the individual. There is also a cost of contribution c common to the whole population. If a sanction is in place, there will be an additional cost for those not contributing $s(1 - a_i)$.[3] In addition, there is an externality gain derived from the contributions of others $e\frac{\sum_{j \neq i} a_j}{2N}$. We consider in section 3.3.2 the case where the production of the public good involves complementarities

Finally, the utility contains a less standard term, $\mu E[v_i|y_i]$ (as in Bénabou and Tirole 2011), which corresponds to a reputation concern. The reputation is based on an observed individual choice y_i by player i. We will consider three cases. First, the case where the vote is secret but the contribution is observable, i.e $y_i = a_i$, an environment we call *public contributions*. This is in fact the main focus of the paper. Second, the case where individual votes are revealed but the individual contributions are kept secret, i.e $y_i = b_i$, what we call *public vote*. Finally we consider the case where both votes and contributions are publicly revealed $y_i = (a_i, b_i)$, what we call *both public*.[4] All cases can be relevant in practice. The agents can value reputation per se because they care about the perception others have of them, or this reputation might have future financial consequences not explicitly specified in the model.

Actions reveal information on the underlying value of v_i, the intrinsic motivation of each agent. Individuals privately observes their v_i, what we will call their type. Specifically v_i are i.i.d. drawn from the distribution of types $f(v)$ with support $[v_{min}, v_{max}]$, continuously differentiable with $f(v) > 0$. The reputation term $E[v_i|y_i]$ is thus the expected value of v_i given action y_i using the distribution of types f.[5]

[3] From the point of view of group members, the sanctions is a pure loss, in particular is not redistributed to the group. We did not want to add this extra consideration that would make our message less clear.

[4] We also consider in section 4.3 the situation *all secret*

[5] Note that we assume that the expectation is taken with respect to the prior distribution f. As in Levy (2007), information could in fact be obtained about individual votes from the aggregate result of the vote. The voters could thus update their belief about the type distribution. This would significantly complicate the resolution and increase the multiplicity of equilibria. We thus assume that the expectation is conditioned only on individual actions. It appears to us to be a reasonable assumption for this term which in any case is a behavioral component of the utility function: other members do not use the aggregate result to base their belief about individual types when the individual actions are so salient.

The timing of the game is thus the following:

1. Types are drawn and privately observed
2. Players vote on the rule with no abstention. The outcome of the vote is publicly revealed
3. Players then simultaneously decide on their contribution decision.

We focus on symmetric Bayesian Nash equilibria (where players with the same type choose the same strategy).

3 Voting on sanctions: secret votes

We first consider situations where individual votes are secret, but the contributions are publicly observed. In most organizations and for most public goods, the contribution is directly observable (for instance participation to meetings, preparation of reports...). The organization might however have in place secret voting. This is the type of situations we first consider, the *public contributions* environment.

3.1 Contribution stage

In the second stage, all players observe whether the sanction was voted or not and decide on their contribution to the public good. For a given sanction s (where $s = 0$ corresponds to the case where voters turned down the sanction), contributing yields intrinsic benefits and costs as well as the honor of doing the right thing:

$$v_i - c + \mu E[v_i|a = 1]$$

Not contributing on the contrary exposes individual i to a sanction and to the stigma of not contributing:

$$-s + \mu E[v_i|a = 0]$$

As in Bénabou and Tirole (2011), the unique equilibrium is a cutoff equilibrium where all individual with type above v_s^* contribute, where the cutoff v_s^* is characterized by:

$$v_s^* = c - s - \mu\triangle(v_s^*)$$

with $\triangle(v_s) = E[v|v > v_s] - E[v|v < v_s]$. As described in Bénabou and Tirole (2011), the condition $1 + \mu\triangle'(v) > 0$ guarantees the unicity of the equilibrium. Overall, we obtain the following result:

Proposition 1 *If $1 + \mu\triangle'(v) > 0$, in the contribution stage after a sanction s has been voted, the unique symmetric Bayesian Nash equilibrium is such that players contribute if and only if they are of type v_i higher than a cutoff v_s^* defined by*

$$v_s^* = c - s - \mu\triangle(v_s^*) \tag{1}$$

The cutoff is increasing in c, decreasing in s and μ.

The cutoff is increasing in c, as a more costly contribution reduces the incentives to participate. The cutoff also decreases with the visibility of contribution (or taste of agents for reputation) μ since more pressure worsens the impact the stigma attached to free-riding and thus provides incentives to contribute. Finally, the voting cutoff is decreasing in s as a higher sanction raises the material cost of free-riding.

3.2 Voting stage

In the first stage, players have to vote whether to approve or not a given sanction s. If there is a majority of votes in favor, sanction s is implemented and the players use in the contribution phase a strategy with cutoff v_s^*, as derived above. On the contrary if a majority votes against the sanction, the players use in equilibrium a strategy with cutoff denoted v_0^*. Proposition 1 implies that $v_s^* < v_0^*$.

Given the behavior in the public good stage, and for a given value of s, agents can be grouped in three categories:

- The *never-participants* who do not contribute regardless of the outcome of the vote: agents with $v_i < v_s^*$.
- The *swing-participants* who contribute if and only if the sanction is voted: agents with $v_s^* \leq v_i \leq v_0^*$.
- The *always-participants* who always contribute regardless of the outcome of the vote: agents with $v_i > v_0^*$.

These different types of individuals have different motivations in voting. One common factor is that they all benefit from a positive externality if the sanction is voted, coming from a higher contribution of the other players in the second stage. Secondly, for the never participants, they can anticipate that if the sanction is implemented it will directly decrease their payoff since they will need to pay the sanction.

There is a third effect that is a key factor in voting: whether a sanction is voted or not will shape social norms. For those who always contribute, a sanction will decrease the honor they derive from doing the right thing since more types will tend to contribute in equilibrium. On the contrary, for the never participants, a sanction increases the stigma attached to not contributing. We now uncover how these different effects interact.

Consider the always participants (i.e players with $v_i > v_0^*$). By definition of this group, the outcome of the vote will not affect their own behavior, since regardless of their particular type v_i they will participate. When considering their voting decision, they thus simply tradeoff the externality gain that a sanction would bring against the decrease in the reputation payoff they obtain from doing an honorable act. It turns out that if e is low enough, the second effect dominates and they actually have an incentive to vote against the sanction.

Proposition 2 *For any sanction s, there exists a value $\overline{e}(s)$ such that if $e < \overline{e}(s)$, it is a weakly dominant strategy for the always-participants to vote against the sanction.*

We obtain the surprising result that individuals who in any case contribute to the public good, have an incentive to vote against a sanction that would force the others to participate as well. This is the case if the loss in reputation dominates the gains from increased contributions of the rest of the group. From a policy perspective this result is important. Note that even if the conditions for Proposition 2 are not met, the fact that this group always suffers from a loss of reputation if the sanction is passed, means that they have fewer incentives to lobby for regulation than what could be expected at first sight.[6]

However, when e is larger, the always participants do have an incentive to vote in favor. We now examine more in details the incentives of the different types to vote for or against the sanction, to determine which group is most likely to support the sanction. The never participants do not change their contribution decision even

[6] In our model there is of course no lobbying phase or stage where players can exert effort to influence legislation. If we added such a stage, the always-participants might have in fact little incentive to lobby for the regulation.

if the sanction is in place. However a sanction will mean a loss in reputation from the increased stigma of not contributing ($\mu(E[v_i|v_i < v_s^*] - E[v_i|v_i < v_0^*]$, which is negative since $v_s^* < v_0^*$), on top of the financial cost s. They however benefit from the expected externality gain obtained if the sanction is approved that we denote G. This externality gain is the difference between the externality obtained with a sanction and that obtained without. Overall, we find that for the never participants, the difference in expected utility comparing the situation with a sanction to the one without, that we denote $D(v_i)$, is given by:

$$D(v_i) = \mu \underbrace{(E[v_i|v_i < v_s^*] - E[v_i|v_i < v_0^*])}_{reputation\ loss\ <0} - s + G \tag{2}$$

For the always participants, as explained above, they suffer a loss of reputation that they need to tradeoff against the externality gain. Note that in equilibrium, the expected externality is the same for all types, and is equal to G as for the never participants. We explain later in detail how this measure is derived. For the always participants we have

$$D(v_i) = \mu \underbrace{(E[v_i|v_i > v_s^*] - E[v_i|v_i > v_0^*])}_{reputation\ loss\ <0} + G \tag{3}$$

Finally, for the swing participants, the decision involves an additional consideration. By definition swing participants contribute if and only if the sanction is in place. Therefore the actual type v_i will play a role. For this group we find:

$$D(v_i) = \mu(E[v_i|v_i > v_s^*] - E[v_i|v_i < v_0^*]) + v_i - c + G \tag{4}$$

We see that for all types, the difference in utilities $D(v_i)$ expressed in conditions 2-4 can be written as $D(v_i) \equiv \mathcal{R}(v_i) + G$, where $\mathcal{R}(v_i)$ is the difference in reputation cost and in direct financial costs, between the case with a sanction and the case without. Note that $\mathcal{R}(v_i)$ will always be negative.

What group has the most incentives to vote in favor of the sanction? the answer is not straightforward. Consider for instance the comparison between the never and always participants. It could a priori be the case that the loss in reputation for the always-participants be greater than for the never participants. This would be the case if $\triangle(v)$ was increasing in v, what Bénabou and Tirole (2011) describe as the case of strategic complements. We however show that in equilibrium, even if that were the case, the difference in reputation could not be greater than s.

Lemma 1 *In all symmetric perfect bayesian equilibria:*

- *If some of the never participants vote in favor of the sanction then all the swing participants and always participants vote in favor.*
- *If some of the swing participants vote in favor of the sanction, then all the always participants vote in favor.*
- *If a swing participant of type v_i votes in favor of the sanction, then all swing participants with type $v_i' > v_i$ vote in favor.*

Lemma 1 reflects the fact that the function $\mathcal{R}(v_i)$ is weakly decreasing in v_i. Types with a higher v_i have a relatively lower cost of having the sanction accepted. In Figure 1 we plot $-\mathcal{R}(v_i)$. The function $-\mathcal{R}(v_i)$ is flat for always and never participants and strictly decreasing for swing participants.

The result of Proposition 1, strongly suggests that symmetric perfect bayesian equilibria should be of the cutoff form, i.e equilibria characterized by a cutoff V^* such that a type votes in favor if and only if $v_i \geq V^*$. This is in fact correct only with a small additional constraint. Indeed there can exist other types of equilibria, due to the fact that all never participants (resp. always participants) face the same tradeoff in voting regardless of their particular type v_i. Without further constraints, there could therefore exist equilibria where all never participants are indifferent between voting in favor or against, some vote in favor and others against (and all swing and always participants vote in favor). In such an equilibrium the identity of those never participants who vote against is not a priori uniquely pinned down.[7] We thus impose the following additional restriction that states that if two types are indifferent in their voting decision then, if the lower type votes in favor, so does the higher type.

Restriction A (tie breaking): *If in equilibrium two types $v_i > v_i^{'}$ are indifferent between voting in favor or against the sanction, then if type $v_i^{'}$ votes in favor, so does type v_i.*

Note that restriction A, which appears minimal, could be rephrased differently: it would be equivalent to assume that a vote for the sanction gives an additional

[7]The specific types of never participants that vote against can be chosen freely under the constraint that the condition of indifference between voting for and against is indeed satisfied.

benefit proportional to the type, ϵv_i, with ϵ very small, and this independently of the behavior in the second phase. When we impose Restriction A, we find:

Proposition 3 *Under Restriction A,*

- *All symmetric perfect bayesian equilibria are cutoff equilibria, characterized by a cutoff V^*.*
- *There is at most one equilibrium where the cutoff belongs to the never participants group (resp. always participant group), i.e $V^* < v_s^*$ (resp $V^* > v_0^*$).*

Proposition 3 establishes that, under Restriction A, all symmetric perfect bayesian equilibria are of the cutoff form. It also shows that even with this restriction, there potentially remains a multiplicity of equilibria. This multiplicity however has a nice economic interpretation that we develop below.

To understand better the intuition of these results, we have to understand what determines the value of the expected externality gain G, which is of course determined in equilibrium. For a given equilibrium with cutoff V^*, G takes a unique value, identical for all groups. However, this value G is not necessarily monotonic in V^*.

As in the literature on information aggregation in voting (Austen-Smith and Banks 1996, Feddersen and Pesendorfer 1996, 1997 and 1998), voters consider only the case where their vote is pivotal. In equilibrium the voter can infer additional information about the distribution of types. In the simple majority voting case, a player is pivotal when there are exactly N yes-voters and N no-voters among the $2N$ other players. Recall however that to determine the expected externality gain from having a sanction, each voter only needs to determine the expected number of swing participants. Indeed they are the only types that change behavior based on whether the sanction is approved or not and they thus determine the added value of having a sanction in place.

Consider the case where the cutoff is among the never participants, i.e $V^* \leq v_s^*$. In this case, the no-voters are necessarily never participants, while some of the yes-voters are going to be swing participants. In fact this proportion is the expected probability of being a swing participant $F(v_0^*) - F(v_s^*)$ conditional on being a yes voter, probability $1 - F(V^*)$. Thus for a given cutoff V^*, the expected externality is given by:

$$G(V^*) = \frac{1}{2}e\left[\frac{F(v_0^*) - F(v_s^*)}{1 - F(V^*)}\right]$$

This expected externality is strictly increasing in V^*: indeed, if the cutoff is higher, it becomes more likely that the yes voters are indeed swing participants. This implies the result in Proposition 3 that there can be only one equilibrium in this zone (since $\mathcal{R}(v_i)$ is flat in this zone as illustrated in Figure 1).

If the cutoff is in the swing participant group, the calculation of G is slightly more intricate. No-voters can in this case either be swing participants or never participants. Specifically, given a voting cutoff V^*, the probability that a no voter is a swing participant is given by $\frac{F(V^*)-F(v_s^*)}{F(V^*)}$. As V^* increases, it becomes more likely that a no voter is in fact a swing participant. On the other hand, the probability that a yes voter is a swing participant (and not an always participant), is given by $\frac{F(v_0^*)-F(V^*)}{1-F(V^*)}$. As V^* increases, it becomes less likely that the yes voter is a swing participant. Overall, the expected externality gain is thus given by the following expression:

$$G(V^*) = \tfrac{1}{2}e\left[\left(\tfrac{F(V^*)-F(v_s^*)}{F(V^*)}\right) + \left(\tfrac{F(v_0^*)-F(V^*)}{1-F(V^*)}\right)\right]$$

The effect of an increase in V^* on the expected externality gain is ambiguous when V^* belongs to the swing participant group. As V^* increases, it becomes more likely that no voters are swing participants but less likely for yes voters, leading to an ambiguous conclusion on the expected number of swing participants, and thus on the size of the externality. This fact leads to the surprising feature that there could be multiple equilibria with $V^* \in (v_s^*, v_0^*)$, i.e the cutoff is among the swing-participants. We call equilibria in this zone *interior equilibria*.

We illustrate this in Figure 1. We plot both the decreasing function $-\mathcal{R}(v_i)$ and the function $G(V^*)$ for the case where f is uniform.[8] There are two equilibria with cutoffs V_1 and V_2 that corresponds to the intersection of function $-\mathcal{R}(v_i)$ and the function $G(V^*)$. Indeed the cutoff is characterized by the fact that the cutoff type V^* is such that $\mathcal{R}(V^*)+G(V^*) = 0$. In equilibrium with cutoff V_1, the pivotal voter expects a large portion of yes voters (to the left of V_1) and of no voters (to the right of V_1) to be swing participants. The expected externality is thus large and justifies

[8]Note that the x axis is v_i for $\mathcal{R}$ and V^* for G.

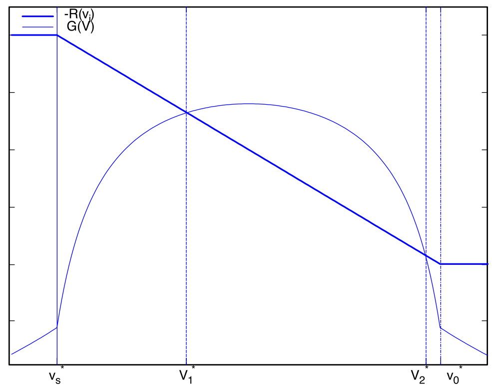

Figure 1: Multiple equilibria

the low voting cutoff. On the contrary in the case of V_2, it is very unlikely that the yes voters are swing participants, the expected externality is thus lower, justifying the higher cutoff.[9]

These different equilibria can be understood as corresponding to different norms of voting. A norm of opposition to sanctions (high cutoff) might prevail and would be based on an expectation of low externality gain. There could also exist norms of voting more favorable to sanctions (lower cutoff) based on an expectation of a high externality gain. Both these norms would be self sustained due to the mechanisms of information aggregation described above.

Figure 1 shows a case with two equilibria but the multiplicity can be larger in general. We present below conditions (satisfied in the case of the uniform distribution presented in Figure 1), where this multiplicity is reduced.

Proposition 4 *Under Restriction A, If $\frac{f}{1-F}(v)$ is weakly increasing and $\frac{f}{F}(v)$ is weakly decreasing,*

- *there are at most two symmetric perfect bayesian equilibria*
- *there is a unique stable interior equilibrium*

[9]This outcome with multiple equilibria is an interesting feature of our model and, to the best of our knowledge, is not present in the literature on aggregation of information in voting. Consider the classic case where voters get information on an underlying state of the world and the expected payoff is increasing in this state. The type of a voter is the signal he obtains, and like in our model the equilibrium will take the form of a cutoff strategy. In this case, the expected payoff is increasing in the cutoff: a higher cutoff means the information obtained when pivotal indicates a higher state of the world. In our model there can be non monotonicities as suggested above.

The conditions $\frac{f}{1-F}(v)$ is weakly increasing and $\frac{f}{F}(v)$ is weakly decreasing, guarantee that the externality gain $G(V^*)$ is a weakly concave function on the interval (v_s^*, v_0^*). Thus there can be at most two equilibria. Furthermore, an interior equilibrium (defined above as $V^* \in (v_s^*, v_0^*)$) is stable if and only if $G(V)$ is increasing at the equilibrium cutoff.

3.3 Robustness

We discuss briefly the robustness of our results to some extensions.

3.3.1 Case of bonuses

Our main model focuses on sanctions for free riding. Naturally, a bonus b for contributions gives very similar results. In the contribution stage, the three same categories will emerge as in Proposition 1, with the cutoff v_0^* unchanged and a cutoff $v_b^* < v_0^*$ in the case with a bonus.

The results of Proposition 2 would still hold, to the extent that we constrained the bonus not to be too large.[10] The result then appears even more forceful: for moderate bonuses, it can be a dominant strategy for group members who would obtain this bonus to vote against the rule implementing them as they might lose too much in reputation. Finally results of Lemma 1 and Proposition 3 would hold and the voting stage would be characterized by a voting cutoff. To preserve the clarity of the presentation, we keep studying, for the rest of the text, the case of sanctions.

3.3.2 Complementarities in public good provision

We considered up till now environments where there were no complementarities between the different individual contributions to the public good. We now briefly show that the results extend to the case of complementarities. Complementarities should be common, in the case of team work for instance. We assume that for given vector of contributions a, the level of the public good is given by $g(a_i, a_{-i})$ with g supermodular in the case of strategic complements and g submodular for substitutes. The utility of agent i is given by:

$$U_i = (v_i - c)a_i - s(1 - a_i) + eg(a_i, a_{-i}) + \mu E[v_i|y_i]$$

[10] Indeed very large bonuses would dominate the loss in reputation.

In this case we show in Appendix B1 that both the contribution and voting behaviors are still characterized by cutoffs. The determination of the cutoffs is more involved in particular because the result of the vote in the first stage provides information on how cooperative the group is and thus affects contributions in the second stage because of complementarities.

4 Changing environments

In this section, we examine how changes in the regulatory environment affects voting and contributions. The designer (leader/manager of the organisation for instance), can affect the level of the sanction s and potentially the details of the voting rule. We return to our main model with no complementarities and focus on environments where the conditions of Proposition 4 hold.

4.1 Changing the level of sanction s

We first examine how the level of s will affect the outcome of the vote. In the public good phase, the level of s will affect the equilibrium size of the groups. As s increases, more never participants become swing participants, i.e as s increases v_s^* decreases.[11] This will of course affect indirectly the equilibrium of the voting stage.

In the voting stage, there will be two effects. First, the expected size of the externality increases since the size of the swing participant group increases with s. However there is a countervailing second effect which is the loss in reputation. These two effects go in opposite directions. The first effect, i.e the externality effect, increases the incentives to vote in favor, while the second, the reputation effect, decreases it. Since the first is proportional to e and the second is proportional to μ, we find the following result:

Proposition 5 *There exists a benchmark $\overline{b}(s)$ such that in the unique stable equilibrium, the voting cutoff in equilibrium V^* is decreasing in the level of sanction s if and only if $\frac{e}{\mu} > \overline{b}(s)$.*

Increasing the sanction can increase the propensity to vote in favor if the externality effect dominates the loss of reputation, i.e e is large compared to μ. If this

[11] Note however that v_0^* is not affected, i.e the size of the group of always participants is not changed.

is not the case, the effect will be reversed. This information is key for a designer wanting to minimize the sanction conditional on approval.

4.2 Voting rule

Another important feature of the voting process is the supermajority chosen for the voting rule. Up till now we have considered the case of majority rule. We now consider institutions that use different rules and require strictly more than K votes in favor to approve the sanction.

The required majority does not impact the contribution stage or the general structure of the equilibrium in the voting phase. It however affects the information aggregated when an individual is pivotal. We find that this rule has a surprising effect on the outcome of the vote.

Proposition 6 *There exists $\widehat{V} \in (v_s^*, v_0^*)$ such that, in all stable equilibria, the voting cutoff V^* is*

- *decreasing in K if $V^* \leq \widehat{V}$*
- *increasing in K if $V^* \geq \widehat{V}$*

Consider an equilibrium such that $V^* < v_s^*$, in other words the pivotal voter is a never-participant. This pivotal voter, to calculate the expected externality, needs to build an expectation on the number of swing participants. Only the yes-voters can be swing participants in such an equilibrium. As K is increased, the number of yes voters is higher when pivotal and thus the expected size of the externality is higher, which makes the pivotal voter more inclined to vote in favor (i.e V^* decreases). In this case a more stringent voting rule, namely an increase in the supermajority rule, makes people more inclined to vote in favor.

In fact there are instances where increasing the supermajority rule can increase the probability of approval. We present such a case in Appendix B2 for a uniform distribution. We thus get the surprising outcome where increasing the number of votes necessary for approval can in fact facilitate approval.

4.3 Keep contributions secret?

The designer in certain instances might have the option to keep the individual contributions secret in the second phase. We examine in this section what are her incentives to do so, if her objective is to maximize contributions.

We first have to derive the equilibrium when both votes and contributions are kept secret, an environment we call *both secret.* To clarify the comparisons, we focus on the case where types are uniformly distributed over $(0, 1)$. In the public good stage members use higher cutoffs than when contributions are public, $\hat{v}_s = c - s > v_s^*$ and $\hat{v}_0 = c > v_0^*$. The fact that they no longer suffer from bad reputation when they do not contribute decreases overall contributions. This would suggest that the designer would want to make contributions public. However, we still need to verify that keeping them secret would not make members more inclined to vote for the sanction (supposing the designer can find ways to commit to keep individual contributions secret in the second phase). We thus examine the voting stage.

In the voting stage, the tradeoff is very similar to the one expressed above: the voters will tradeoff the cost of contributing (that we denote $-\widehat{\mathcal{R}}(v_i)$) against the expected externality gain $\widehat{G}(v_i)$. The functions $-\mathcal{R}$, $-\widehat{\mathcal{R}}$, G and $\widehat{G}$ are plotted in Figure 4. The financial and reputational cost is higher for always and never participants to vote in favor in the *public contributions* environment compared to the *both secret* (i.e at the extremities, for $v < v_s^*$ or $v > \hat{v}_0$, $-\mathcal{R}$ is above $-\widehat{\mathcal{R}}$). Consider for instance the case of never participants: by voting in favor they will not only incur the financial cost s but will also suffer from the increased bad reputation, a concern not present when contributions are kept secret. For the intermediate zone, and in particular in the zone of swing-participants for both (i.e $V \in (\hat{v}_s, v_0)$),[12] the cost is higher in the *both secret* environment. Indeed image concerns make swing participants more inclined to vote in favor as they benefit from the good reputation of contributing when the sanction is passed.

We now compare the expected externality gains G and $\widehat{G}$. We see a pattern emerge in Figure 4. For low values of $V < v_s^*$, where types are never participants in both cases, $G(V) = \widehat{G}(V)$, and similarly for high values $V > \hat{v}_0$. Then for V slightly above v_s^*, $G(V)$ starts increasing faster, as some of the no voters can now be swing participants in the *secret vote* environment. It eventually decreases as V approaches v_0^*, while $\widehat{G}(V)$ starts increasing. We find that in the uniform case, there is a single intersection between the two curves. This feature turns out to be more general as shown in the following result.

Lemma 2 *If f is uniformly distributed over $(0, 1)$ and if $\hat{v}_s < \frac{1}{2} < v_0^*$, then if $V < \frac{1}{2}$, $G(V) \geq \widehat{G}(V)$ and if $V > \frac{1}{2}$, $G(V) \leq \widehat{G}(V)$*

[12] This zone does not necessarily exist, in particular if μ is high enough, we will have $v_0 < \hat{v}_s$.

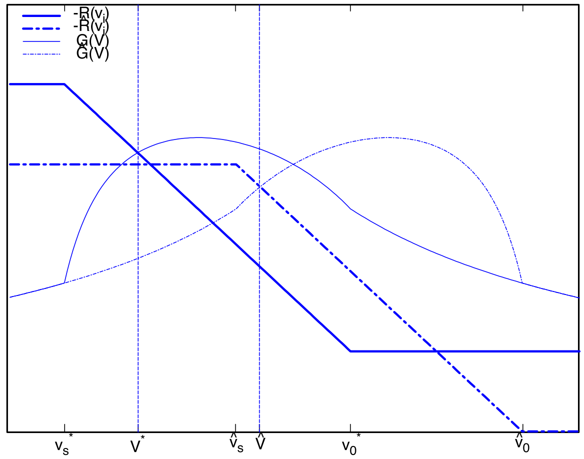

Figure 2: Comparing the cases *both secret* and *public contributions*

We can now conclude on whether the designer who wants to maximize contributions should commit to keep contributions secret in the second stage. In Figure 4 we illustrate a case where $V^* < \widehat{V}$. In this case, the designer should clearly not commit: in the first phase members are more inclined to vote in favor of a sanction and in the second stage they are more likely to contribute. We provide in the next proposition general conditions for this pattern to emerge.

Proposition 7 *If f is uniformly distributed over $(0,1)$ and if $\hat{v}_s < \frac{1}{2} < v_0^*$, there exists e_l and e_h, such that if $e \in (e_l, e_h)$, the sanction is more likely to be adopted under public votes than under both secret, i.e $V^* < \widetilde{V}$.*

5 Public voting

We now consider the *public vote* environments, i.e situations where the individual votes are public but the contribution decisions are not observed by the group members. This could be the case if the individual contribution to the public good are hard to identify, which is often the case when it is the outcome of team work. Of course, to impose the sanctions, the designer/manager will have to observe individual actions. We thus consider cases where he commits or is unable to credibly make the actions public. Note that in the case of bonuses in firms, which is typically a reward corresponding to a team work, individual bonuses are not revealed to other members. There can exist in groups resistance to the public disclosure of sanctions or bonuses.

In these cases, the reputation of individuals will therefore be based on their votes.[13] In the contribution phase, given a sanction s, individual i contributes if and only if $v_i > c - s$. As in the previous sections, there will be three categories in equilibrium, always participants, swing participants and never participants. However the cutoffs between these categories do not include a reputation component anymore: $v_0^* = c$ and $v_s^* = c - s$.

The voter cares not only about the event where his vote is pivotal, but also about the other events since his vote will always be observed, and other players will make inferences on his type based on it. All voters care in the same way about reputation, so would vote the same way if they knew they were not pivotal. However, in the pivotal case, those with lower v_i will have higher incentives to vote against the sanction.

We denote Piv the event of being pivotal when voting and $\triangle^* \equiv E[v_i|b_i = 1] - E[v_i|b_i = 0]$ the reputation derived in equilibrium from voting in favor of the sanction rather than against. Both these measures are determined in equilibrium and do not depend on the individual types of players.

In equilibrium, the net benefit of the never participants to vote in favor of the sanction is given by:

$$\mu\triangle^* + P\left[Piv\right](-s + G)$$

for swing participants

$$\mu\triangle^* + P\left[Piv\right](v_i - c + G)$$

and for always participants

$$\mu\triangle^* + P\left[Piv\right](G)$$

where G is the expected externality gain from the sanction, conditional on the event of being pivotal.

As in the environments studied in the previous sections, under Restriction A, all symmetric perfect bayesian equilibria are of the cutoff form.

[13] We consider the case of majority voting.

Lemma 3 *Under Restriction A, all symmetric perfect bayesian equilibria are cutoff equilibria, characterized by a cutoff V^*.*

Lemma 3 reveals that under the same minor restriction as before, all symmetric equilibria are of the cutoff form. As in the case of observable actions and unobservable votes, the problem is characterized by a large multiplicity of equilibria. One of the reasons is non-monotonicities in the expected externality gain G as a function of V^*, as in the case of *public contributions*. There are however in this setting additional sources of multiplicity.

The first is due to the reputation term $\triangle^*$. Given that we are faced with cutoff equilibria, this term can be written $\triangle^* = E[v_i|v_i > V^*] - E[v_i|v_i < V^*] \equiv \triangle(V^*)$. As described in detail in Benabou and Tirole and in Jewitt, the function $\triangle(V)$ is not necessarily montonic. These nonmontonicities are the first source of additional mutliplicity of equilibria.

The second is more interesting and is present even when we abstract from the externality effect. Consider the case where $e = 0$. For a given N, the probability of being pivotal is not monotonous in V^*. For V^* lower than the median, the probability of being pivotal is increasing in V^* while it is decreasing otherwise. There could thus be two potential equilibria: one where V^* is low and one where it is close to 1/2. In the first case, players have high incentives to vote in favor because the probability of being pivotal is low and the reputation effect dominates, thus justifying the low V^*. This is a self sustaining norm of general support for sanctions, self sustained because the chance of pivotal is small if everyone votes in the same way. In the second equilibrium (different norm), players have more incentives to vote against (higher V^*) as this higher V^* is coherent with a higher probability of being pivotal.

We have shown that the key tradeoff in the case of public voting is between the reputation derived from voting in favor and the effect of the vote on the outcome in the event the vote is pivotal. It is thus natural to ask the question of what occurs for large organizations when the chance of being pivotal decreases. We find that any level of sanction will be adopted in equilibrium for large assemblies. We need to focus on sequential equilibria to eliminate the unreasonable equilibrium where everyone votes against the sanction and the belief, off the equilibrium path when someone votes in favor, is that the deviator has a very low type (i.e $\triangle^* < 0$).[14]

[14]such an equilibrium is not sequential since if voting in favor happens with some probability on

Proposition 8 *If $\Delta(V^*)$ is decreasing in V^*, then an increase in the size of the organization N will decrease the probability of acceptance.*

However, when $N \to +\infty$, in all sequential equilibria of the game where only votes are observable, any sanction $s > 0$ is approved with probability converging to one.

The intuition when N becomes very large is quite clear. Low types vote against the sanction when the probability of being pivotal is high enough that it compensates for the loss of reputation. When N becomes large, the probability of being pivotal goes to zero and the proportion of people ready to vote against shrinks. Note that this is independent of the level of the sanction and in particular, it could well be the case that the sanction decreases total welfare. This result generalizes the result of Feddersen, Gailmard and Sandroni (2009) who consider the case where one alternative is exogenously given as the ethical outcome.

The first result of Proposition 8 however shows that this is only a limit result, and in a sense qualifies the finding of Feddersen, Gailmard and Sandroni (2009). Increasing N can actually decrease the probability of acceptance of the sanction in the case where $\triangle(V)$ is a decreasing function at V^*.

6 Observing votes and contributions

In small committees it is often the case that the contribution decision is observable, or as discussed in the previous section, that the designer is unable or unwilling to keep the contributions secret. Therefore, if the institutional choice was to choose public voting, the inferences made by other group members about an individual's type, will be based both on her vote and contribution choices.

To clarify the forces at play, we focus on the case where the sanction s submitted to a vote is large, $s > v_{min} - c$, so that if the sanction is voted, all types participate. Consider the case where the sanction was not voted in the first stage of the game. The behavior of the players will depend on the way they voted in the first phase. Conditional on a vote, the players will choose a cutoff strategy. We denote the cutoff $v_0(1)$ for the players who voted in favor of the sanction in the first phase and $v_0(0)$ for those who voted against.

---

the equilibrium path, then given that the incentives to vote in favor are increasing in the type, the reputation $\triangle^*$ attached to a yes-vote should necessarily be positive.

Using the notation $E_0(b,a) = E[v_i|(b_i = b) \cap (a_i = a) \cap (s = 0)]$ (for instance $E_0(1,0)$ is the expected value of v given that the player voted for the sanction, the sanction was not passed and he did not participate), the cutoff is defined by

$$v_0(i) = c - \mu [E_0[i,1] - E_0[i,0]]$$

There is no clear ordering between $v_0(0)$ and $v_0(1)$. On the one hand, those who already sent a bad signal by voting against the sanction, might have little to lose by not participating. On the other hand, those who already voted for the sanction, can afford to send a bad signal of not participating. The ranking will depend on inferences made in equilibrium.

Proposition 9 *In the voting phase:*

1. *There exist equilibria where $v_0(0) < v_0(1)$ and others where the ordering is reversed*

2. *The equilibrium voting strategy is not necessarily a cutoff strategy. In particular there exists an equilibrium with two cutoffs $\underline{v}$ and $\overline{v}$ such that:*

    - *for $v < \underline{v}$, players vote against the sanction and do not participate, i.e choose (0,0)*
    - *for $\underline{v} \leq v \leq \overline{v}$, players vote for the sanction and do not participate (1,0)*
    - *for $v > \overline{v}$, players vote against the sanction and participate (0,1)*

Proposition 9 presents properties of equilibria such that certain types vote in favor of the sanction and some against. The full set of equilibria is described in the proof. The first key property is that the voting strategy is not necessarily of the cutoff form. In particular there is an equilibrium where low types vote against and do not participate, intermediate types vote for and do not participate and high types vote against and participate. In this equilibrium, the low types do not want to vote in favor because contributing is too costly and they consider the case where their vote can be pivotal. The intermediate types are ready to take the risk of voting in favor and potentially losing if pivotal, to benefit from the increased reputation. They however do not want to deviate to action (0,1) since contributing is still too costly.

7 Discussion and examples

We conclude by discussing potential avenues for testing empirically our results, both with field and lab data, and incidentally give some further exemples of applications of our setup.

We first return to the case, mentioned in the introduction, of the French Assemblee Nationale voting a law imposing sanctions for not attending.[15] Following the vote and the introduction of the sanction, participation significantly increased: from an average of 7 meetings attended per year, it jumped to 19 meetings. There was however a large heterogeneity in reactions. Some people still did not participate. The maximum fine imposed was 4615 euros for someone failing to attend a single meeting.

This setting with elected members of parliament voting on rules that regulate their activity appears to be a very good environment to test our theoretical results. Data is available: for instance we can observe both voting and contribution behaviors. Furthermore, the level of observability evolved. Initially in 2010, the members did not expect their presence to be publicly revealed, but an independent association, called *Regards Citoyens*, decided to make public the information. The participation behavior can thus initially be considered as partially secret since the parliamentaries were not aware that the data would be published.

There are however several specific features that we need to highlight. First, there is a tendency, particular strong in France, for parliamentaries to follow party lines, making it harder to identify individual decisions. Second, the members of parliament do not care only about their image within the group, but also in the eyes of their electorate.[16]

Even though there is strong party discipline, we code a party member abstaining or not coming to vote as someone voting against the party line. For instance in this particular case, the right wing party (UMP) called to vote in favor of the proposal, while the left called to vote against. We consider those UMP members who did not vote as voting against the proposal.[17] Among the top 7 shirkers in 2010, 3 voted against the sanction and out of the four UMP ranking in the top 7, three were not

[15] If they missed more than 2 sessions per month, their compensation for these meetings would be reduced by 25 percent (representing more or less 353 euros).

[16] Both these dimensions are compatible with our preference structure.

[17] In fact, the final result turned out to be quite close, with 312 voting in favor when a majority of 266 was required.

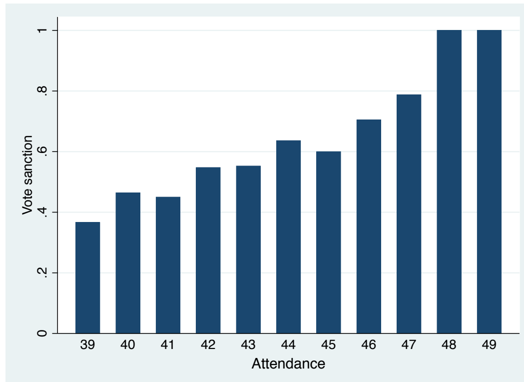

Figure 3: Attendance and voting on sanctions

present on voting day.

We can go further and examine the link between voting behavior and attendance in the general assembly. This is presented in Figure 3. We see a clearly increasing trend indicating that those with a higher v_i are more likely to vote against the sanction, consistent with our result that equilibria will be of the cutoff form.

This exemple suggests one possible avenue for testing some of our results. Ideally data could also be obtained from the functioning of smaller groups where the image concern would only be relevant within the group (whereas politicians care also about the image their constituents have of them). A different direction would be to use experimental data.[18]

In fact our paper can inform some of the existing experimental evidence in the literature, in particular the results in Tyran and Feld (2006). The authors examine a two stage game very similar to ours, involving 3 players. The second stage is a public good provision game. In some treatments a sanction for free riding is exogenously imposed, while in others it is voted by the group. We focus on the results they obtain in the mild sanction case, i.e a sanction that would not deter a purely rational player from free riding.

The authors show that contributions in the public good games are significantly higher following a sanction endogenously chosen versus exogenously set. They suggest one explanation based on selection: if participants vary in their generosity (i.e our v_i), those cases where the sanction is adopted are cases where the group is

[18]There are now new procedures to measure individual sensitivity to image concerns.

overall more generous. They discard this explanation based on the following two additional results.[19] Yes and no voters contribute more or less in the same way in the public good game that follows. Yes and no voters contribute significantly less if the sanction is not imposed than if it is imposed. The authors claim that if selection was an issue, yes voters should contribute the same regardless of whether they are in a group that accepted vs rejected the sanction.

Our analysis suggests a more nuanced view. Consider the case where most subjects in the lab are swing participants. If that were the case, by definition of swing participants, they would behave differently if the group adopted or rejected the sanction. Moreover, both yes and no voter if they are swing participants, even though they vote differently, would behave in the same way later on. The two facts mentioned above are thus not incompatible with selection. However, if most subjects are actually swing participants, we should not observe a significant difference between exogenous and endogenous sanctions, which might point to the fact that additional forces as those proposed by Tyran and Feld might be at play, including conditional cooperation. Understanding how this modifications of preferences affects our results, could be the object of interesting future work.

[19]Shown in Table 3 of their paper.

REFERENCES

Andreoni, J. 1990. "Impure altruism and donations to public goods: a theory of warm-glow giving." *The economic journal*, 464-477.

Andreoni, J. and R. Petrie (2004): "Public goods experiments without confidentiality: a glimpse into fund-raising", *Journal of Public Economics*, 88, 1605-1623.

Ariely, D., Bracha, A. and S. Meier. 2009. "Doing good or doing well? Image motivation and monetary incentives in behaving prosocially." *The American Economic Review*, 544-555.

Austen-Smith, D. and J. S. Banks. 1996. "Information aggregation, rationality, and the Condorcet jury theorem." *American Political Science Review*, 34-45.

Bénabou, R. and J. Tirole. 2003. "Intrinsic and extrinsic motivation." *The Review of Economic Studies*, 70(3), 489-520.

Bénabou, R. and J. Tirole. 2006. "Incentives and Prosocial Behavior." *The American Economic Review*, 1652-1678.

Bénabou, R. and J. Tirole. 2011. "Laws and norms"(No. w17579). *National Bureau of Economic Research.*

Callander, S. 2008. "Majority rule when voters like to win." *Games and Economic Behavior*, 64, 393420.

Dana, J., D. M. Cain, and R. M. Dawes (2006): "What you don't know won't hurt me: Costly (but quiet) exit in dictator games", *Organizational Behavior and Human Decision Processes*, 100, 193-201.

De Condorcet, N. 1785. "Essai sur l'application de l'analyse la probabilit des dcisions rendues la pluralit des voix."

DellaVigna, S., J. A. List, and U. Malmendier (2012): "Testing for altruism and social pressure in charitable giving," *The Quaterly Journal of Economics*, 127, 1-56.

Feddersen, T. and W. Pesendorfer. 1996. "The swing voter's curse." *The American economic review*, 408-424.

Feddersen, T. and W. Pesendorfer. 1997. "Voting behavior and information aggregation in elections with private information." *Econometrica*, 1029-1058.

Feddersen, T. and W. Pesendorfer. 1998. "Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting." *American Political Science Review*, 23-35.

Galbiati, R., Schlag K. and J. van der Weele (2013), "Sanctions that Signal: An Experiment" *Journal of Economic Behavior and Organization,* 94, 3451.

Galbiati, R. and P. Vertova (2008) "Obligations and Cooperative Behaviour in Public Good Games" *Games and Economic Behavior*, 64(1), 146-170.

Godefroy, R. and E. Perez-Richet. 2013 "Choosing Choices: Agenda Selection With Uncertain Issues." *Econometrica*, 81(1), 221-253.

Henry, E. and J. Sonntag, 2015. "Measuring image concerns." *CEPR DP10831*.

Levy, G. 2007. "Decision making in committees: Transparency, reputation, and voting rules." *The American Economic Review*, 150-168.

Midjord, R., T. Rodriguez and J. Valasek 2013 "Over-Caution of Large Committees of Experts. " *working paper*

Ottaviani, M. and P. Sorensen(2001), "Information aggregation in debate: who should speak first?" *Journal of Public Economics*, 81, 393-421.

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

Samek, A. S. and R. M. Sheremeta (2013): "Recognizing contributors: an experiment on public goods", *Experimental Economics*, 17, 673-690.

Tyran, J. R. and L. P. Feld. 2006. "Achieving Compliance when Legal Sanctions are Non-deterrent." *The Scandinavian Journal of Economics*, 108(1), 135-156.

APPENDIX

Proposition 1

As derived in the main text, the contribution cutoff for a given sanction s is given by:

$$v_s^* = c - s - \mu\triangle(v_s^*)$$

with $\triangle(v_s) \equiv E[v|v > v_s] - E[v|v < v_s]$.

Let $F(v) = \mu\triangle(v) + v - c + s$. The equation characterizing v_s^* can be rewritten $F(v_s^*) = 0$. Furthermore, we have $F'(v) = 1 + \mu\triangle'(v)$. Thus, the condition $1 + \mu\triangle'(v) > 0$ guarantees the unicity of the equilibrium as in Bénabou and Tirole (2011).

To prove the comparative statics, we use the implicit function theorem, yielding:

$$\frac{\partial v_s^*}{\partial s} = -\frac{1}{1 + \mu\triangle'(v_s^*)} < 0$$
$$\frac{\partial v_s^*}{\partial c} = \frac{1}{1 + \mu\triangle'(v_s^*)} > 0$$
$$\frac{\partial v_s^*}{\partial \mu} = -\frac{\triangle(v)}{1 + \mu\triangle'(v_s^*)} < 0$$

Proposition 2

Denoting G_s the expected externality obtained if the sanction is approved and G_0 the expected externality obtained if the sanction is not approved, we denote $G = G_s - G_0$ as in the main text, the expected externality gain from sanctions, which is the same for all players regardless of their type.

The expected payoff of always participants if the sanction is approved is:

$$v_i - c + \mu(E[v_i|v_i > v_s^*]) + G_s$$

If it is rejected:

$$v_i - c + \mu(E[v_i|v_i > v_0^*]) + G_0$$

The difference between the payoffs is thus:

$$D(v_i) = \mu(E[v_i|v_i > v_s^*] - E[v_i|v_i > v_0^*]) + G \tag{5}$$

$D(v_i)$ is independent of the particular value of v_i and furthermore we have $E[v_i|v_i > v_s^*] - E[v_i|v_i > v_0^*] < 0$ since $v_s^* < v_0^*$. In other words, more players participate following the sanction so there is less value of reputation from participating.

The maximum value for G is achieved when all other players contribute if and only if the sanction is passed, i.e $G_s = G = e$ and $G_0 = 0$. Therefore, rewriting equation (5) using the fact established above that $G \leq e$, we have:

$$D(v_i) \leq \mu(E[v_i|v_i > v_s^*] - E[v_i|v_i > v_0^*]) + e$$

For a given s, define $\overline{e}(s)$ as:

$$\overline{e}(s) = -\mu(E[v_i|v_i > v_s^*] - E[v_i|v_i > v_0^*])$$

We thus conclude that for all $e \leq \overline{e}(s)$ it is a best response for the always participants to vote against the sanction, even if the externality gain is at its maximum. We thus conclude that for $e \leq \overline{e}(s)$, voting for the sanction is a weakly dominated strategy.

Lemma 1

As derived in the main text, in equilibrium, never participants vote in favor of the sanction if and only if:

$$\mu(E[v_i|v_i < v_s^*] - E[v_i|v_i < v_0^*]) + G - s \geq 0 \tag{6}$$

swing participants vote in favor if and only if:

$$\mu(E[v_i|v_i > v_s^*] - E[v_i|v_i < v_0^*]) + G + v_i - c \geq 0 \tag{7}$$

and always participants vote in favor if and only if:

$$\mu(E[v_i|v_i > v_s^*] - E[v_i|v_i > v_0^*]) + G \geq 0 \tag{8}$$

Taking the difference between the left-hand side of condition (7) and (6), we find that this difference equals:

$$\begin{aligned} & \mu\left(E[v_i|v_i > v_s^*] - E[v_i|v_i < v_s^*]\right) + v_i - c + s \\ = \quad & \mu\triangle(v_s^*) + v_i - c + s \end{aligned}$$

Since v_i corresponds to the type of swing-participants, we have $v_i \geq v_s^*$ and by definition of v_s^* given in Proposition 1, we have $v_i \geq c - s - \mu\triangle(v_s^*)$. So that

$$\mu\triangle(v_s^*) + v_i - c + s \;\geq\; 0$$

We thus conclude that if condition (7) is satisfied then condition (6) will also be satisfied.

Similarly, using the fact that v_i for swing participants is such that $v_i \leq v_0^*$, we can show that if condition (8) is satisfied then condition (7) will also be satisfied. This establishes the first two points of the lemma. The last point directly follows using the fact that for the swing participants, the incentive to vote in favor of the sanction is strictly increasing in v_i (i.e the left hand side of condition (7) is increasing in v_i).

Proposition 3

The first result of Proposition 3 directly follows from Lemma 1 and using in addition restriction A to guarantee that if a never participant (resp. always participant) votes in favor, then all never participants (resp. always participant) with higher types will vote in favor.

The second result directly follows from the derivations in the main text. For $V < v_s^*$, we have that $\mathcal{R}(V)$ is constant and $G(V)$ is strictly increasing. An equilibrium is characterized by the intersection of $-\mathcal{R}(V)$ and $G(V)$. Thus, there is at most one equilibrium where $V^* < v_s^*$. The same reasoning applies for the always participants.

Proposition 4

We first show that under the conditions of Proposition 4, $G(V)$ is concave for $V \in (v_s^*, v_0^*)$. In this region, we have

$$G(V) = \tfrac{1}{2}e\left[\left(\tfrac{F(V)-F(v_s^*)}{F(V)}\right) + \left(\tfrac{F(v_0^*)-F(V)}{1-F(V)}\right)\right]$$

Thus, the derivative is given by

$$G^{'}(V) = \tfrac{1}{2}e\left[\left(\tfrac{f(V)F(V)-f(V)(F(V)-F(v_s^*))}{(F(V))^2}\right) + \left(\tfrac{-f(V)(1-F(V))+f(V)\left(F(v_0^*)-F(V)\right)}{(1-F(V))^2}\right)\right]$$

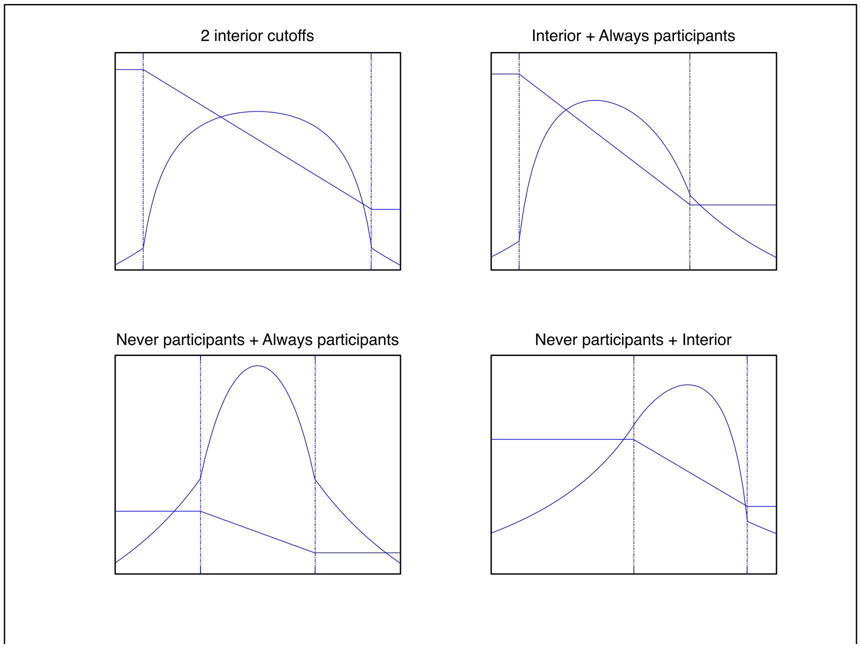

Figure 4: Four possible outcomes

Looking at the second derivative, we have

$$G^{''}(V) = \frac{1}{2}e\left[\left(\frac{F(v_s^*)F(V)\big(f'(V)F(V)-2(f(V))^2\big)}{(F(V))^4}\right) + \left(\frac{(1-F(v_0^*))(1-F(V))\big(-f'(V)(1-F(V))-2(f(V))^2\big)}{(1-F(V))^4}\right)\right]$$

Thus, a sufficient condition to establish $G^{''}(V) < 0$ is that

$$f'(V)F(V) - 2(f(V))^2 \leq 0 \tag{9}$$

and

$$-f'(V)(1 - F(V)) - 2(f(V))^2 \leq 0 \tag{10}$$

This two conditions are implied by the conditions introduced in Proposition 3 namely $\frac{f}{1-F}(v)$ is weakly increasing implies condition 10 and $\frac{f}{F}(v)$ is weakly decreasing implies condition 9 .

We thus establish that under these conditions, $G(V)$ is concave in V in the interval (v_s^*, v_0^*). This implies that there are at most two perfect bayesian equilibria. The four cases are illustrated in Figure 4.

- Suppose G and $-\mathcal{R}$ intersect for $V < v_s^*$ and for $v_s^* < V < v_0^*$. By result 2 of Proposition 3, there cannot be another intersection with $V < v_s^*$. Moreover, because of the concavity of $G(V)$, there cannot be another intersection for V such that $v_s^* < V < v_0^*$. Finally there cannot then be an intersection with $V > v_0^*$ since G is decreasing on that interval while $\mathcal{R}$ is linear and at v_0^* we have $G(v_0^*) < \mathcal{R}(v_0^*)$. This corresponds to the case "Never participants+interior" in Figure 4.

- Similar reasoning implies that if there is an intersection for $V < v_s^*$ and for $V > v_0^*$, then there cannot be an intersection for $v_s^* < V < v_0^*$. This corresponds to the case "Never participants+always participants" in Figure 4.

- Similarly, if there are two intersections for $v_s^* < V < v_0^*$ there cannot be another intersection. This corresponds to the case "2 interior cutoffs" in Figure 4.

- The last case is one where there are two intersections, one with $v_s^* < V < v_0^*$ and one with $V > v_0^*$. This corresponds to the case "Interior+always participants" in Figure 4.

Proposition 5

We examine the case of interior equilibria (where V^* is the swing-participant group). In this case V^* is implicitly defined by:

$$\begin{aligned} &\mu(E[v_i|v_i > v_s^*] - E[v_i|v_i < v_0^*]) + V^* - c + G = 0 \\ \Leftrightarrow \quad &\mu(E[v_i|v_i > v_s^*] - E[v_i|v_i < v_0^*]) + V^* - c + \frac{1}{2}e\left[\left(\frac{F(V^*) - F(v_s^*)}{F(V^*)}\right) + \left(\frac{F(v_0^*) - F(V^*)}{1 - F(V^*)}\right)\right] = 0 \end{aligned}$$

We use the implicit function theorem

$$\frac{\partial V^*}{\partial s} = -\frac{1}{1 + \frac{e}{2}f(V^*)\left[\frac{F(v_s^*)}{(F(V^*))^2} + \frac{F(v_0^*)-1}{(1-F(V^*))^2}\right]}\left(-\frac{e}{2}\frac{f(v_s^*)}{F(V^*)} + \mu\frac{\partial E[v_i|v_i > v_s^*]}{\partial v_s^*}\right)\left(\frac{\partial v_s^*}{\partial s}\right)$$

In stable equilibria, $G(V)$ needs to be increasing in V. This guarantees that $[\frac{F(v_s^*)}{(F(V^*))^2} + \frac{F(v^*)-1}{(1-F(V^*))^2}]$ is positive.

Thus, since $\frac{\partial v_s^*}{\partial s} < 0$ we have that V^* is decreasing in s if and only if:

$$-\frac{e}{2}\frac{f(v_s^*)}{F(V^*)} + \mu\frac{\partial E[v_i|v_i > v_s^*]}{\partial v_s^*} < 0 \tag{11}$$

Given that $E[v_i|v_i > v_s^*] = \int_{v_s^*}^{v_{max}} \frac{vf(v)}{1-F(v_s^*)}dv$, we have

$$\begin{aligned}\frac{\partial E[v_i|v_i > v_s^*]}{\partial v_s^*} &= \frac{-v_s^* f(v_s^*)(1-F(v_s^*)) + f(v_s^*)\int_{v_s^*}^{v_{max}} vf(v)}{(1-F(v_s^*))^2} \\ &= \frac{f(v_s^*)}{1-F(v_s^*)}\left(E[v_i|v_i > v_s^*] - v_s^*\right)\end{aligned}$$

Overall condition (11) can be rewritten as:

$$\frac{e}{\mu} > 2\frac{F(V^*)}{(1-F(v_s^*))}\left(E[v_i|v_i > v_s^*] - v_s^*\right) \tag{12}$$

When $V^* < v_s^*$, the cutoff is defined by:

$$\mu(E[v_i|v_i < v_s^*] - E[v_i|v_i < v_0^*]) + \frac{e}{2}\frac{F(v_0^*) - F(v_s^*)}{1-F(V^*)} = 0$$

Similar computation shows that V^* is decreasing in s if and only if $\frac{e}{\mu} > \bar{b}(s)$ with:

$$\bar{b}(s) = 2\frac{1-F(V^*)}{F(v_s^*)}(E[v_i|v_i > v_s^*] - v_s^*)$$

Defining $b(s) \equiv 2\frac{F(V^*)}{(1-F(v_s^*))}\left(E[v_i|v_i > v_s^*] - v_s^*\right)$, this establishes the result.

Proposition 6

We first consider the case where V^* belongs to the swing participant group.

As specified in the main text, when a supermajority of K votes is required to pass the law, the indifference condition characterizing the cutoff is given by:

$$\begin{aligned}& V^* + e\left[\frac{K}{2N} \times \frac{F(v_0^*) - F(V^*)}{1-F(V^*)} + \frac{2N-K}{2N} \times \frac{F(V^*) - F(v_s^*)}{F(V^*)}\right] \\ = \quad & c - \mu(E[v_i|v_i > v_s^*] - E[v_i|v_i < v_0^*])\end{aligned}$$

Considering K as a continuous variable, we can apply once again the implicit function theorem, we have:

$$\frac{\partial V^*}{\partial K} = -\frac{1}{1+\frac{e}{2N}f(V^*)\left[(2N-K)\frac{F(v_s^*)}{(F(V^*))^2}+K\frac{F(v_0^*)-1}{(1-F(V^*))^2}\right]}\frac{e}{2N}\left[\frac{F(v_0^*)-F(V^*)}{1-F(V^*)}-\frac{F(V^*)-F(v_s^*)}{F(V^*)}\right]$$

Given that in stable equilibria, the denominator will be positive, we have that $\frac{\partial V^*}{\partial K}$ is thus of the same sign as $\frac{F(V^*)-F(v_s^*)}{F(V^*)} - \frac{F(v_0^*)-F(V^*)}{1-F(V^*)}$, which is an increasing function of V^*, negative at v_s^* and positive at v_0^*. There is a unique value $\widehat{V}$ defined by

$$\frac{F(\widehat{V})-F(v_s^*)}{F(\widehat{V})} = \frac{F(v_0^*)-F(\widehat{V})}{1-F(\widehat{V})} \tag{13}$$

such that for interior equilibria, if $V^* \leq \widehat{V}$, V^* is decreasing in K and the opposite for $V^* > \widehat{V}$

Furthermore we have, as explained in the main text, that for $V^* < v_s^*$, V^* is decreasing in K, while equilibria such that $V^* > v_0^*$ are not stable.

We thus have the result of the Proposition that there exists $\widehat{V} \in (v_s^*, v_0^*)$, with $\widehat{V}$ defined by equation (13), such that the voting cutoff V^* is

- decreasing in K if $V^* \leq \widehat{V}$
- increasing in K if $V^* \geq \widehat{V}$

Lemma 2

For f uniform over $(0,1)$, we have that $\Delta(v)$ is constant, so that

$$v_0^* - v_s^* = \hat{v}_s - \hat{v}_0 \tag{14}$$

Using equation (14), for $V < v_s^*$

$$G(V) = \widehat{G}(V) = \frac{1}{2}e\left[\frac{v_0^*-v_s^*}{1-V}\right]$$

For $v_s^* < V < \hat{v}_s$

$$G(V) = \frac{1}{2}e\left[\frac{V-v_s^*}{V}+\frac{v_0^*-V}{1-V}\right] > \widehat{G}(V) = \frac{1}{2}e\left[\frac{v_0^*-v_s^*}{1-V}\right]$$

the inequality derives from the fact that on this interval $V < \frac{1}{2}$.

For $\hat{v}_s < V < v_0^*$, we have:

$$G(V) = \frac{1}{2}e\left[\frac{V - v_s^*}{V} + \frac{v_0^* - V}{1 - V}\right]$$
$$\widehat{G}(V) = \frac{1}{2}e\left[\frac{V - \hat{v}_s}{V} + \frac{\hat{v}_0 - V}{1 - V}\right]$$

So that $G(V) = \widehat{G}(V) \Leftrightarrow V = \frac{1}{2}$
The comparison of $G(V)$ and $\widehat{G}(V)$ is then symmetric for $V > \frac{1}{2}$

Proposition 7

Lemma 2 establishes that for $V < \frac{1}{2}$, $G(V)$ is above $\widehat{G}(V)$. Changing the value of e shifts $G(V)$ and $\widehat{G}(V)$ without affecting $\mathcal{R}$ or $\widehat{\mathcal{R}}$. Thus for e not too high, G an $-\mathcal{R}$ will intersect for lower values of V than $\widehat{G}$ and $-\widehat{\mathcal{R}}$, as visible in Figure 2. If e is sufficiently high, so that $\widehat{G}(0) > \widehat{\mathcal{R}}(0)$, then the cutoff will be zero for the *both secret* case. There is therefore an upper bound e_h on e for the result to hold. Similarly, if e is very low, the unique equilibrium in the *secret vote* environment is for all types to vote against the sanction. There is therefore a lower bound e_l on e for the result to hold.

Lemma 3

As stated in the main text, the incentive to vote in favor of the sanction is, for the never participants, given by:

$$\mu\triangle^* + P\left[Piv\right](-s + G)$$

for swing participants

$$\mu\triangle^* + P\left[Piv\right](v_i - c + G)$$

and for always participants

$$\mu\triangle^* + P\left[Piv\right](G)$$

The incentive is weakly increasing in v_i. Restriction A then implies that all symmetric perfect equilibria are cutoff equilibria as in Proposition 3.

Proposition 8

For all intervals, the indifference condition characterizing the equilibrium is of the form

$$\mu\triangle^* = P\left[Piv\right]\Lambda$$

where Λ can take different values depending on which interval the equilibrium belongs to.

An increase in N, for a given V^* decreases $P\left[Piv\right]$. Thus, if $\triangle^*$ is decreasing in V^*, we see that an increase in N will lead to an increase in V^* (i.e decrease in probability of acceptance). This establishes the first result.

To establish the second result, we first show that the equilibrium where all types vote against the sanction is not a sequential equilibrium. Indeed for all totally mixed strategies, it has to be the case that $E[v|b_i = 1] - E[v|b_i = 0] \geq 0$, because of the cutoff property. So in all sequential equilibria, always participants will always have an incentive to vote in favor of the sanction to benefit from the expected externality and since a vote in favor cannot bring a bad reputation.

We now show that when $N \to +\infty$ the probability that there is a vote against the sanction goes to zero.

Never participants are those who have the least incentives to vote in favor of the sanction. In equilibrium, the net benefit of the never participants to vote in favor of the sanction are given by:

$$\mu\left[E[v|b_i = 1] - E[v|b_i = 0]\right] + P\left[Piv\right](-s + G)$$

We have that $-s + G$ is bounded and $P\left[Piv\right]$ converges to zero, so that in equilibrium, if some individuals vote against the sanction, it has to be the case that $[E[v|b_i = 1] - E[v|b_i = 0]]$ converges to zero. This implies that the proportion of those voting against converges to zero.

Proposition 9

As indicated in the main text, in the public good stage, players will use cutoff strategies conditional on their voting behavior in the first stage. There are therefore two relevant cutoff: $v_0(0)$ and $v_0(1)$.

Conditional on a particular equilibrium, denote P_p the probability that the sanction is adopted independently of individual i's vote, P_r the probability that the sanction is rejected independently of individual i's vote, and P_{piv} the probability that the individual is pivotal. All these probabilities are independent of the player's actual type.

We also use the notation $E_0(b,a) = E\left[v|(b_i = b) \cap (a_i = a) \cap (s = 0)\right]$. For instance $E_0(1,0)$ is the expected value of v given that the player voted for the sanction, the sanction was not passed and he did not participate

Consider group members with $v_i < min(v_0(0), v_0(1))$. If he votes in favor, his expected benefit is:

$$P_p(\mu E_s[1,1] + v_i - c + G_s) + P_r(\mu E_0[1,0] + G_0) + P_{piv}(\mu E_s[1,1] + v_i - c + G_s)$$

where G_s (resp. G_0) denotes as before the expected externality gain when the sanction is passed (resp. not passed).

If the group member votes against, his expected benefit is:

$$P_p(\mu E_s[0,1] + v_i - c + eG_s) + P_r(\mu E_0[0,0] + eG_0) + P_{piv}(\mu E_0[0,0] + eG_0)$$

The net benefit of voting in favor for that individual is thus

$$D(v_i) = P_p\mu(E_s[1,1] - E_s[0,1]) + P_r\mu(E_0[1,0] - E_0[0,0]) + P_{piv}(\mu(E_s[1,1] - E_0[0,0]) + v_i - c + G)$$

$D(v_i)$ is increasing in v_i on that interval.

Similarly, if $v_i > max(v_0(0), v_0(1))$, the net benefit of voting in favor is given by

$$D(v_i) = P_p\mu(E_s[1,1] - E_s[0,1]) + P_r\mu(E_0[1,1] - E_0[0,1]) + P_{piv}(\mu(E_s[1,1] - E_0[0,1]) + G) \quad (15)$$

which is independent of v_i.

We now consider the intermediate regions.

Suppose first $v_0(0) < v_0(1)$ and consider the case $v_0(0) < v_i < v_0(1)$. Such a group member participates when the sanction did not pass, if and only if he voted against the sanction.

If he votes in favor, his expected benefit is:

$$P_p(\mu E_s[1,1] + v_i - c + G_s) + P_r(\mu E_0[1,0] + G_0) + P_{piv}(\mu E_s[1,1] + v_i - c + G_s)$$

If he votes against, his expected benefit is:

$$P_p(\mu E_s[0,1] + v_i - c + G_s) + P_r(\mu E_0[0,1] + v_i - c + G_0) + P_{piv}(\mu E_0[0,1] + v_i - c + +G_0)$$

The net benefit of voting in favor for that individual is thus

$$\begin{aligned} D(v_i) &= P_p\mu(E_s[1,1] - E_s[0,1]) + P_r(\mu(E_0[1,0] - E_0[0,1]) - (v_i - c)) \\ &+ P_{piv}(\mu(E_s[1,1] - E_0[0,1]) + G) \end{aligned}$$

$D(v_i)$ is then decreasing in v_i on that interval.

Suppose on the contrary that $v_0(1) < v_0(0)$ and consider types with $v_0(1) < v_i < v_0(0)$, then the net benefit of voting in favor for that individual is thus

$$\begin{aligned} D(v_i) &= P_p\mu(E_s[1,1] - E_s[0,1]) + P_r\mu(E_0[1,1] - E_0[0,0] + (v_i - c)) \\ &+ P_{piv}(\mu(E_s[1,1] - E_0[0,0]) + (v_i - c) + G) \end{aligned}$$

which is increasing in v_i

Consider case A: $v_0(1) < v_0(0)$.

In this case as shown above, if we impose restriction A as before,the voting strategy is a cutoff strategy with cutoff V^* since the net benefit function $D(v_i)$ is weakly increasing in v_i on all intervals and continuous. There are three situations

- $V^* < v_0(1)$ then there are three zones with respective outcomes (0,0), (1,0) and (1,1)
- $v_0(1) < V^* < v_0(0)$ with two zones: (0,0) and (1,1)
- $V^* > v_0(0)$ with three zones (0,0), (0,1) and (1,1)

We now check whether these equilibria are compatible with the condition $v_0(1) < v_0(0)$. We have

$$v_0(0) = c - \mu(E[0,1] - E[0,0])$$
$$v_0(1) = c - \mu(E[1,1] - E[1,0])$$

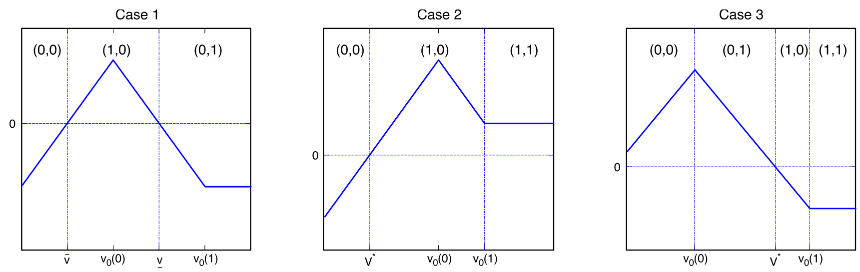

Figure 5: Comparing the cases *both secret* and *secret vote*

So that $v_0(1) < v_0(0)$ is equivalent to:

$$E[1,1] - E[1,0] > E[0,1] - E[0,0] \tag{16}$$

Only that values $E[1,1]$, $E[0,0]$ and $E[1,0]$ (in the second equilibrium) are pinned down in equilibrium, and thus $E[0,1]$ can be chosen low enough to guarantee that condition (16) is satisfied.

Consider case B: $v_0(0) < v_0(1)$. In this case as shown above, the voting strategy is no longer necessarily a cutoff strategy since the net benefit curve $D(v_i)$ first increases in v_i then decreases. There are then potentially three cases for an equilibrium with some types voting in favor and some against.

Consider the case $e = 0$, then there exists a value of μ such that $D(v_i)$ intersects the zero line twice and you thus have outcomes (0,0) for low values of v_i, outcomes (1,0) for intermediate values and (0,1) for high values, as indicated in the result of the Proposition. This is represented in case 1 in Figure 5.

Suppose now that $D(v_i)$ intersects the zero line once for $v < v_0(0)$ (case 2 in Figure 5). Agents will then choose (0,0) for low values, (1,0) for intermediate values and (1,1) for high values.

Finally, if $D(v_i)$ intersects the zero line once for $v_0(0) < v < v_0(1)$ (case 3 in Figure 5), there are four zones (0,0), (0,1), (1,0) and (1,1).

Note furthermore that the condition $v_0(0) < v_0(1)$ is equivalent to

$$E[1,1] - E[1,0] < E[0,1] - E[0,0] \tag{17}$$

We can again find beliefs that will imply condition (17).

Appendix B

B1: complementarities

Consider the contribution stage and fix the contributions a_{-i} of other group members. If individual i contributes, he gets,

$$v_i - c + \mu E[v_i|a=1] + eg(1, a_{-i})$$

Not contributing yields

$$-s + \mu E[v_i|a=0] + eg(0, a_{-i})$$

So contribution behavior is going to be characterized by a cutoff, as in the main model. However, the exact determination of the cutoff is more involved for two reasons. First, a_{-i}, the contribution of others, is an equilibrium choice and thus a function of the equilibrium cutoff. Second, the outcome of the vote in the first phase, gives information about the distribution of types in the population and therefore affects the expectation of a_{-i}. We restrict our attention to cases where only the aggregate result of the vote is revealed.

In spite of these differences, the results of Proposition 1 and 3 are preserved. In particular the shape of the equilibrium is unaffected.

Proposition 10

- *For the case of substitutes, if $1 + \mu\triangle'(v) > 0$ and $e < \widetilde{e}$, the contribution stage is characterized by two cutoffs $\tilde{v}_0 > \tilde{v}_s$ such that group member i contributes if and only if $v_i > \tilde{v}_0$ if there is no sanction and $v_i > \tilde{v}_s$ with a sanction.*
- *For the case of complements, if $1+\mu\triangle'(v) > 0$ and $e < \widetilde{e}$, there exists a bound ϵ such that if $\frac{\partial(g(1,a_{-i})-g(0,a_{-i}))}{\partial a_j} < \epsilon$, then the contribution stage is characterized by two cutoffs $\tilde{v}_0 > \tilde{v}_s$ such that group member i contributes if and only if $v_i > \tilde{v}_0$ if there is no sanction and $v_i > \tilde{v}_s$ with a sanction.*
- *Under Restriction A, all symmetric perfect bayesian equilibria are cutoff equilibria at the voting stage.*

Proof

The aggregate result of the vote (pass or fail), allows the group members to update on the type distribution. We denote $f_s(v)$ the updated belief over the types of other group members (vector v of size 2N) given that the sanction was passed in the first phase and $f_0(v)$ the updated belief given that the sanction was rejected.

Consider the case where the sanction was adopted. Not contributing yields

$$v_i - c + \mu E[v_i|a=1] + eE_s[g(1, a^*_{-i})]$$

Not contributing yields

$$-s + \mu E[v_i|a=0] + eE_s[g(0, a^*_{-i})]$$

where

$$E_s[g(1, a^*_{-i})] = \int g(1, a^*_{-i}(v)) f_s(v) dv$$

is the expected value of the public good, given belief f_s over types and given that the relation in equilibrium between type and action, $a^*_j(v_j)$.

The net incentive to vote in favor is thus given by:

$$v_i - c + s + \mu \left(E[v_i|a=1] - E[v_i|a=0]\right) + e\left(E_s[g(1, a^*_{-i})] - E_s[g(0, a^*_{-i})]\right)$$

So the contribution behavior is going to be characterized by a cutoff. The cutoff is defined by:

$$v^*_s - c + s + \mu\Delta(v^*_s) + e\left(E_s[g(1, a^*_{-i})|v^*_s] - E_s[g(0, a^*_{-i})|v^*_s]\right) = 0$$

where a^*_{-i} depends on the cutoff v^*_s used in equilibrium.

We reexpress this equation as $F(v) = 0$. The implicit function theorem

$$\begin{aligned} \frac{\partial v^*_s}{\partial s} &= -\frac{\frac{\partial F}{\partial s}}{\frac{\partial F}{\partial v^*_s}} \\ &= -\frac{1 + \frac{\partial e\left(E_s[g(1,a^*_{-i})|v^*_s] - E_s[g(0,a^*_{-i})|v^*_s]\right)}{\partial s}}{\frac{\partial F}{\partial v^*_s}} \end{aligned}$$

$E_s[g(1, a^*_{-i})|v^*_s]$ depends on s through the changes in f_s. The numerator is thus negative, provided e is low enough.

The denominator is given by:

$$1 + \mu\Delta'(v_s^*) + e\sum_{j\neq i} \frac{\partial a_j}{\partial v_s^*}\frac{\partial\left(E_s[g(1,a_{-i})] - E_s[g(0,a_{-i})]\right)}{\partial a_j} \tag{18}$$

We have $\frac{\partial a_j}{\partial v_s^*} < 0$. In the case of substitutes $\frac{\partial(E_s[g(1,a_{-i})]-E_s[g(0,a_{-i})])}{\partial a_j} < 0$ so the condition $1 + \mu\triangle'(v) > 0$ is sufficient to guarantee that the overall $\frac{\partial v_s^*}{\partial s} < 0$.

For the case of complements $\frac{\partial(E_s[g(1,a_{-i})]-E_s[g(0,a_{-i})])}{\partial a_j} > 0$, so the extra constraint is required to guarantee that the left hand side is increasing:

$$\frac{\partial\left(g(1,a_{-i}) - g(0,a_{-i})]\right)}{\partial a_j} < \epsilon$$

This establishes the first two results of the Proposition.

So in the second stage the same three categories emerge as in our main model: never-participants, swing-participants and always-participants. The same derivations as those used for Lemma 1 apply, in particular the calculations in equations (2), (3), (4). The only difference involves the calculation of G, but since G is identical for all groups, Lemma 1 and the first result of Proposition 3 follow directly.

B2: changing supermajority rule

We present a particular exemple where the sanction is more likely to be implemented when we increase the supermajority requirement.

We consider a case where the type distribution f is a uniform distribution on (0,1) and we restrict our attention to a group with 3 members. Participation cutoffs simplify to

$$v_0 = c - \frac{\mu}{2}$$
$$v_s = c - \frac{\mu}{2} - s$$

We use the following parameters: $c = 3$, $s = 0.5$, $\mu = 4$ and $e = 1.5$. It implies that $v_0 = 1$ and $v_s = 0.5$. All yes voters are swing participants.

Suppose that majority is required (K=2) Voting cutoff if interior is determined by:

$$\mu(\frac{v_s+1}{2}-\frac{v_0}{2})+V^*-c+\frac{1}{2}e\times[\frac{V-v_s}{V}+1]=0$$

Plugging the parameters, the voting cutoff that satisfies this equation is $V_2 \simeq 0.91$.

Because of the simple majority assumption, the probability that the sanction is accepted is:

$$\begin{aligned} P(\text{sanction is accepted}) &= P(3\text{ votes for})+P(2\text{ votes for}) \\ &= (1-V_2)^3+V_2\times(1-V_2)^2\times 3 \\ &\simeq 0.02 \end{aligned}$$

Now suppose that unanimity is required. If a voter is pivotal, the other two must have voted for the sanction and thus are swing participants. If he votes for, pivotal player will therefore convince all others to contribute. The equation defining the cutoff becomes:

$$\mu(\frac{v_s+1}{2}-\frac{v_0}{2})+V^*+e=0$$

And the voting cutoff is the smallest possible interior cutoff: $V_3 = v_s = 0.5$. This cutoff is much smaller than before because in our example, being pivotal when unanimity is required increases a lot the expected externality gain.

As a result, we have:

$$\begin{aligned} P(\text{sanction is accepted}) &= P(3\text{ votes for}) \\ &= (1-V_3)^3 \\ &= \tfrac{1}{8} \end{aligned}$$

Overall, the sanction is more likely to be implemented when we increase the supermajority in this example.