Organizing for Collective Action: Olson Revisited*

Marco Battaglini
Cornell University, Bocconi University, NBER, CEPR

Thomas R. Palfrey
California Institute of Technology, NBER

February 9, 2023

Abstract

We study a standard collective action problem in which successful achievement of a group interest requires costly participation by some fraction of its members. How should we model the internal organization of these groups when there is asymmetric information about the preferences of their members? How effective should we expect it to be as we increase the group's size n? We model it as an optimal honest and obedient communication mechanism and we show that for large n it can be implemented with a very simple mechanism that we call the Voluntary Based Organization. Two new results emerge from this analysis. Independently of the assumptions on the underlying technology, the limit probability of success in the best honest and obedient mechanism is the same as in an unorganized group, a result that is not generally true if obedience is omitted. An optimal organization, however, provides a key advantage: when the probability of success converges to zero, it does so at a much slower rate than in an unorganized group. Because of this, significant probabilities of success are achievable with simple honest and obedient organizations even in very large groups.

JEL Classification: D71, D72, C78, C92, H41, H54 **Keywords:** Collective Action; Free Riding; Volunteering; Lobbying; Mechanism Design

*The paper has benefited from comments and discussions by participants the 2021 ETH Workshop on Theoretical Political Economy, the 2021 SITE Workshop on Political Economic Theory, and seminar participants at Berkeley, Caltech, Oxford University and the University of Zurich. We thank Francesco Billari, Francesco Billotta and Julien Neves for their research assistance.

1 Introduction

Problems of collective action are among the most basic and ubiquitous forms of strategic interaction in societies. Examples of collective action problems range from the case of private citizens banding together in public demonstrations; to dissatisfied workers participating in union activities; to voters bearing up against bad weather to cast their ballots; to community members donating their time to organize charity or cultural events. At a more macro level, the choices by countries contemplating to join an international environmental agreement also constitute a collective action problem. These are all instances of environments in which a common goal can be achieved by a community, but only if a sufficiently large number of its members are willing to make individual contributions, thus overcoming the incentives to free ride. There are many concrete examples testifying that societies are indeed able to partially solve collective action problems; theories of voluntary behavior and free riding, however, find hard to explain significant levels of individual participation, except assuming that citizens like it or feel morally obliged to it.

In his seminal work, Mancur Olson [1965] provided a taxonomy of the factors determining success of collective action, and highlighted the presence of an organization as a key factor. This observation is intuitive, but it opens up practical and theoretical questions that, as we will argue, have not yet been fully explored in the literature. A first set of questions is positive: what type of organizations should we plausibly expect in collective action problems, and how effective should we expect them to be? A second related set of questions is normative: how do empirically plausible organizations compare to the theoretically optimal organization? To what extent, the presence of an organization (plausible or even optimal) can explain the observed effectiveness of collective action even with a large number of agents? Understanding these questions is important to make sense of the limits and opportunities of collective action and may provide normative insights to improve it.

In this paper we make progress on these issues by studying the effectiveness of organizations in a classic threshold contribution game, widely studied in economics, biology, political science and sociology.[1] In the game, a group of n agents pursue a collective goal that, if achieved, generates a benefit v per agent. The goal is achieved if at least a m_n out of n agents choose to make a personal contribution. The cost of a personal contribution is private information to each agent: it is independently distributed across agents according to some $F(c)$ with support $[0, \overline{c}]$, where typically (but not necessarily) one assumes $\overline{c} > v$. The agents may or may not have an organization, and the organization may be strong (allowing for transfers and/or some form of coercion) or, more plausibly, weak (no transfers and no coercion). We ask how the probability of success changes as n increases, depending on (1) the rate of increase in m_n, (2) whether or not there is an organization, and (3) whether the organization, if it exists, is strong or weak. We also ask under what conditions the group of agents will endogenously form an organization.

[1] Classic contributions are Palfrey and Rosenthal [1984] in economics and Diekmann [1985] in sociology, who coined the term the "volunteer's dilemma" for the special case in which $m_n = 1$. A survey of the work using these games in biology is presented by Archetti and Scheuring [2012]. Applications to environmental economics include, for example, Tavoni et al. [2011] and Barrett et al [2014]. Recent contributions in economics included Harrington [2011], Bergstrom [2017], Battaglini and Benabou [2003], Battaglini [2017], Bergstrom and Leo [2020], Nöldeke and Peña (2020), Dziuda et al. [2021], among others.

Our analysis produces four new theoretical insights. As a preliminary step, we first revisit the equilibrium analysis without an organization when the threshold m_n is a general increasing function of n. Our first finding is that, even without an organization and with a threshold m_n that grows to infinity, failure of collective action is not inevitable: the key factor is the rate of increase of m versus n. Perhaps surprisingly, we show that, regardless of the shape of $F(c)$, success is achieved with probability one if m_n grows at a rate slower than $n^{2/3}$; success is instead impossible if n grows faster than $n^{2/3}$. When m_n grows faster than $n^{2/3}$, moreover, the is a critical group size, n_U, such that the probability of success falls precipitously from a strictly positive success probability to becomes exactly zero for $n \geq n_U$. Collective action, therefore, does not require an organization to be successful if m_n grows sufficiently slowly; but it can be really valuable otherwise.

The other three main findings address the questions of how and to what extent the performance of collective action can be improved by an organization. The key issue here is how to model an organization. The standard approach in mechanism design theory has focused on the study of optimal organizations with transfers that are Bayesian incentive compatible (IC) and interim individually rational (IR), what we refer to as *strong mechanisms*.[2] This approach, through the (IC) constraint, captures the problem of honestly aggregating the dispersed private information regarding the agents' types; it also partially captures, through the IR constraint, a moral hazard problem at the interim stage by guaranteeing a minimal *expected* utility to all types. In most environments of interest, however, this approach bypasses the moral hazard problem faced by the group, since some types might choose to disobey a recommendation by the mechanism if carrying out the recommendation would not be optimal. In the standard Bayesian mechanism design problem, a direct mechanism maps each reported profile of types to an allocation and a payment by each agent, which is then imposed on all agents even if the allocation/payment makes some agents worse off at the reported type profile. In contrast, in a collective action problem, a mechanism lacks the power to simply impose the outcome on all agents, and can only suggest recommended (i.e., not imposed) actions, one for each agent ("go protest", "sign a petition", "volunteer", "do nothing", etc.). The final outcome ultimately depends in the individual willingness of the agents to voluntarily carry out these recommendations.

To clarify this point with an example, consider a community asking for volunteers to organize an event. The event requires at least 3 out of 10 agents to spend one afternoon at the community center and yields a value $v = 1/2$ per person if the quota is met. If $\bar{c} = 1$, then a simple (IC) and (IR) mechanism can achieve the goal with probability 1 using a simple lottery draft mechanism:[3] just randomly select 3 agents and ask them to volunteer. This is (IC), since the information on the types is not used; and it is (IR) since the interim expected cost is lower than the benefit even for the highest type ($v = 1/2 > 3/10$). The problem with this mechanism is that it violates the moral hazard (obedience) constraint: no type $c \in (1/2, 1)$ would agree to volunteer if asked.[4] In such a situation one must add an obedience constraint, requiring that the agents asked to volunteer find it optimal to carry

[2]An exception is Dixit and Olson [2000], as we will discuss below.

[3]This mechanism is also optimal under some weak conditions as we will see in Section 3. This fact is however irrelevant for the present discussion.

[4]Indeed, the mechanism described above is also ex post incentive compatible, since the willingness to participate does not depend on the vector of reports by the agents, before the realization of the randomization

out the mechanism's recommendation. We refer to mechanisms that also satisfy obedience and no transfers as *weak mechanisms.*

This distinction between (IC) and (IR) mechanisms and honest and obedient mechanisms was not especially important for limiting results with large groups in the early literature that assumed constant returns to scale; i.e., the cost of the common project grows linearly with the number of agents. In that case, the limit probability of success is zero even if we ignore the obedience constraint (Rob [1989], Mailath and Postlewaite [1990], Ledyard and Palfrey [1994, 1999]). If one generalizes the constant returns assumption, however, the distinction becomes important. As we are able to show, when m_n grows slower than n, even if at a speed arbitrarily close to n, then optimal (IC) and (IR) mechanisms achieve a probability 1 of success for a large enough n even if we adopt a simple lottery with no transfers as outlined before. Such mechanisms, however, violate the obedience constraint, as shown in the example above. It therefore becomes important to understand what can be achieved with an honest and obedient mechanism.

Our second theoretical contribution is to show that a simple class of honest and obedient mechanisms that we call Volunteer-Based Organizations (VBO) is asymptotically optimal. The mechanism is a simple extension of the lottery described above. In a general VBO, agents are asked to report whether they are willing to be activated (*volunteers*) or not (*free riders*). If the number of agents who state they are willing to be volunteers is lower than some threshold $k_n \geq m_n$,[5] then no agent is asked to be active and the group fails, but wastes no cost of action by any agent. If the number of volunteers is greater than or equal to k_n, then the collective goal is achieved by randomly and anonymously selecting m_n volunteers. These volunteers are willing to follow the recommendation because they know that exactly $m_n - 1$ volunteers will also carry out similar recommendations. Free riders are never asked to be active. Indeed, we show that an even simpler mechanism in which $k_n = m_n$ obtains approximately the same probability of success as the optimal mechanism and a positive probability of success for any n. While these mechanisms appear plausible because of their simplicity, there is also empirical evidence that mechanisms similar to these are likely to emerge endogenously if agents are allowed to communicate (see van de Kragt et al. [1983], Palfrey et al. [2017]).

In our third theoretical result we use the previous characterization to explore the limits of optimal honest and obedient organizations. This allows us to extend the negative limit results of the early literature which hold only for environments with constant return to scale. That literature showed that the limit probability of success in an optimal (IC) and (IR) with constant returns to scale is zero both with an organized and unorganized group (Rob [1989], Mailath and Postlewaite [1990], Ledyard and Palfrey [1994, 1999, 2002]). As mentioned above, this finding does not extend when we generalize the assumption of constant marginal costs. We instead show that with honest and obedient mechanisms the limiting probability of success is the same with an organized and an unorganized group for *any* rate of growth of m_n: as with the Bayesian Nash equilibrium for unorganized groups, the optimal honest and obedient mechanism achieves a limiting success probability of 1 if m_n grows at a rate

in the mechanism is realized. The mechanism however is not obedient for types with $c > v$ since after they are drafted to be active by the random mechanism, they are not willing to be active.

[5] Obviously k_n cannot be less than m_n.

slower than $n^{2/3}$; and it achieves a limit probability of 0 if it grows faster than $n^{2/3}$. An implication of this result is that the probability of success converges to zero even if the total benefit for the group is strictly higher than the total cost in the worst possible scenario in which all types are at the theoretical maximum cost (a case in which, as we will see, strong mechanisms are successful with probability 1 even with no transfers).

So is there any value in having an organization? The fourth lesson from our analysis is that organizations are indeed very useful, and that focusing only on limit results for infinite-sized groups misses an important part of the problem. We show that even when m_n grows faster than $n^{2/3}$, the limit probability of success in a honest and obedient organization converges to zero at a rate that is infinitely slower than without an organization (which indeed achieves exactly zero probability after a finite threshold, $\overline{n}$). In Section 4.2 we also present numerical simulations of the model to quantify how the probability of success with a VBO varies as group size changes. Even when m_n grows faster than $n^{2/3}$, significant (though obviously lower than 1) probabilities of success can be achieved even with n in the tens of thousands (when unorganized groups can achieve success probabilities equal to zero).

Taken together these results confirm and sharpen Olson's intuition for the importance of organizations for collective action, and also highlight important limitations to the power of organizations. Even an ideally optimal honest and obedient organization is not necessarily a perfect solution of the collective action problem for arbitrarily large numbers of interested agents. Simple forms of cooperation such as a VBO, however, can be approximately optimal for large but finite n and provide an effective institution for group success. These results may help explain why numerous cases of successful collective action have been documented, even if collective action is not a panacea for all social problems. This message is confirmed and strengthened when we endogenize the formation of an organization, as we do in Section 6. That analysis suggests that even when successful collective action is only possible with an organization, we should observe the formation of organizations only for values v larger than a threshold $v(n)$, increasing in n. These observations complement and formalize Olson's observation that groups with higher and more concentrated benefits are more successful, which is not necessarily correct if we ignore the endogeneity of an organization.

1.0.1 Related literature

As mentioned above, Olson [1965] was arguably the first to highlight the importance of an organization in solving collective action problems, providing a first informal description of the features of an organization useful to solve them.[6] Formal analysis of this issue, however, had to wait for the development of the theory of optimal mechanisms in Bayesian environments. Our work follows this tradition, departing from it in two ways: first, because we assume no transfers; second, and most importantly, because we require the optimal mechanism to be honest and obedient as discussed above. Most previous theoretical research on optimal mechanisms for public good provision in Bayesian environments consider only strong organi-

[6]Other factors for the success of collective action that have been emphasized by Olson [1965] and the following literature include the cohesiveness of the preferences of the group's members, the elasticity of their cost function as a function of the contributions, and the degree of excludability of the common goal's benefits. Important works on these dimensions see Chamberlin [1974], and more recently Esteban and Ray [2001].

zations, which allow unlimited side payments and interim individual rationality constraints, ignoring the obedience constraint (Mailath and Postlewaite [1990], [1989], Ledyard and Palfrey [1994], [1999], [2002], Hellwig [2003]).[7] As far as we know, the problem of optimal public goods mechanisms in Bayesian environments that satisfy obedience has never been studied. The first three groups of authors have presented negative results of strong organizations assuming constant returns to scale, showing that limit probabilities converge to zero with or without an optimal (IC) and interim (IR) mechanism. Hellwig [2003] has shown that with increasing returns, limit probabilities equal to 1 are feasible with an optimal (IC) and (IR) mechanism with unlimited transfers, indeed always achieved when the demand for the public good is bounded above. When we consider honest and obedient mechanism, results are very different, both with constant returns and without. Allowing for increasing returns, we extend the insight that organizations are not useful in the limit, since we show with (HO) mechanisms they can only obtain the same limit probabilities of success than unorganized groups as $n \to \infty$: with sufficiently increasing returns, however, this probability may be one both with and without an organization, a case that we precisely characterize. With constant returns, we also show that the failure of organizations in the limit is a more severe phenomenon than previously believed, since it extends to cases in which the total societal value of the collective action goal, i.e. vn, is strictly higher than the expected cost in the worst scenario, i.e. when all types are equal to the maximal so $\alpha n \cdot \overline{c}$ (a case in which success is guaranteed with strong organizations and that was ruled out by assumption in previous work).[8]

Following Olson [1965], a significant literature has also studied organizations for collective action from a positive perspective, providing empirical studies of the type of organizations that emerge in concrete examples, both using case studies (Ostrom [1990], for instance) and laboratory experiments (De Kragt et al [1983], Braver and Wilson [1986], Palfrey and Rosenthal [1991], Ostrom and Walker [1991], Ostrom et al. [1992], Palfrey et al. [2017] among others). Several of these experimental papers study public good games very similar to ours, by allowing players to communicate before the contribution stage and ruling out coercion, and report the endogenous emergence of mechanisms similar to the VBO mechanism that we show to be asymptotically optimal.

[7] Other research on Bayesian mechanism design with public goods analyzes "super-strong" organizations that require incentive compatibility, but allow for unlimited side payments and no participation constraints (d'Aspremont and Gerard-Varet [1979], Cremer and McLean [1985], d'Aspremont, Cremer, and Gerard-Varet [1990], Ledyard and Palfrey [1999], [2002]).

[8] The limits of organizations for collective action are also explored by Dixit and Olson [2000], who focus on the the study of incentives to join organized groups. They take a cooperative perspective, assuming that organizations achieve the efficiency frontier through Coaseian bargaining; agents, however, have incentives to stay out, free riding on those who join the organization (for a similar approach in a dynamic setting, see also Battaglini and Harstad [2016]). Passarelli and Tabellini [2017] present a model of political unrest that incorporates psychological rewards for activism. Besides the contributions cited above, moreover, a recent significant literature has studied the limits of organizations in Bayesian mechanisms. See, for example, Healy [2010], Goldlucke and Troger [2020], and Bierbrauer and Hellwig [2016].

2 The Collective Action Model

2.1 Setup

A group with n members, $I = \{1, 2, ..., n\}$, desires an outcome generating a total value of W_n, with each member in the group receiving a personal, direct benefit of $v = W_n/n \in (0, 1)$ which is independent of n.[9] The policy is obtained if and only if at least m_n out of the n members of the group are active. The fraction of agents that are required to be active for success is denoted by $\alpha_n = m_n/n \in (0, 1)$.

Different members have different activity costs, and we denote by c_i the cost of being active for member i. Member i's payoff is given by:

$$\begin{aligned} u_i &= 0 \text{ if } i \text{ is not active and fewer than } m \text{ members are active} \\ &= v \text{ if } i \text{ is not active and at least } m \text{ members are active} \\ &= -c_i \text{ if } i \text{ is active and fewer than } m \text{ members are active} \\ &= v - c_i \text{ if } i \text{ is active and at least } m \text{ members are active} \end{aligned}$$

Costs are i.i.d. and distributed in $[0, \overline{c}]$ according to a distribution $F(c)$ with density $f(c)$. We normalize without loss of generality $\overline{c} = 1 > v$ and we assume $0 < f(c) < \overline{f}$ for some bound $\overline{f} < \infty$ and all $c \geq 0$.[10]

We do not need to assume that m_n is monotonic in n, though typically we expect it to be non decreasing with $m_n \to \infty$ as $n \to \infty$.[11] We refer to the case in which $m_n = \alpha n$ for some fixed constant $\alpha \in (0, 1)$ as the *constant returns to scale* case, since it represents a situation in which the fraction of active members required for success, m_n/n, is constant in n (or equivalently converges to a constant). We refer to the case in which m_n/n declines in n as the *increasing returns to scale* case, since in this case the average cost of the common goal declines in n.

There are two basic forms of organization of the group. The first is no organization at all. In this case each member decides to be active or to free ride independently, given rational expectations about the other members activity decisions. This corresponds to a pure voluntary participation game with a threshold.

The second form is an organized group. We are interested in studying the benefits from organizing even when the organization has very limited tools at its disposal. To this end, we assume that the organization cannot directly observe the types of its members, it cannot exert any form of coercion on the members' actions and it cannot even commit to monetary transfers. We refer to such organizations as *weak organizations*. The organized group can only design an optimal communication mechanism. In such a mechanism, group members can exchange messages through the mechanism; given the exchanged messages, the mechanism sends each agent a recommended action. While the set of such mechanisms can

[9] It is straightforward to extend the analysis to the case in which we have a value v_n depending on on n and $v_n \to v$ as $n \to \infty$.

[10] The analysis directly extends to more general environments. To keep the discussion of the main case simple, we discuss the case with unbounded support and the case with $f(0) = 0$ in Section 10.

[11] We give an example in the next section of a boundary case where $m_n = M$, a constant, for all $n \to \infty$.

be very large, Myerson [1982] has shown that the characterization of the set of all Bayesian Nash equilibria of all such communication mechanisms can be accomplished by considering only honest and obedient direct communication mechanisms.[12] In Section 2.2.1 below we provide a formal characterization of this class of mechanisms and its relationship to the class of incentive compatible and individually rational mechanisms.

In Section 2.2.2 we describe a stronger form of organization in which only incentive compatibility and interim individual rationality is required. This class of mechanism is a useful benchmark since the previous literature has focused on these mechanisms in the form presented here or in close variants (Rob [1989], Mailath and Postlewaite [1990] and Ledyard and Palfrey [1994, 1999, 2002]). We refer to these as *strong organizations*.

2.2 Modelling organizations

2.2.1 Weak organizations

In the absence of monetary transfers, a direct communication mechanism is fully characterized by a mapping from the set of possible type profiles into the set of probability distributions over the subsets of I, $\mu : [0,1]^n \to \Delta\left(2^I\right)$, where we call μ either the *mechanism* or the *activity function*, $\Delta 2^I$ is the set of probability distributions over subsets of I, and we denote by $\mu_g(\mathbf{c})$ the probability the activity function selects subset $g \subseteq I$ of the group to be active at type profile $\mathbf{c}$. Members independently report their types to the mechanism; given the messages $\mathbf{c}$ the mechanism selects a coalition g to activate according to $\mu_g(\mathbf{c})$ and sends the corresponding recommended action to each member; then each member observes their own recommendation and decides whether to comply.

In the following it is sometimes useful to denote a coalition $g \subseteq I$ as an n-dimensional vector of zeros and ones, in which the ith component, g_i, is equal to 1 if $i \in g$ and equal to 0 if $i \notin g$. In this notation $(g_{-i}, 0)$ is a coalition with g_{-i} that excludes i; and $(g_{-i}, 1)$ is the coalition of g_{-i} plus i. We denote $|g| = \sum_i g_i$.

Define $I_i = \{g \subseteq I | i \in g\}$ as the subsets of I containing i and define $I^m = \{g \subseteq I | |g| \geq m\}$ as the set of subsets containing at least m members. Given an activity function, μ, the probability i is active at type profile $\mathbf{c}$ is given by $A_i(\mathbf{c}; \mu) = \sum_{g \in I_i} \mu_g(\mathbf{c})$, and the probability that enough members are active so the group is successful is given by $P(\mathbf{c}; \mu) = \sum_{g \in I^m} \mu_g(\mathbf{c})$. A mechanism is *balanced* if and only, for all $\mathbf{c}$, $\mu_g(\mathbf{c}) > 0 \;\Leftrightarrow |g| = m$. A mechanism has *undercontribution at* $\mathbf{c}$ if $\mu_g(\mathbf{c}) > 0$ for some $|g| < m$ and a mechanism has *overcontribution at* $\mathbf{c}$ if $\mu_g(\mathbf{c}) > 0$ for some $|g| > m$. Thus a mechanism is balanced if and only if it never has overcontribution or undercontribution.

For any mechanism μ define its *reduced form mechanism* by the functions $p_i(c_i) = E_{c_{-i}}[P((c_i, c_{-i}); \mu)]$ and $a_i(c_i) = E_{c_{-i}}[A_i(c_i, c_{-i}); \mu)]$, which are, respectively the expected probability of success and the expected probability i is active, condition on i's cost. We assume the mechanism is symmetric, i.e., for any $i, j \in I$, $c \in [0,1]$, $p_i(c) = p_j(c)$ and $a_i(c) = a_j(c)$.[13] To simplify notation, we drop the member subscripts and simply write these reduced form functions as $p : [0,1] \to [0,1]$ and $a : [0,1] \to [0,1]$. We call a reduced form

[12] This set is closely related to the set of correlated Bayesian equilibrium outcomes of the game.

[13] The restriction to symmetric mechanisms is without loss of generality. To see this, consider any honest

mechanism, (p, a), *feasible* if and only if there exists an activity function μ that generates (p, a). Given any activity function, μ, the interim expected utility for type c who reports to be a type c' is denoted by $U(c', c) = vp(c') - ca(c')$ with $U(c) \equiv U(c, c)$.

In Myerson [1982] a coordination mechanism is honest and obedient (HO) if it provides incentive to reveal the true type and/or to follow the recommendations of the mechanism. Define $\chi(g)$ as the success indicator function when a coalition $g \in I$ is activated: so $\chi(g) = 1$ if $|g| \geq m_n$ and $\chi(g) = 0$ if $|g| < m_n$. Given this, the utility for agent i when the vector of types is c and the activated coalition is g can be written as:

$$u_g^i(c) = \begin{cases} v\chi(g) - c & \text{if } g \in I_i \\ v\chi(g) & \text{if } g \notin I_i \end{cases}$$

Using this notation, condition (HO) requires:

$$U(c) = E_{c_{-i}}\left[\sum_{g \in I} \mu_g\left(c, c_{-i}\right) u_g^i\left(c\right)\right] \geq E_{c_{-i}}\left[\sum_{g \in I} \mu_g\left(c', c_{-i}\right) u_{g_{-i}, \delta_i(g_i)}^i(c)\right] \tag{HO}$$

for any $i = 1, ..., n$, c, $c' \in [0, 1]$, and any function $\delta_i(g_i)$ mapping g_i to $\{0, 1\}$. If we fix $\delta_i\left(g_i\right) = g_i$, (HO) implies the standard interim incentive compatibility condition:

$$U(c) \geq U(c', c) = E_{c_{-i}}\left[\sum_{g \in I} \mu_g\left(c', c_{-i}\right) u_g^i(c)\right] \tag{IC}$$

for any $c, c' \in [0, 1]$.

If we fix $c_i = c$, (HO) implies the following *interim moral hazard condition* (IMH):

$$E_{c_{-i}}\left[\sum_{g \in I} \mu_g\left(c, c_{-i}\right) u_g^i\left(c\right)\right] \geq \max_{\delta_i} E_{c_{-i}}\left[\sum_{g \in I} \mu_g\left(c, c_{-i}\right) u_{g_{-i}, \delta_i(g_i)}^i\left(c\right)\right] \tag{IMH}$$

This inequality states that members find it optimal to follow the mechanism's recommendation on the equilibrium path in which types are truthfully revealed. Condition (HO) however also rules out joint deviations, in which a member misreports his/her type and then disobeys to the recommendation that follows the misreport.

Condition (IMH) has two implications. First, since the right hand side is non negative and the left hand side is $U(c) = E_{c_{-i}}\left[U(c, c_{-i})\right] \geq 0$, it *implies* interim individual rationality (INTIR):

$$U(c) = E_{c_{-i}}\left[U(c, c_{-i})\right] \geq 0 \tag{INTIR}$$

It follows that an (HO) mechanism is also an (IC) and (INTIR) mechanism. Second, (IMH)

and obedient asymmetric mechanism, μ. For any permutation, ρ, of the member indices, define the mechanism μ_ρ by $p_i(c; \mu_\rho) = p_{\rho(i)}(c; \mu)$ and $a_i(c; \mu_\rho) = a_{\rho(i)}(c; \mu)$. Now define the symmetric mechanism, $\overline{\mu}$, by uniformly randomizing among all possible such permutations. Linearity of the member utility function will guarantee that $\overline{\mu}$ is also honest and obedient, and it generates the same total surplus as μ.

implies:

$$c > v \Rightarrow a(c) = 0 \tag{1}$$

since $U((g_{-i}, 0), c) > U((g_{-i}, 1), c)$ for any g_{-i} if $c > v$. Condition (1) is not required in an (IC) and (INTIR) mechanism, so an (IC) and (INTIR) mechanism is not generally an (HO) mechanism (as we will see in Section 3.2).[14]

2.2.2 Strong organizations

The standard approach in the literature to study collective action and public good provision with an organization is to study the best direct mechanism allowing for monetary transfers and requiring incentive compatible as in (IC) and interim individually rational (INTIR), see Rob [1989], Mailath and Postlewaite [1990] and Ledyard and Palfrey [1994, 1999]. As we will see in Section 2.2., monetary transfers are not necessary when n is large, so it will not be important to allow them. We refer to a mechanism without transfers requiring (IC) and (INTIR) as a strong organizations, since in both cases it needs to satisfies weaker constraints than in the weak organization defined in 2.2.1, and thus they can achieve more.

3 Unorganized and strongly organized groups

In this section we study the equilibrium behavior of the activist group in two benchmarks: with no organization in Section 3.1; and with a strong organization in Section 3.2.

3.1 Equilibrium for an unorganized group

For an unorganized group, the payoff function and distribution of costs described above define a Bayesian game where each member simultaneously choose to be active or not. We consider only symmetric equilibria of the game. The symmetry assumption reflects the idea that an asymmetric equilibrium implicitly requires some degree of organization or communication.

For any n denote by p_n^U the probability a member is active in the voluntary contribution game. Given any value of $p_n^U \in [0, 1]$, each member has a best reply that is characterized by a cutpoint, $\widehat{c}_n^U(p_n^U)$, with the property that member i is active if and only if $c_i \leq \widehat{c}_n^U(p_n^U)$. If success requires at least m_n of the members to be active, then an equilibrium cutpoint must satisfy:

$$\widehat{c}_n^U(p_n^U) = vB(m_n - 1, n - 1, p_n^U), \tag{2}$$

where $B(m_n - 1, n - 1, p_n^U) \equiv \binom{n-1}{m_n-1} \left(p_n^U\right)^{m_n-1} \left(1 - p_n^U\right)^{n-m_n}$. In equilibrium, it must be that p_n^U coincides with the probability a member has $c_i \leq \widehat{c}_n^U(p_n^U)$, which is simply equal to $F\left(\widehat{c}_n^U(p_n^U)\right)$. Hence the following condition is necessary and sufficient for p_n^U to be an

[14] An alternative assumption for the participation constraint is *ex post individual rationality* (EXIR). It is interesting to note that (EXIR) does neither imply (IMH) nor, more generally, it implies (HO). An example is presented in Section 3.2, where we show conditions under which the optimal (IC) and (INTIR) mechanism is (EXIR), but fails (IMH) and (HO).

equilibrium probability that a member is active:

$$p_n^U = F\left(vB(m_n - 1, n-1, p_n^U)\right). \tag{3}$$

The equilibrium cutpoint is then given by $c_n^U = \widehat{c}_n^U(p_n^U)$ from (2). An equilibrium exists, trivially, because $p_n^U = 0$ is always a solution to equation (3). It is possible there are also equilibria with $p_n^U \in (0, v)$. In all the analysis that follows, p_n^U always refers to the largest solution to equation 3: this is without loss of generality since we just intend to find an upper bound to the effectiveness of unorganized groups.

An organized group succeeds with positive probability only if there is a strictly positive solution p_n^U. Given a $p_n^U > 0$ and associated $c_n^U > 0$, the equilibrium probability an unorganized group is successful is:

$$P_n^U(p_n^U, \alpha_n) = \sum_{k=m_n}^{n} B(k, n, p_n^U) \tag{4}$$

It is relatively straightforward to see that in the extreme case with constant returns to scale, i.e. $m_n = \alpha n$ for some $\alpha \in (0,1)$, large groups completely fail for sufficiently large n, in the sense that nobody is ever active, including members with arbitrarily small costs: formally, there is a finite $\overline{n}_U(\alpha, v)$ such that $p_n^U = c_n^U = 0$ for $n > \overline{n}_U(\alpha, v)$. To see this point, assume here for simplicity that F is uniform in $[0,1]$ (a more general argument is presented in the proof of Theorem 1 below), consider any $p_n^U \in (0,1)$ and divide both sides of equation (3) by vp_n^U to obtain:

$$\begin{aligned}
\frac{1}{v} &= \binom{n-1}{m_n - 1}\left(p_n^U\right)^{m_n-2}\left(1-p_n^U\right)^{n-m_n} \\
&= \left(\frac{1}{1-p_n^U}\right)\left(\frac{n-m_n+1}{m_n-1}\right)\binom{n-1}{m_n-2}\left(p_n^U\right)^{m_n-2}\left(1-p_n^U\right)^{n-m_n+1} \\
&\Leftrightarrow \\
\frac{1-p_n^U}{v} &= \left(\frac{(1-\alpha)n+1}{m_n-1}\right)\binom{n-1}{m_n-2}\left(p_n^U\right)^{m_n-2}\left(1-p_n^U\right)^{(1-\alpha)n+1}
\end{aligned} \tag{5}$$

The limit of the left hand side of equation (5) converges to $\frac{1-p_n^U}{v} > \frac{1-v}{v}$, and the right hand side converges to $\left(\frac{1-\alpha}{\alpha}\right) B(m_n - 2, n-1, p_n^U)$ which converges to 0. Hence there exists $\overline{n}_U(\alpha, v)$ such that for $n > \overline{n}_U(\alpha, v)$ there does not exist a value of $p_n^U \in (0, v)$ that satisfies equation (3).

While it is natural to assume that m_n increases in n, it is also natural to expect that it grows slower than n. This opens the question of whether and to what extent an unorganized group can achieve success if m_n grows sufficiently slow. The following example shows that at least in the polar extreme case in which m_n is constant, this is not true. This is a particularly extreme example of increasing returns to scale of activism, in which the ratio of required participation to population, m_n/n declines at the speed of $1/n$.

Example 1. The Volunteer's dilemma: Assume $F(c)$ is uniform in $[0,1]$ and consider

the so called "volunteer's dilemma," in which only 1 volunteer is required, regardless of group size, so $m_n = 1$.[15] Will the group be able to send 1 volunteer as $n \to \infty$? It is straightforward to see that the answer is yes. From equation (3), a cutpoint equilibrium solves: $c_n^U = v\left(1 - c_n^U\right)^{n-1}$ which has a unique positive solution c_n^U for all n, with $\lim_{n\to\infty} c_n^U = 0$. The probability of success, from equation 4 is:

$$P_n^U = 1 - \left(1 - c_n^U\right)^n = 1 - \left(\frac{c_n^U}{v}\right)^{\frac{n}{n-1}}$$

so $\lim_{n\to\infty} P_n^U = 1 - \left(\frac{\lim_{n\to\infty} c_n^U}{v}\right) = 1.$ ■

Can the logic of the volunteer's dilemma be generalized to the more realistic case in which m_n grows without bound? We say that m_n grows slower (resp., faster) than a sequence f_n, i.e. $m_n \prec f_n$ (resp., $m_n \succ f_n$) if $m_n/f_n \to 0$ (resp. $m_n/f_n \to \infty$). We say that m_n grows at the same speed as f_n, i.e. $m_n \simeq f_n$ if $m_n/f_n \to \rho$ for some finite ρ.[16] The following theorem shows that, independently of the shape of F, unorganized groups are successful with probability 1 in the limit as $n \to \infty$ if $m_n \prec n^{2/3}$, and are completely unsuccessful if $m_n \succ n^{2/3}$.[17]

Theorem 1. *With an unorganized group:*

(1) if $m_n \prec n^{2/3}$ then $\lim_{n\to\infty} P_n^U = 1$ for all $v \in (0,1)$.

(2) if $m_n \succ n^{2/3}$ then, for all $v \in (0,1)$, there exists $\overline{n}_U(\alpha, v)$ such that the unique equilibrium is $c_n^U = 0$ for all $n > \overline{n}_U(\alpha, v)$, and hence $\lim_{n\to\infty} P_n^U = 0$.

Theorem 1 shows that we do not need constant returns for a group to fail: when the rate of growth of m_n is sufficiently high, i.e. $m_n \succ n^{2/3}$, the probability of success in the unorganized group collapses to *exactly* zero for n large enough. Perhaps more significantly, however, the theorem also shows that the negative results on collective action cannot generalize to all cases in which m_n grows slower than n. Even with no organization the group can achieve a limit success probability of 1; but the no organization case can be seen as a trivial (HO) mechanism, so full success is also possible in the limit in an optimal HO mechanism when $m_n \prec n^{2/3}$.

When m_n grows slower than n, an increase in n has two effects on the right hand side of equation (3). First, it pushes it down, since the probability of exactly $m_n - 1$ active agents goes down: this makes it harder to have a positive intersection. Second, however, it moves the curve to the left as illustrated in Figure 1, since the share of required active agents $\alpha_n = m_n/n$ is also reduced: this makes it easier to have a positive intersection even if $F\left(vB(m_n - 1, n - 1, c_n^U)\right)$ is lower. As we increase n, the probability of success remains bounded above zero and eventually converges to 1 if the sequence of intersections c_ns remains

[15] The volunteer's dilemma is not consistent with our assumption that $m > 1$, which is assumed to hold throughout the rest of the paper, but it is an illustrative boundary case. The argument presented here generalizes to the case in which m_n is constant and equal to any integer $M > 1$.

[16] We will also use the notation $m_n \succeq f_n$ (resp. $m_n \preceq f_n$) to the note the case in which m_n does not grow slower (resp., faster) than a sequence f_n.

[17] All omitted proofs are presented in the appendix.

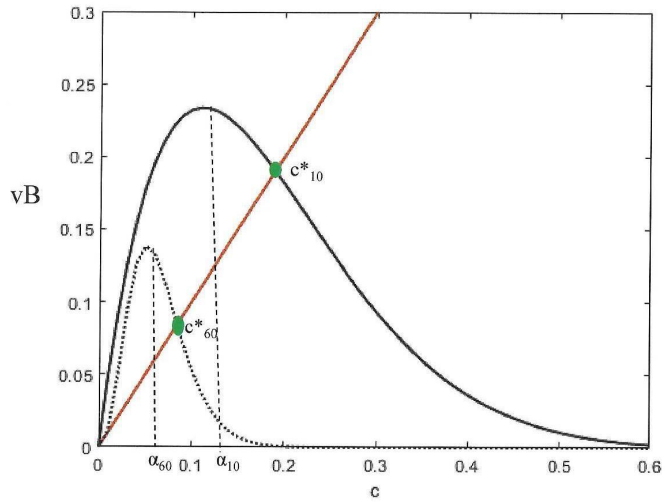


Figure 1: Intuition for equilibria with positive limit probability of success in unorganized groups: $v = .6$, $n = 10, 60$, $m_n = \lceil .5n^{.6} \rceil$.

sufficiently higher than the sequence of thresholds α_n. From Figure 1 we can see that a necessary condition for this is that for all n sufficiently large the right hand side of equation (3), evaluated at α_n, is higher than the 45° degree line, so higher than α_n itself,[18] i.e.:

$$\frac{F\left(vB(m_n - 1, n - 1, \alpha_n)\right)}{\alpha_n} \geq 1. \tag{6}$$

When this is the case, then the highest intersection point, c_n, remains on the right of the threshold α_ns. The proof of Theorem 1 shows that, for any choice of F, a necessary and essentially sufficient condition for this to happen is that m_n declines faster than $n^{2/3}$: when $m_n \prec n^{2/3}$, the right hand side of (6) diverges at infinity, so (6) is satisfied; when $m_n \succ n^{2/3}$, the right hand side of (6) converges to zero, so (6) fails to be satisfied.

The logic behind the "magic number" 2/3 in Theorem 1 can be heuristically explained as follows. Let us first see why, when $m_n \succ n^{2/3}$, the expected share of volunteers in equilibrium $F(c_n^U)$ falls short of the threshold α_n for all n sufficiently large, thus leading to a limit probability of success equal to zero. As $n \to \infty$, $c_n^U \to 0$, so $F(c_n^U) \simeq f(0)c_n^U$. We therefore have $\alpha_n \leq F(c_n^U)$ only if :

$$\frac{m_n}{n} = \alpha_n \leq F(c_n^U) \simeq f(0)c_n^U = vf(0)B(m_n - 1, n - 1, F\left(c_n^U\right)) \tag{7}$$

where the last equality follows from the equilibrium condition for c_n^U. Since $B(m_n - 1, n - 1, c)$ is maximized at $(\alpha_n - 1/n)/(1 - 1/n) \simeq \alpha_n$, in the limit $B(m_n - 1, n - 1, F\left(c_n^U\right) \leq B(m_n - 1, n - 1, \alpha_n)$, so (7) is implied by $vf(0) \cdot B(\alpha_n n - 1, n - 1, \alpha_n)/\alpha_n \geq 1$. The binomial

[18] The function of c_n $F\left(vB(\alpha_n n - 1, n - 1, c_n)\right)$ has a maximum at $c_n = (\alpha_n n - 1)/(n - 1) < \alpha_n$, and it is increasing (resp. decreasing) in c for $c < (\alpha_n n - 1)/(n - 1)$ (resp., $c > (\alpha_n n - 1)/(n - 1)$).

probability of $\alpha_n n - 1$ successes converges to zero at the slowest rate when the probability of success is α_n, and this rate is on the order of $1/\sqrt{\alpha_n n}$. This implies that:

$$B(\alpha_n n - 1, n - 1, F\left(c_n^U\right)) \preceq B(\alpha_n n - 1, n - 1, \alpha_n) \simeq m_n^{-1/2},$$

where $\preceq$ here means that the right hand side converges to zero at the same or slower rate than the left hand side. A necessary condition for (7) therefore is that $\frac{m_n}{n}$ converges to zero faster than $m_n^{-1/2}$, but this condition cannot hold if $m_n \succ n^{2/3}$. When $m_n \prec n^{2/3}$, we have $\alpha_n < F\left(vB(m_n - 1, n - 1, \alpha_n)\right)$ for all n sufficiently large. Since $F\left(vB(m_n - 1, n - 1, 1)\right) = 0 < 1$, continuity implies that there is a solution $c_n^U > \alpha_n$ for all n sufficiently large. Indeed, the proof of Theorem 1 establishes that this solution remains sufficiently larger than α_n, so that the probability of success converges to 1. When instead $m_n \succ n^{2/3}$, then $\alpha_n > F\left(vB(m_n - 1, n - 1, \alpha_n)\right)$ for all n sufficiently large. Again, the proof of Theorem 1 establishes that the only solution of $c_n^U = F\left(vB(m_n - 1, n - 1, F(c_n^U))\right)$ is actually zero in this case for n sufficiently large, so that the probability of success is 0 for n sufficiently large.

3.2 Strong Organizations

The best (IC) and (IR) mechanism can be characterized as the solution of the following maximization problem:

$$\begin{aligned} \max_{p,a} \textstyle\int_0^1 U(c)dF(c) \qquad & (8) \\ \text{s.t. } vp(c) - ca(c) \geq vp(c') - ca(c') \ \forall c, c' \in [0,1] \\ vp(c) - ca(c) \geq 0 \ \forall c \in [0,1] \\ p, a \text{ feasible} \end{aligned}$$

where the first constraint is the (IC) constraint, the second is the (INTIR) constraint and the third is the feasibility constraint discussed in Section 2. Following standard methods, we can show that the (IC) constraint is equivalent to requiring $U'(c) = -a(c)$ and $a(c)$ is non-increasing. Substituting (IC) into the objective function, moreover, we can write:

$$\begin{aligned} \int_0^1 U(c)dF(c) &= -\left[U(c)\left[1 - F(c)\right]\right]\Big|_0^1 + \int_0^1 U'(c)\left[1 - F(c)\right]dc \\ &= vp(0) - \int_0^1 a(c)\left[1 - F(c)\right]dc = vp(0) - E\left[a(c) \cdot \frac{1 - F(c)}{f(c)}\right] \end{aligned}$$

This leads to the problem:

$$\begin{aligned} \max_{p(0), a(\cdot)} \left\{ vp(0) - E\left[a(c) \cdot \frac{1 - F(c)}{f(c)}\right] \right\} \qquad & (9) \\ \text{s.t. } U'(c) = -a(c),\ a(c) \in [0,1] \text{ and non-increasing,} \\ U(c) \geq 0 \, \forall c \in [0,1], \text{ and } p, a \text{ feasible} \end{aligned}$$

To solve (9), consider a relaxed version in which we ignore the (INTIR) constraint. In the online appendix, we prove that when $F(c)$ satisfies the Monotone Hazard Rate Assumption (MHRA) the optimal way to solve this relaxed problem is to keep $a(c)$ flat. Intuitively, when $F(c)$ satisfies MHRA, then in the objective function $a(c)$ is weighted by an increasing function, $-\left[\frac{1-F(c)}{f(c)}\right]$. In this case, if $a(c)$ is strictly decreasing, then it is optimal to shift the probability of participation $a(c)$ from lower to higher values of c; since $a(c)$ is non increasing, the best way to do it satisfying feasibility is to keep $a(c)$ and (by incentive compatibility) $p(c)$ constant: $a_n(c) = \alpha_n$ and $p_n(c) = p_n$. Since we do not have an (INTIR) constraint and it is efficient for the group to be successful, we have $p_n = 1$ and a_n is chosen to be the smallest possible, so $a_n(c) = \alpha_n$. The solution of the relaxed problem is therefore a simple lottery draft mechanism in which m agents are randomly selected with equal probability to be active and the group is always successful. This mechanism, however, is also a solution of the full problem (9) when n is large. To see this, note that for n sufficiently large, $m_n < (v/c) \cdot n$ when $m_n \prec n$, so $U(c) = vp_n - \frac{m_n}{n} \cdot c = v\left(1 - \frac{m_n}{n} \cdot \frac{c}{v}\right) > 0$.

Theorem 2. *If F satisfies (MHRA), then with a strong organization subject to (IC) and (INTIR), we have:*

(1) If $m_n \prec n$, then there is a n^S such that for $n > n^S$ the optimal direct mechanism satisfying (IC) and (INTIR) is a random mechanism in which each g such that $|g| = m_n$ is activated with probability $1/\binom{n}{m_n}$ and each g such that $|g| \neq m_n$ is activated with probability 0. The probability of success converges to one as $n \to \infty$;

(2) if $m_n = \alpha n$ for some $\alpha \leq 1$ and $\alpha < v$, then for all n the optimal direct mechanism satisfying (IC) and (INTIR) is a random mechanism in which each g such that $|g| = \alpha n$ is activated with probability $1/\binom{n}{\alpha n}$ and each g such that $|g| \neq \alpha n$ is activated with probability 0. The probability of success equals 1.

Theorem 2 is relevant for two reasons. First, because it shows that the optimal (IC) and (IR) mechanism is not obedient. This is can be seen from the fact that it requires all types to be active with positive probability, but this directly violates (1) since no type with $c > v$ would find it optimal to be active. The mechanism satisfies (INTIR) since the probability of being drafted is small, so $v\left(1 - \alpha_n \frac{c}{v}\right) > 0$ even if $c > v$; but this only guarantees *interim* participation in the mechanism, not that a type $c > v$ will obey a recommendation to be active. Second, Theorem 2 is relevant because it highlights the need to study more realistic, honest and obedient mechanisms: by ignoring moral hazard, the optimal mechanism achieves complete success as $n \to \infty$. In order to understand why collective action can only be partially successful in more realistic environments, we need to integrate the obedience constraint into the analysis of the optimal mechanism.

We should note that the monotone hazard rate assumption in Theorem 2 is used only for the characterization of the shape of the mechanism, not for the substantive result that there is an (IC) and (INTIR) mechanism that achieves success with probability one for n large when $m_n \prec n$. To see this observe that even without MHRA, the mechanism described in Theorem 2 is (IC) and (INTIR) and achieves success with probability one for $m_n \prec n$ and n sufficiently large.

Note also that Theorem 2 is not in conflict with the main result in Mailath and Postlewaite [1990] where it was shown that the probability of success converges to zero in the best (IC) and (INTIR) mechanisms (even allowing for monetary transfers). That earlier result relied on the assumption that the total benefit of success, nv, is strictly lower than the cost of obtaining it in the worst scenario in which all types have cost equal to 1, an assumption that reduces to $v < m_n/n = \alpha_n$, which is not satisfied for large n in our environment when $m_n \prec n$.[19] The main reason to assume $v < \alpha_n$ is that without it the model generates the perhaps implausible conclusion that a strongly organized group achieves success with probability 1. As we will see in the next two sections, when instead we consider honest and obedient mechanisms, $v > \alpha_n$ does not imply success with probability one. Failure of collective action, therefore does not require this condition. Indeed, there are many situations in which it very natural to assume $v > \alpha_n$, such as situations in which the "sacrifice" of a small share of population is needed to guarantee a successful public protest.

The observation that when we relax this assumption then the collective good can be financed in an (IC) and (INTIR) mechanism with monetary transfers is not completely new, as it was previously made by Hellwig [2003] in a more general environment in which the public good can be chosen as a continuous variable. Theorem 4 differs from Hellwig's result in two ways: it dispenses with the assumption of unlimited monetary transfers and it provides a full characterization of the optimal mechanism, even for large but finite n, as a simple lottery draft mechanism.

4 A simple Voluntary Based Organization (VBO)

As discussed in the previous section, strong mechanisms that only require (IC) and (INTIR) are not a good description of organizations for collective action because they ignore the obedience constraint, implicitly allowing agents to commit to obey to the mechanisms' recommendations. An optimal strong organization, moreover, cannot explain collective action as an empirical phenomenon, since it either predicts complete success of any group when $m_n \prec n^{2/3}$; or complete failure in the limit when $m_n \simeq n$ and $v < \overline{c}$.

What type of (honest and obedient) organization should we then expect? A natural mechanism that has repeatedly emerged in experimental investigation of this question is what we call a *Voluntary Based Organization* (VBO).[20] In this mechanism, which is naturally

[19] Formally, the key assumption in Mailath and Postlewaite [1990] is assumption (iv) in Theorem 2 (p. 357). Translating it to apply to our environment, it is equivalent to requiring that there exists $\epsilon > 0$ such that $nv + n\epsilon < n\alpha_n$.

[20] A significant experimental literature has studied how groups self-organize when given the possibility communication before playing a threshold public good game as described in Section 2. Simmons [1980], De Kragt et al [1983] and Braver and Wilson [1986], among others, studied the case with complete information, a special case of the model described in Section 2. Palfrey and Rosenthal [1991] and Palfrey et al. [2017] studied the case with asymmetric information described in Section 2. These works do not impose a specific communication protocol, they just empirically observe how unstructured pre-play communication affects coordination in this game. In all these experiments, groups that can coordinate through messaging achieve higher success probabilities and welfare. Communication allows the members to self-select as volunteers, and the group to coordinate on a subset of volunteers of minimal size in which the volunteers exceed the requirement, allowing the group to achieve success coordinating on the activation of subgroup of volunteers

honest and obedient, members self-identify as volunteers and coordinate in order to activate a minimal coalition for success when the number of volunteers are more than the success threshold. In this section, we start the analysis of honest and obedient mechanisms by studying this simple type of mechanism. This class has independent interest because of its simplicity; as we will show in the next section, it has the additional virtue of being approximately optimal as $n \to \infty$. The basic intuition presented in this section for why a VBO cannot achieve a higher limit probability than a group with no organization at all, moreover, will prove useful for understanding the more general result that also holds in the context of the optimal honest and obedient mechanism.

4.1 Characterization and properties

In a VBO each member reports his or her type: if the reported type is higher than some threshold c_n^O, the agent is excused and not asked to be active, irrespective of what the other members' reports; if the type is below c_n^O, then the agent is deemed a *volunteer* and is activated with positive probability, determined by the following rule. If the number of volunteers is greater than m_n, then a coalition of exactly m_n volunteers is selected and activated, thus triggering success of the group. If the number of volunteers is fewer than m_n, then no volunteer is activated and the groups is unsuccessful. In case the group activates m_n volunteers, then all volunteers have the same probability of being included. Using the notation introduced in Section 2.2.1 a VBO is defined formally as follows:

Definition 1. *For any $c \in [0,1]$ and any profile of types, $\mathbf{c}$, let $k(\mathbf{c};c) = |\{j \in I | c_j \leq c\}|$. For any given m_n and n, a simple VBO mechanism is defined by a volunteer cutoff $c_n^O \in (0,v)$ such that (1) $A_i(\mathbf{c}) = 0$ for all $\mathbf{c}$ and for all i such that $c_i > c_n^O$; (2) $k(\mathbf{c};c_n^O) < m_n \Rightarrow P(\mathbf{c}) = 0$ and $A_i(\mathbf{c}) = 0$ for all i; (3) $k(\mathbf{c};c_n^O) > m_n \Rightarrow P(\mathbf{c}) = 1$ and $A_i(\mathbf{c}) = \frac{m_n}{k} P(\mathbf{c})$ for all i such that $c_i \leq c_n^O$.*

The following result characterizes the unique incentive compatible VBO. Define the function:

$$Y_n(p) = F\left(\frac{vB(m_n - 1, n-1, p)}{\sum_{k=m_n-1}^{n-1} \frac{m_n}{k+1} B(k, n-1, p)}\right) \tag{10}$$

Proposition 1. *For any m_n and n, the function $Y_n(p)$ has a unique fixed point $p_n^O > 0$ and a VBO is (IC) if and only it has a volunteer threshold $c_n^O = F^{-1}\left(p_n^O\right)$.*

Condition (*33*) provides a simple way to compute the equilibrium threshold c_n^O and characterize its qualitative properties. We will use this condition below for the qualitative analysis and in Section 4.2 where we illustrate equilibrium behavior with numerical simulations. The next result shows that an incentive compatible VBO respects the moral hazard problem of the organization.

Proposition 2. *An incentive compatible VBO is Honest and Obedient as defined in (HO).*

of minimal size.

Propositions 1 and 2 make clear why it is natural to refer to such a mechanism as "volunteer based". When c_n^O is chosen so that the VBO is incentive compatible as in (IC), then the VBO can be implemented with a simple menu with two options. In this implementation, an agent is asked to choose to be a "volunteer" or an "inactive member". The groups is successful if the number of volunteers is greater than or equal to m_n, in which case exactly m_n of them are randomly selected to be active. If the number of volunteers is fewer than m_n, then the group is unsuccessful and no member is activated.

Compared to an unorganized group, a group using a VBO mechanism eliminates two possible sources of ex post efficiency. The first source of inefficiency in the unorganized voluntary contribution mechanism is *undercontribution*, which occurs when fewer than m_n members contribute, as this creates unwanted costs (borne by all members with $c_i \leq c_n^U$) without any benefit to any member of the group. In a VBO, no members at all contribute in such an event. The second source of inefficiency in the unorganized group is *overcontribution*, which occurs when strictly more than m_n members contribute, as the extra members who participate add nothing to benefit the group, but bear the cost. This inefficiency is avoided by always activating the minimum number of volunteers required for success.

These properties have an immediate positive impact on the willingness to volunteer, which ultimately leads to a higher probability of success of the group. The first result is that a VBO guarantees higher participation and higher success probability for any n. Recall that c_n^U and c_n^O are the volunteer thresholds for an unorganized group and a group with a VBO, respectively. Denote by P_n^O the probability that the organized group is successful in the optimal mechanism, and recall that P_n^U is the probability that a unorganized group is successful.

Proposition 3. *For any n, $c_n^O > c_n^U$ and $P_n^O / P_n^U > 1$.*

To evaluate how the group performs when n is large it is important to compare participation and the probabilities of success as $n \to \infty$. The next result shows that the VBO guarantees a much higher participation rate even in the limit as $n \to \infty$.

Proposition 4. *With an unorganized group:*

- *If $m_n = \alpha n$ for some $\alpha \in (0,1)$, then $c_n^O > 0$ for any n and $\lim_{n\to\infty} F\left(c_n^O\right) = F\left(c_\infty^O\right) > 0$.*
- If $m_n \prec n$, then $c_n^O > 0$ for any n and $\lim_{n\to\infty} \left(F\left(c_n^O\right)/\alpha_n\right) \geq 1$

Proposition 4 has important implications for the probability of success of the organized group and the welfare of its members. When $m_n = \alpha n$, the fact that $c_\infty^O > 0$ implies that there are always volunteers with positive probability for any n. When $m_n \prec n$, the fact that $\lim_{n\to\infty}(F\left(c_n^O\right)/\alpha_n) \geq 1$ implies that the rate of participation converges to zero as $n \to \infty$, but at the same rate as the threshold fraction, α_n.

The intuition behind Proposition 4 is as follows. Volunteers are always willing to follow a recommendation to be active, since they know that, conditional on receiving such a recommendation, the mechanism has also activated exactly $m_n - 1$ other volunteers. At the

interim stage, however, an agent might still have an incentive to misreport, since it would prefer some other agent to be called in case of $k \geq m_n + 1$. The cost of misreporting as a free rider is that indeed that there are exactly $m_n - 1$ volunteers in the rest of the group, so the misreport would be pivotal in inducing the group's failure. The expected value of this cost is $B(m_n - 1, n - 1, F\left(c_n^O\right))v$, pretty much the same as in the unorganized group in (3), except that the probability that a member volunteers is $c_n^O > c_n^U$ instead of c_n^U, because there is no fear of being activated when the group is unsuccessful and no concern that when one is activated one's participation is inessential to the group's success. These qualifications are reflected in the denominator of (10). The numerator converges to zero as $n \to \infty$, but c_n^O does not converge to zero because the denominator also converges to 0, and at the same rate as the numerator. The critical difference for the organized group is that volunteers are not called to action indiscriminately, but only if they are needed, and never in excess. The proof of Proposition 4 in the appendix uses the fact that the numerator and the denominator of the ratio on the right hand side of (10) both converge to zero at the same rate and indeed the ratio is strictly positive in the limit. See the appendix for details.

Does the fact that c_n^O remains bounded or converges to zero at the same speed of α_n imply that we can we can get strictly positive probability of success even with large or arbitrarily large groups, and/or we can achieve a higher limit probability of success than without an organization? Consider the case with constant returns to scale first, i.e. $m_n = \alpha n$ for some $\alpha \in (0, 1)$. From the previous literature studying (IC) and (INTIR) optimal mechanisms, we already know that in this case success in the limit as $n \to \infty$ is impossible. Indeed, this literature has shown that a positive limit probability is impossible even if we consider direct mechanisms with fewer constraints and we allow for monetary transfers. The following result confirms these results, but it also shows that limit results do not capture the full story concerning VBO. The fact that individual participation remains strictly positive even in the limit, guarantees that the probability of success converges to zero much slower than without an organization (which indeed is exactly zero for after some finite n).

Proposition 5. *If $m_n = \alpha n$ for some $\alpha \in (0, 1)$, then for all $v < 1$:*

- $\lim_{n\to\infty} c_n^O \in (0, \alpha)$ *and thus* $P_\infty^O = \lim_{n\to\infty} P_n^O = 0$.
- *There exists $\overline{n}_U(\alpha, v)$ such that for all $n > \overline{n}_U(\alpha, v)$, $P_n^U / P_n^O = 0$.*

Proposition 5 follows from the incentive compatibility constraint (33), requiring that an agent with type c_n^O is indifferent between reporting to be a volunteer or not. The constraint implies that the expected cost of being activated $c_n^O a(c_n^O)$ must be equal to $v\left[p_1(c_n^O) - p_2(c_n^O)\right]$, the net increase in the probability of being pivotal from reporting to be a volunteer, which converges to zero as $n \to \infty$. Since $c_n^O \to c_\infty^O > 0$, it must be that the probability of being activated converges to zero: $a(c_n^O) \to 0$. In this case, the expected share of agents willing to be active is lower than the threshold: so, by the law of large numbers, the probability of passing the threshold converges to zero as $n \to \infty$. Hence $P_\infty^O = 0$ so it must be that $F\left(c_n^O\right) \leq \alpha$.

When $m_n \prec n$, then *both* the required threshold and participation converge to zero and $F\left(c_n^O\right)/\alpha_n \to 1$. The key question is exactly how $F\left(c_n^O\right)/\alpha_n$ converges to 1. If

$F\left(c_n^O\right)/\alpha_n$ converges from below and convergence is slow, then the probability that the number of volunteers passes the threshold converges to zero; if instead convergence is fast or $F\left(c_n^O\right)/\alpha_n$ converges from above, then the probability of success of the group will be strictly positive even for an arbitrary large number of activists. We might expect that the presence of an organization that allows for coordination and eliminates wasteful participation makes it possible to achieve higher limit probabilities of success in the limit, at least for some parameterizations. This conjecture is however incorrect. We have:

Proposition 6. *For any $v \in (0,1)$:*

- *If $m_n \prec n^{2/3}$, then $\lim_{n\to\infty} P_n^O = 1$ for organized groups using the VBO mechanism.*
- *If $m_n \succ n^{2/3}$, then $P_n^O > 0$ for all n* and *$\lim_{n\to\infty} P_n^O = 0$ for organized groups using the VBO mechanism. Hence, $P_n^U/P_n^O = 0$ for sufficiently large n.*

Proposition 6 establishes two results that, as we will show in the next section, will hold more generally for HO mechanisms. First, perhaps surprisingly, the limit probability of success is the same with a VBO or without it, in an unorganized group. When $m_n \prec n^{2/3}$, a limit probability of success is 1, but this was also true for an unorganized group; when $m_n \succ n^{2/3}$, the limit probability of success is zero with a VBO, once again, just as in an unorganized group. The second result is that with a VBO the probability of success is positive for any n, a feature that is not shared by the equilibrium in an unorganized group (where the probability is exactly zero after a finite n): this can be a significant benefit of adopting a simple VBO compared to having an unorganized group. As we will show in the next section where we quantify by numerical methods the success probability in a VBO, for reasonable parameter values groups with VBO can achieve high probabilities of success even for large groups, even if $m_n \succ n^{2/3}$.

It is useful to go over the intuition of this result since, suitably generalized, it will also help with understanding the extension to general HO mechanisms in the next section. In equilibrium, the condition that induces the marginal type to volunteer is:

$$a_n^O(c_n^O) \cdot c_n^O = vB(m_n - 1, n-1, F\left(c_n^O\right)) \tag{11}$$

where $a_n^O(c) = \sum_{k=m_n-1}^{n-1} \frac{m_n}{k+1} B(k, n-1, c)$ is the probability that a volunteer is activated when the other members use a threshold c, so the left hand side is the expected cost of being a volunteer;[21] and the right hand side is the expected benefit of being a volunteer, i.e. v times the probability that a single additional volunteer is useful. We can rewrite (11) as:

$$c_n^O = \frac{vB(m_n - 1, n-1, F\left(c_n^O\right))}{a_n^O(c_n^O)} = Y(c_n^O) \tag{12}$$

[21] Note that in Section 2, we used the notation $a(c)$ to the note the interim probability of success for a type c. In a VBO, the interim probability of success for a volunteer, i.e. a type $c \le c_n^O$, is independent of the exact volunteer's type; however it depends on the threshold used by the other players, c_n^O. The notation $a_n^O(c)$ makes this dependence explicit, representing the probability of activation as a function of a general threshold c used by the other players.

where $Y(\cdot)$ is defined by (10). The share of volunteers does not fall short of the threshold for success only if $\alpha_n \leq F\left(c_n^O\right)$. Since $c_n^O \to 0$, we have $F\left(c_n^O\right) \simeq f(0)c_n^O$, implying:

$$\frac{\alpha_n}{f(0)} \lesssim c_n^O = Y(c_n^O) \leq Y(\alpha_n) = \frac{vB(m_n - 1, n-1, \alpha_n)}{a_n^O(\alpha_n)}$$

where $\lesssim$ here means that the LHS converges to a value less than or equal to the RHS. The second inequality, $Y(c_n^O) \leq Y(\alpha_n)$ follows from the fact that, $Y(\cdot)$ is a decreasing function, by Proposition 1. Hence:

$$a_n^O(\alpha_n) \lesssim vf(0) \cdot \frac{B(m_n - 1, n-1, \alpha_n)}{\alpha_n} \tag{13}$$

If $m_n \succ n^{2/3}$, we know from the discussion of Theorem 1 that the right hand side of (13) converges to zero. Condition (13), therefore, implies that $a_n^O(\alpha_n)$ must also converge to zero. It is however intuitive to see that this is impossible. Note that $a_n^O(\alpha_n)$ is the probability that a volunteer is activated when the threshold used by the other members for volunteering is α_n. But when this is the case, for large n there will be a share of volunteers roughly equal to α_n. In this case, the probability that a volunteer is activated cannot be arbitrarily small since, even conditioning on having at least a share α_n of volunteers, the share of volunteers will almost surely be only marginally greater than α_n, the minimal requirement for success.

4.2 Numerical Computations

The results of the previous section - that very large organized groups are infinitely more successful than unorganized groups - are limiting results. It is also insightful to compare the performance of organized and unorganized groups of finite size.[22] As the examples below indicate, the performance of organized groups is many orders of magnitude greater than the performance of unorganized groups, even with a small number of members.

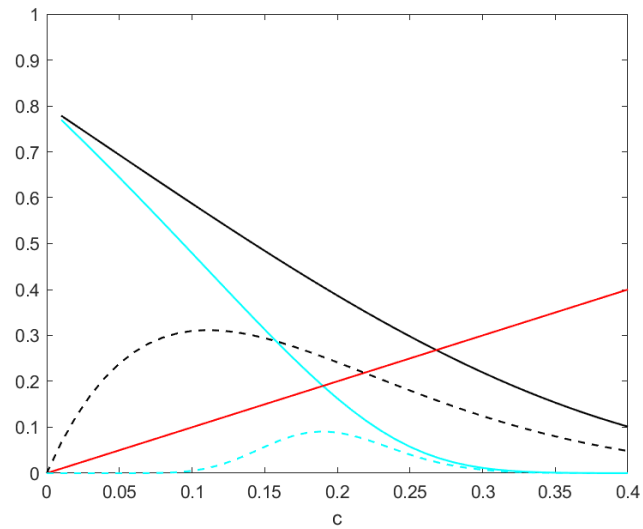
Example 1: $\alpha_n = 0.2$**,** $v = 0.8$**,** $F \sim \mathcal{U}_{[0,1]}$. In this example, we compare the organized and unorganized group, varying the size of the group, but keeping α fixed (constant returns to scale). From Proposition 1, the organized group's optimal threshold satisfies: $c_n^O = Y_n(c_n^O, \alpha, v)$ (where we make explicit the dependence of Y_n on α, v for convenience), where:

$$Y_n(c_n^O, \alpha, v) = v\frac{B(m_n - 1, n-1, c_n^O)}{\sum_{k=m_n-1}^{n-1} \frac{m_n}{k+1} B(k, n-1, c_n^O)}. \tag{14}$$

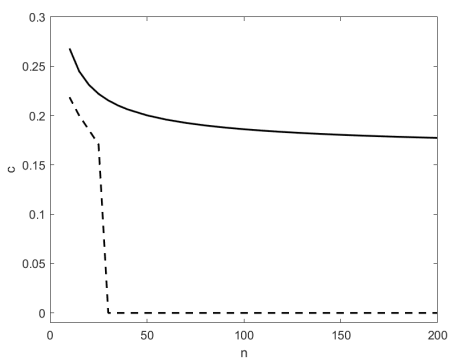
With no organization the equilibrium condition is: $c_n^U = Z_n(c_n^U, \alpha, v)$, where:

$$Z_n(c_n^U, \alpha, v) = vB(m_n - 1, n-1, c_n^U) \tag{15}$$

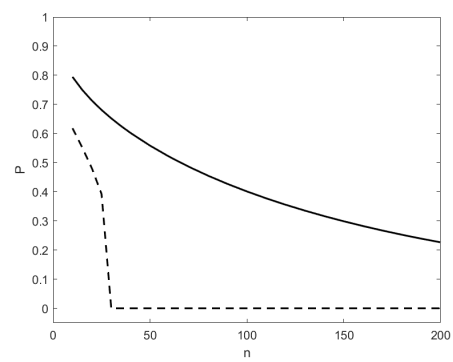
[22] A theoretical lower bound on the speed of convergence of the probability of success in a VBO is presented in Section 5.3.



(a) Equation 2 (Unorganized; dashed) and Equation 12 (VBO; solid) $n = 10$ (black) and $n = 80$ (cyan)



(b) c_n^U (dashed) and c_n^O (solid).



(c) P_n^U (dashed) and P_n^O (solid).

Figure 2: Comparison of VBO mechanism (solid lines) and unorganized group equilibrium (dashed lines). $\alpha_n = 0.2$. $v = 0.8$. F(c) Uniform.

The equilibrium probabilities of success for the unorganized and organized groups are, respectively:

$$P_n^U(c_n^U, \alpha) = \sum_{k=m_n}^{n} B(k, n, c_n^U) \text{ and } P_n^O(c_n^O, \alpha) = \sum_{k=m_n}^{n} B(k, n, c_n^O)$$

The top panel of Figure 2 illustrates the $c_n^O = Y_n(c_n^O, \alpha, v)$ and $c_n^U = Z_n(c_n^U, \alpha, v)$ conditions for group with 10 and 80 members, fixing $\alpha = 0.2$, $v = 0.8$. The monotonically downward sloping solid darker curve is $Y_{10}(c, 0.2, 0.8)$ and the downward sloping lighter shaded curve is $Y_{80}(c, 0.2, 0.8)$, corresponding to the RHS of equation 12. The respective equilibrium cutpoints, c_{10}^O and c_{80}^O, are given by the intersection of these two $Y(\cdot)$ curves with the 45° line (in red). The darker dashed hump-shaped curve in top panel of Figure 2 is $Z_{10}(c, 0.2, 0.8)$,

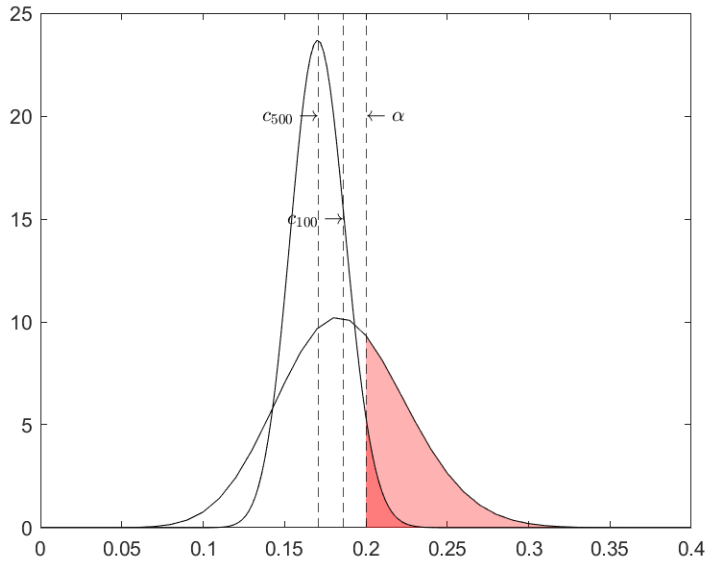
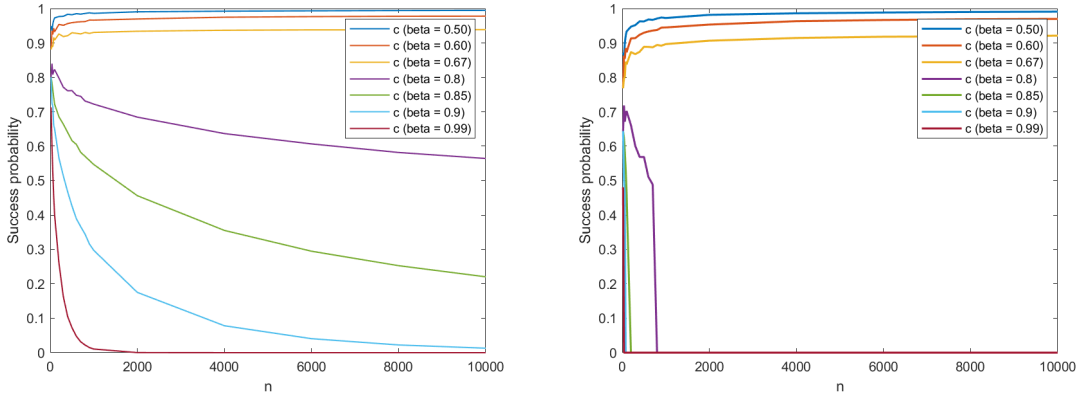


Figure 3: The distribution of the fraction of group members who are volunteers (i.e. $c \leq c_n^*$) for $n = 100$ and $n = 500$.

and c_{10}^U is given by the highest intersection of this curve with the 45° line. The lighter dashed hump-shaped curve in top panel of Figure 2 is $Z_{80}(c, 0.2, 0.8)$, which does not intersect the 45° line at any positive value, so the unique equilibrium for the unorganized group is $c_{80}^U = 0$ and the unorganized group has zero participation.[23] The bottom right panel of Figure 2 graphs the success probabilities of the organized group (solid) and the unorganized group (dashed) as a function of group size for n ranging from 10 to 300. The bottom left panel shows how the thresholds c_n^O and c_n^U change as a function of n. In this case, $c_n^U = 0$ for $n \geq 30$.

As proved in Proposition 5, as n increases $c_n^O \to c_\infty^O > 0$, and c_n^O does not change very much with n for n high, a property illustrated the bottom left panel of Figure 2. Still, the probability of success converges to zero as $n \to \infty$. The intuition for this is illustrated in Figure 3. The distribution of the fraction of volunteers, i.e. those with cost below c_n^O, is a binomial with expected value c_n^O. As n becomes large, the expected value becomes insensitive to n, but the distribution becomes more concentrated around c_n^U. This implies that the tail probability that the share of group members who are available is larger than α converges to zero. Figure 3 illustrates the distribution of the share of group members who are available for $n = 100$ and $n = 500$, again fixing $\alpha = 0.2$, $v = 0.8$. These tail probabilities are shaded in red. With $n = 100$, the probability of success is $P_{100}^O(0.19, 0.2)) = 0.31$ and with $n = 500$, the probability of success is $P_{500}^O(0.17, 0.2) = 0.04$.

[23] If $\alpha n = m > 1$, there is always a solution at $c^{U*} = 0$. When n is sufficiently small, there are can also be at most two positive, equilibrium cutpoints for the unorganized group, in addition to the 0 equilibrium cutpoint.



(a) Organized Groups (b) Unorganized Groups

Figure 4: Probability of Success as a function n for different values of β. Comparison of organized and unorganized groups. Higher curves correspond to lower values of β. $v = 0.8$. $F(c)$ Uniform.

Example 2: $m_n \simeq n^\beta$**.** Figure 4 illustrates the equilibrium probability of success for organized (left panel) and unorganized (right panel) groups of size up to 10^4 for $m_n \propto n^\beta$ for and various levels of β. As in the previous example, $v = 0.8$ and F is Uniform. One can see that for $\beta < .66$, both unorganized and organized group can achieve success with probability that is high even for small groups and rapidly converges to 1. For $\beta = .67$, the probability of success also starts out relatively high and is essentially flat for $n < 10,000$ for both organized and unorganized groups, although it will theoretically converge to zero eventually in both cases. For $\beta > .67$, instead the probability converges to zero, and groups without an organization fail completely after reaching a threshold size that is decreasing in β. In contrast, organized groups obtain high probability of success even for $\beta = 0.8$ and this success falls relatively slowly in group size. The figure shows the importance of considering finite n even when the limit probability of success is 0. For organized groups and $\beta = 0.85$, the probability of success converges to zero as $n \to \infty$, but the probability of success is more than 20% even for groups with more than $10,000$ members. In contrast, for unorganized groups and $\beta = 0.85$, the probability of success is *exactly zero* for all $n \geq 200$.

5 When is a VBO optimal?

The previous section showed that a group can achieve significant (albeit imperfect) levels of success even when n is large by organizing with a very simple honest and obedient mechanism. Proposition 6 suggests that the advantage of a VBO is not in the limit probability of success that can be achieved with infinitely large groups, which is equal to the limit probability obtainable without an organization; instead, the real benefit of a VBO is that when the probability converges to zero in the limit, it does so at a much slower rate compared with unorganized groups. Is this a general result? Indeed, Proposition 6 does not rule out the possibility that there is an even better (HO) mechanism that achieves higher probability of

success than in an unorganized group. In this section we prove two key results that answer affirmatively to this question, but qualify its implications. First, we show that a minimal generalization of the simple VBO is asymptotically optimal: so for large n, there is little benefit to consider more complicated mechanisms. Second, we show that in the limit, as $n \to \infty$, the probability of success in an optimal (HO) mechanism is the same as without an organization: zero if $m_n \succ n^{2/3}$, and one if $m_n \prec n^{2/3}$. We prove this by showing that the result holds for a generalized class of VBO mechanisms, and then using the fact that a generalized VBO is asymptotically optimal.

5.1 The optimal honest and obedient mechanism

The optimal honest and obedient mechanism is defined as the solution $a^*(c)$, $p^*(c)$ of the following problem:[24]

$$\max_{a(c),\, p(c)} \int_0^1 U(c)dc \tag{16}$$
$$\text{s.t.(HO) and } a(c),\, p(c) \text{ is feasible.}$$

where (HO) is the honest and obedient constraint described in Section 2.2. As discussed in Section 2, the (HO) constraint implies (IC), (INTIR) and Condition (1). In the following we will first study the solution of the relaxed problem in which only (IC) and (1) are considered, we will then prove that this solution satisfies the omitted constraints and thus solves (16). By standard methods and formally proven in the appendix, we can write the relaxed problem as:

$$\max_{p(0)\in[0,1], a(c)\in[0,1]} \left\{ vp(0) - E\left[a(c)\cdot\frac{1-F(c)}{f(c)}\right]\right\} \tag{17}$$
$$\text{s.t. } U'(c) = a(c) \text{ with } a(c) \text{ non-increasing}$$
$$a(c) = 0 \text{ for } c > c^*, \text{ where } c^* = \min\{c \le v \,|vp(c) - ca(c) \le vp_2\}$$
$$\text{and } p, a \text{ feasible}$$

In (17) we derived the objective function using the (IC) constraint in a similar way as in (9). The constraints in the second line of (17) is a monotonicity constraint implied by IC, also present in (9); the constraint in the third line follows from (1) and incentive compatibility, and it is new to (17).[25] Note that in the problem above we have no (IR) constraint; IR, however, follows from the monotonicity of $U(c)$ and the definition of c_n^*.[26]

To study the solution to this constrainted optimization problem it is useful introduce a class of mechanisms that generalizes the simple VBO of the previous section. We call these mechanisms general Voluntary Based Organizations, or general VBO. A general VBO is defined by a threshold $k_n^G \ge m_n$ and a volunteer cutoff c_n^G. A k_n^G-generalized VBO with a threshold k_n^G greater than or equal to m_n works as follows. If there are more than

[24] Whenever it does not create confusion, as here and when we take n as given, we omit the subscript n in the equilibrium variables $a_n^*(c)$, $p_n^*(c)$, c_n^*.

[25] See the proof of Theorem 3 for the details on how it follows from (1) and incentive compatibility.

[26] Indeed, $U(c) \ge U(c^*) \ge vp_2 \ge 0$ for all $c \in [0,1]$.

k_n^G volunteers, then the mechanism selects and activates exactly m_n volunteers, each with equal probability, thus guaranteeing that the group succeeds. If there are fewer than k_n^G volunteers, then the mechanism selects 0 volunteers and the group fails. If there are exactly k_n^G volunteers, then the mechanism selects and activates exactly m_n volunteers with some probability $q_{k_n^G} \in (0, 1]$ and selects exactly 0 volunteers with probability $1 - q_{k_n^G}$. Thus, the simple VBO analyzed above corresponds to the special case $k_n^G = m_n$ and $q_{k_n^G} = 1$. The lottery draft mechanism that in Section 3.2 we showed solves problem (8) for the best (IC) and (IR) mechanism when n is large can also be seen as a VBO with $k_n^G = m_n$, $q_{k_n^G} = 1$ and $c_n^G = 1$.

Definition 2. *For any $c \in [0, 1]$ and any profile of types, $\mathbf{c}$, let $k(\mathbf{c};c) = |\{j \in I | c_j \leq c\}|$. For any given m_n and n, a generalized VBO mechanism is defined by a volunteer cutoff $c_n^G \in (0, v)$ and a critical mass threshold $k_n^G \geq m_n$, such that (1) $A_i(\mathbf{c}) = 0$ for all $\mathbf{c}$ and for all i such that $c_i > c_n^G$; (2) $k(\mathbf{c};c_n^G) < k_n^G \Rightarrow P(\mathbf{c}) = 0$ and $A_i(\mathbf{c}) = 0$ for all i; (3) $k(\mathbf{c};c_n^G) > k_n^G \Rightarrow P(\mathbf{c}) = 1$ and $A_i(\mathbf{c}) = \frac{m_n}{k} P(\mathbf{c})$ for all i such that $c_i \leq c_n^G$.*

Intuitively, one might conjecture that a *simple* Voluntary Based Mechanism (VBO) would be the optimal honest and obedient binary mechanism; i.e., the group succeeds if and only if there are enough volunteers (at least m_n). Using a higher threshold than m_n seems wasteful and is ex post suboptimal since it implies that there are events in which the group fails even thought the number of volunteers is known (by the mechanism) to exceed the minimum number required for success. However, in principle, it could be ex ante optimal for the mechanism to commit to failure in some such events (e.g., $k_n^G = m_n + 1$) in order to create better incentives for the agents to self identify as volunteers and more generally relax the (HO) constraint. This could happen if increasing k_n^G above m_n leads to a higher volunteer cutoff, c_n^G.

In Section 3.2 we showed that the optimal (IC) and (INTIR) mechanism for any $m_n \prec n$ can be easily characterized when n is large after we note that we can improve the objective function in (9) without violating the constraints by "flattening" the mechanism, i.e. by making the mechanism less sensitive to an agent's type c. The optimal mechanism for n large is indeed a simple lottery in which all types are asked to be activated with positive probability and randomly selected with the same probability. Because the mechanism ignores the type reports, (IC) is always trivially satisfied; moreover, (INTIR) is also satisfied for n large since given that all types are activated, the probability of being activated is α_n, smaller than v for n large.

The same logic cannot be applied to (17). Here too, the objective function (which is the same in (9) and (17)) improves if we "flatten" the mechanism; however, now flattening the mechanism may affect the obedience constraint. To see this note that a mechanism in which all types are activated with positive probability is certainly impossible, since no type with $c > v$ will ever accept to be activated even if it is interim rational (INTIR) to commit to participate in the mechanism. In the optimal (HO) mechanism we have a maximal type $c_n^* < v < 1$, who is indifferent between volunteering, in which case he will be activated with some probability (the third line in (17)), and free riding. By flattening the mechanism, we now necessarily require higher expected participation from this type c_n^* which would break that indifference. Having a flatter mechanism therefore involves a trade off: on the one hand,

for a given c_n^*, it improves the objective function since it relaxes the (IC) constraint; on the other hand, however, it may imply lower participation, in the form of a lower c_n^*. The shape of the mechanism in this case may be difficult to pin down since it depends on the trade-off between the benefit of keeping the volunteer cutoff high (a higher c_n^*), which produces a larger pool of volunteers, and keeping the mechanism flat.

In the reminder of this section we provide two key results. In Section 5.2, we prove that a general VBO is approximately optimal as $n \to \infty$. This implies that it is impossible that an optimal mechanism achieves positive probability in the limit if the optimal VBO mechanism does not achieves the same utility in the limit.

In Section 5.3, we use the characterization of Section 5.2 to prove that in the limit as $n \to \infty$ unorganized groups are just as likely to be successful as groups with the best honest and obedient mechanism. This result thus allows us generalize the equivalence result of Mailath and Postlewaite [1990]. The previous authors showed that for $m_n = \alpha n$ for some $\alpha \in (0, 1)$, both an unorganized group and a group with the optimal (IC) and (INTIR) mechanisms are equally unsuccessful in the limit. When m_n grows slower than n, the optimal (IC) and (IR) always achieves probability of success equal to 1, so the equivalence result does not generalize for optimal (IC) and (INTIR) mechanisms. An implication of this, is that the simple VBO is also an asymptotically optimal mechanism that guarantees positive probability of success for any finite n.

5.2 A positive result: the asymptotic optimality of a general VBO

In this section we prove that a generalized VBO is asymptotically optimal as $n \to \infty$. The result is immediate if the limit probability of success of a mechanism that solves (17) converges to zero (i.e., if $m_n \succ n^{2/3}$). But we know from the previous analysis that this limit probability is strictly positive if $m_n \prec n^{2/3}$, which is the more difficult case to prove. In this section, we show that this limit probability is asymptotically the same as the probability (and the associated welfare level) that can be achieved with a generalized VBO. This will allow us to prove that the limit probability of success in an optimal VBO is the same as with no organization.

We proceed in two steps. We first prove an intermediate result of independent interest. We define a mechanism as *binary*, if it allows the agents to send at most two messages: volunteer or not volunteer. In the next result, we establish that the optimal binary mechanism is a generalized VBO.

Lemma 1. *For any $v \in (0, 1), m_n$ and n, there exists $c_n^b \in (0, 1)$ and $k_n^b \geq m_n$ such that a k_n^b-generalized VBO mechanism with volunteer cutoff c_n^b and critical mass threshold k_n^b is an optimal honest and obedient binary mechanism.*

Proof: The proof is carried out in two steps. In step 1, we establish that the optimal binary mechanism is non-wasteful, meaning that it never activates more volunteers than necessary. In step 2, we show that if the optimal binary mechanism activates m_n agents with positive probability with k volunteers, then it must activate m_n agents with probability one with more than k volunteers. This implies that the optimal non-wasteful binary mechanism

is characterized by a threshold k_n^b as specified in Definition 1. Proof details are in the appendix. ■

In an optimal binary mechanism, the mechanism does not use detailed information regarding the type of the agent, just whether the agent is willing to be volunteer. The optimal binary mechanism is therefore asymptotically optimal if the best (HO) mechanism also does not use detailed information regarding the types c, even if the probability of success is positive in the limit. The following result establishes this property:

Lemma 2. *Let $a_n^*(c), p_n^*(c)$ be an optimal honest and obedient mechanism. For any two types c' and c'' with $c' > c'' > 0$, $p_n^*(c') - p_n^*(c'') \to 0$ and $a_n^*(c') - a_n^*(c'') \to 0$ as $n \to \infty$.*

To see the intuition of this result, note that by IC, $p_n^*(c)$ is non-increasing in c, so $p_n^*(c') \leq p_n^*(c \leq c')$ and $p_n^*(c'') \geq p_n^*(c \geq c'')$, where $p_n^*(c \leq c')$ and $p_n^*(c \geq c'')$ are the interim probabilities of success conditioning on, respectively, $c \leq c'$ and $c \geq c''$. Moreover,

$$p_n^*(c \leq c') = \tau_{0,n-1} P_B^n + (1 - \tau_{0,n-1}) P_0^n$$

where $\tau_{0,n-1}$ is the probability that, out of the remaining $n-1$ agents, there is at least one type $\widetilde{c} \geq c''$, P_B^n is the expected probability of success conditioning on the presence of a type $\widetilde{c} \geq c''$ and a type $\widetilde{c} \leq c'$, and P_0^n is the expected probability of success conditioning on the presence of at least one type $\widetilde{c} \geq c''$. Similarly, we have:

$$p_n^*(c \geq c'') = \tau_{1,n-1} P_B^n + (1 - \tau_{1,n-1}) P_1^n$$

where $\tau_{1,n-1}$ is the probability that, out of the remaining $n-1$ agents, there is at least one type $\widetilde{c} \leq c'$, and P_1^n is the expected probability of success conditioning on the presence of at least one type $\widetilde{c} \leq c'$. But then we have:

$$0 \leq p_n^*(c') - p_n^*(c'') \leq (\tau_{0,n-1} - \tau_{1,n-1}) P_B^n + (1 - \tau_{0,n-1}) P_0^n - (1 - \tau_{1,n-1}) P_1^n$$

As $n \to \infty$, both $\tau_{0,n-1}$ and $\tau_{1,n-1}$ converge to 1. Since P_0^n, P_1^n and P_B^n are all bounded, we have that for any $\varepsilon > 0$, there is a n_ε such that $p_n^*(c') - p_n^*(c'') < \varepsilon$ for all $n > n_\varepsilon$.

An implication of Lemma 2 is that when n is large the optimal mechanism is characterized by a c_n^* such that for $c > c_n^*$, the required participation $a_n^*(c)$ is zero; and for $c \leq c_n^*$, participation is a non-increasing function which is approximately flat, even when the probability of success converges to a positive value. The next result shows that the utility obtained in such a mechanism converges to the utility that can be obtained in a binary mechanism. This fact combined with Lemma 1 implies that the optimal VBO is asymptotically optimal. Let V_n^G and V_n^* be the expected welfare generated in, respectively, the best general (VBO) and in the best (HO) mechanism when the number of agents is n. Putting this all together, we have:

Theorem 3. $\lim_{n\to\infty} V_n^G = \lim_{n\to\infty} V_n^*$.

This theorem allows us to rule out situations in which the limit probability of success is positive in the optimal (HO) mechanism, but is zero in the optimal VBO. An implication

of this is that whenever the limit probability of the optimal VBO converges to zero, then it converges to zero in *every* honest and obedient mechanism. We will use this fact in the next section to show that the limit probability in the best HO mechanism is the same as in the unorganized case.

5.3 A negative result: the irrelevance of an optimal organization in the limit (but only in the limit)

In our earlier analysis of the simple VBO mechanism (Proposition 6), we proved that in the limit large groups succeed with probability 1 if $m_n \prec n^{2/3}$ and large groups fail with probability 1 if $m_n \succ n^{2/3}$. However that left open the question of whether the optimal honest an obedient mechanism might succeed with positive probability for some values of $m_n \succ n^{2/3}$. Since the simple VBO is not necessarily optimal, it could be the case that the optimal mechanism does much better than a simple VBO in large groups. In fact it's possible that the optimal generalized VBO mechanism (which also might not be optimal) does much better than the simple VBO. In this section, we prove that the limiting properties of the simple VBO are shared by the optimal honest and obedient mechanism.

Let P_n^* denote the probability of success in the optimal honest and obedient mechanism as a function of n, for a given threshold m_n and value v. We have:

Theorem 4. *For any $v \in (0,1)$:*

- *If $m_n \prec n^{2/3}$, then $\lim_{n\to\infty} P_n^* = 1$.*
- *If $m_n \succ n^{2/3}$, then $\lim_{n\to\infty} P_n^* = 0$.*

Proof: The first bullet point is trivial, since the simple VBO achieves success with probability 1 in large groups from Proposition 6.

The bullet point is proved in two steps. The first step establishes that if $m_n \succ n^{2/3}$, then the probability of success in the best generalized VBO converges to zero. The second step prove that this implies that the probability of success in the best mechanism (not necessarily VBO) converges to zero as well. We outline the first step here and leave the second step for the online appendix.

Denote by $P_n^{\theta_n}$ the probability of success in the best general VBO with a threshold fraction equal to θ_n ($\geq \alpha_n$), so m_n members are activated if and only if there are at least $k_n = \theta_n n$ volunteers. Suppose $m_n \succ n^{2/3}$ and assume, by contradiction, that $\lim_{n\to\infty} P_n^{\theta_n} > 0$. It must be that:

$$\lim_{n\to\infty} \frac{1}{\theta_n} \cdot F\left(\frac{B(\theta_n n - 1, n-1, \theta_n)}{\sum_{j=\theta_n n-1}^{n-1} \frac{m_n}{j+1} B(j, n-1, \theta_n)}\right) \geq 1 \tag{18}$$

If this inequality is not satified then the threshold $c_n^{\theta_n}$ below wich an agent is willing to be activated is such that $F(c_n^{\theta_n})$ is lower than θ_n, making it impossible to achieve positive

probability in the limit as $n \to \infty$.[27] By the Mean Value Theorem we can write:

$$F\left(\frac{B(\theta_n n-1, n-1, \theta_n)}{\sum_{j=\theta_n n-1}^{n-1} \frac{m_n}{j+1} B(j, n-1, \theta_n)}\right) = F(0) + vf(\xi)\frac{B(\theta_n n-1, n-1, \theta_n)}{\sum_{j=\theta_n n-1}^{n-1} \frac{m_n}{j+1} B(j, n-1, \theta_n)}$$

where $\xi \in \left[0, \frac{B(\theta_n n-1, n-1, \theta_n)}{\sum_{j=\theta_n n-1}^{n-1} \frac{m_n}{j+1} B(j, n-1, \theta_n)}\right]$ and the last term uses the residual in the Lagrange form. Thus, if we define $\overline{f} = \max_{c\in[0,1]} f(c)$, we have:

$$\begin{aligned} & \frac{1}{\theta_n} \cdot F\left(\frac{B(\theta_n n-1, n-1, \theta_n)}{\sum_{j=\theta_n n-1}^{n-1} \frac{m_n}{j+1} B(j, n-1, \theta_n)}\right) \\ \leq \quad & \left(v\overline{f}\right) \cdot \frac{B(\theta_n n-1, n-1, \theta_n)}{\theta_n} \cdot \frac{1}{\sum_{j=\theta_n n-1}^{n-1} \frac{m_n}{j+1} B(j, n-1, \theta_n)} \end{aligned} \tag{19}$$

since $F(0) = 0$. Consider the second term. Following the same steps as in Theorem 1, we can write it as:

$$\frac{B(\theta_n n-1, n-1, \theta_n)}{\theta_n} \simeq \frac{1}{\theta_n}\sqrt{\frac{1}{2\pi\theta_n\left(1-\theta_n\right)n}} \simeq \sqrt{\frac{1}{\theta_n^3 n}}$$

where $\simeq$ means that it converges to zero at the same speed. Consider now the denominator of the third term, we can write:

$$\sum_{j=\theta_n n-1}^{n-1} \frac{m_n}{j+1} B(j, n-1, \theta_n) = \frac{\alpha_n}{\theta_n} \sum_{j=\theta_n n-1}^{n-1} \frac{\theta_n n}{j+1} B(j, n-1, \theta_n) \simeq \frac{\alpha_n}{\theta_n}$$

where the last step follows since $\sum_{j=1}^{n-1} \frac{\theta_n n}{j+1} B(j, n-1, \theta_n)$ can be shown to be bounded away from zero.[28]

Putting all of the above together, (18) can be written as:

$$\theta_n^3 n\left(\frac{\alpha_n}{\theta_n}\right)^2 \leq \left(v\overline{f}\right)^2 \Leftrightarrow \alpha_n^2\theta_n n \leq \left(v\overline{f}\right)^2 \Leftrightarrow \alpha_n^3 n\left(\frac{\theta_n}{\alpha_n}\right) \leq \left(v\overline{f}\right)^2$$

But we have $\alpha_n^3 n \cdot (\theta_n/\alpha_n) \geq \alpha_n^3 n$ since $\frac{\theta_n}{\alpha_n} \geq 1$. So if $\alpha_n^3 n > \left(v\overline{f}\right)^2$, then $\alpha_n^3 n \cdot (\theta_n/\alpha_n) > \left(v\overline{f}\right)^2$ as well. For $m_n \succ n^{2/3}$, we have that $\alpha_n^3 n = (m_n)^3/n^2 \to \infty$, implying that $\theta_n^3 n \cdot (\theta_n/\alpha_n)^2 \to \infty$ as well. We conclude that for $m_n \succ n^{2/3}$, (18) is not satisfied and the probability of success does not converge to a positive value. Essentially, having $\theta_n > \alpha_n$ helps making the probability of activation smaller (the second term); but it slows down convergence to zero too much for the first term.

For the second step (see Appendix) we rely on Theorem 3. By Theorem 3, welfare in

[27] When (18) is not satisfied, then $Y(\theta_n) > \theta_n$, so $Y(\theta_n)$ intersects the 45° line on the left of θ_n, implying that $\theta_n > p_n^{\theta_n} = F(p_n^{\theta_n})$.

[28] The fact can be formally proven following the same steps in Proposition 7, just replacing θ_n for α_n.

the best general VBO converges to the welfare in the best (HO) mechanism. But then, as $n \to \infty$, it is impossible that the the probabilities of success under the two mechanisms converge to different values. ■

We conclude this section discussing three implications of Theorem 4. The first is that the moral hazard problem is worse than previously assessed. If we believe that the groups can adopt a strong organization, then success will always be achieved if the total benefit for society $V_n = nv$ is larger than the cost in the worst case scenario in which $c_i = \overline{c} = 1$ for all $i \in I$. In our setting this situation emerges when $\alpha_n < v$; in other environments, such a situation emerges, for example, when total demand for a public good is bounded above (Hellwig [2003]). When we need to satisfy honesty and obedience, success is not generally guaranteed even in this case:

Corollary 1. *An optimally organized group fails to achieve its goal even if the total benefit is strictly higher than the cost in the worst case scenario in which* $c_i = \overline{c}$ for all $i \in I$ *if* $m \succ n^{-2/3}$.

The second implication of Theorem 4 is that the real benefit of an organization occurs for finite n, not in the limit. Theorem 4 shows that the optimal HO, general VBO and no organization all have the same expected per capita welfare in the limit if $m_n \prec n^{2/3}$ and 0 if $m_n \succ n^{2/3}$. The real benefit of having an efficient organization is derived by the fact that when the limit probability of success of an unorganized group converges to zero, the limit probability with an organized group converges to zero much slower. The next result bounds below this rate of convergence. We say that P_n^* converges at a strictly lower rate than exponential if $P_n^*/e^{-\upsilon n} \to \infty$ for any $\upsilon > 0$. We have:

Proposition 7. *For any* $m_n \succ n^{-2/3}$, P_n^* *converges to zero at a rate that is strictly slower than exponential.*

Finally, we have that:

Corollary 2. *The simple VBO is asymptotically optimal.*

The simple VBO has the appealing properties of being intuitive, asymptotically optimal and, when the probability of success converges to zero, has a very slow convergence rate.

6 Endogenous organizations: when do groups choose to organize?

The key insight in Olson [1965] is that we should expect successful collective action only when the free rider problem is not too severe: that is, for a given value v of the public good, when the number of interested agents n is not too large; or for a given n, when the value of the public good v is sufficiently large.[29] These observations motivate his claim that small groups with strong individual incentives will be much more effective than large groups in which individuals have weak incentives.

[29] See the discussion in chapters 1-2 in Olson [1965].

The results obtained above have implications that relate to these conjectures and also some additional insights about when one might expect successful groups to arise endogenously. On the one hand, the marginal impact of v or n on the probability of success depends on whether the group is organized or not and the extent to which the underlying technology displays increasing returns to scale: for example, without an organization, the marginal impact of v and n is exactly zero when n is large; it is positive only with an organization. So we cannot understand the true impact of individual preferences and the size of the population without first specifying their impact of the presence and quality of a group's organization.

On the other hand, whether the group might become organized or not depends on the underlying fundamentals of the economy, thus on v and n as well. We can evaluate the importance of these variables for the success of a group *only* when we include in the analysis their impact on the presence and effectiveness of an organization. To explore this idea, in this section we capitalize on the previous analysis to endogenize the presence of an organization and study how its endogeneity affects the impact of v and n on the ultimate success of a group.

We model the process of formation of an organization in a stylized, yet general way. Suppose the agents composing a group evaluate the opportunity of establishing an organization ex ante, before they know their individual costs of activism c. A VBO is formed if the increase in expected utility with the organization is larger than a given organizational fixed cost, κ: $\Delta V_n^* = V_n^O - V_n^U \geq \kappa(n)$, where V_n^O and V_n^U are, respectively, the expected utilities with and without an organization; and $\kappa(n)$ is the per person cost of forming and operating an organization with these n agents. We assume $\kappa(2) > 0$, and $\lim_{n\to\infty} \kappa(n) = \kappa > 0$.[30]

The expected utility in a VBO with cut off c_n^O can be written as:

$$V_n^O = v\left[F\left(c_n^O\right) p_{1,n}(c_n^O) + (1 - F\left(c_n^O\right))p_{2,n}(c_n^O)\right] - F\left(c_n^O\right) E\left(c; c_n^O\right) a_n^*, \tag{20}$$

where $p_{1,n}(c_n^O)$ and $p_{2,n}(c_n^O)$ are the probability of success for an agent, conditioning on being a volunteer and on not being a volunteer; and $E\left(c; c'\right)$ is the expected c, conditioning on not being larger than c'. In a VBO, we must have $c_n^O a_n^* = v(p_{1,n}(c_n^O) - p_{2,n}(c_n^O))$, the expected utility with an optimal organization can therefore be written as:

$$V_n^O = v\left[\begin{array}{c} F\left(c_n^O\right)\left(1 - \frac{E\left(c;c_n^O\right)}{c_n^O}\right) B(m_n - 1, n-1, F\left(c_n^O\right)) \\ + \sum_{j=m_n}^{n-1} B(j, n-1, F\left(c_n^O\right)) \end{array}\right] \tag{21}$$

[30]This simple model is intended capture a variety of environments. Consider these two polar examples. First, assume perfect substitutability and that there is an elite of $l \leq n$ agents each of whom can form the organization paying a fixed cost $\widehat{\kappa}$. The organization is created if at least one member of the elite pays the cost; if the organization is formed, the members of the elite capture a share $v \leq 1$ of the total surplus $n\Delta V^*$. In this case, there is an equilibrium in which each member of the elite pays the cost with probability $\phi < 1$ and the condition for the establishment is $\Delta V^* \geq \kappa$ with $\kappa = \widehat{\kappa}/\left[nB(0, l-1, \phi)v\right]$. The elite members internalize only a share of the benefit because they themselves may face a free rider problem.

Second, assume perfect complementarity in the technology for the formation of the organization, so that each member of the elite needs to pay a cost $\widehat{\kappa}$. In this case the organization forms if and only if $\Delta V^* \geq \kappa$ is satisfied with $\kappa = \widehat{\kappa}/nv$ as described in the main text.

Following similar steps, the expected utility without an organization can be written as:

$$V_n^U = v \left[\begin{array}{c} F\left(c_n^U\right)\left(1 - \frac{E\left(c;c_n^U\right)}{c_n^U}\right) B(m_n - 1, n-1, F\left(c_n^U\right)) \\ + \sum_{j=m_n}^{n-1} B(j, n-1, F\left(c_n^U\right)) \end{array} \right] \tag{22}$$

The effect of n on ΔV_n^* is complicated by the fact that n indirectly affects the thresholds for equilibrium participation c_n^O and c_n^U. Still, from the continuity of the functions in the square parentheses with respect to c_n^O and c_n^U, and the fact that we know for n large enough $c_n^U = 0$ and $c_n^O \to 0^+$, we can deduce the no organization will ever be formed for arbitrarily large groups:[31]

Proposition 8. *There is a $n_\kappa > 0$ such that a VBO is formed only if $n \leq n_k$.*

A similar discontinuity as highlighted above is generated by a change in v if we keep n constant. Again, signing the comparative statics in full generality is difficult because it involves evaluating how the mechanism cutoff for volunteers, c^O, changes relative to c^U as we change v. However, the effect can easily be signed when n is sufficiently large. We have:

Proposition 9. *There is a $n^* > 0$ such that for any $n > n^*$, ΔV_n^* is strictly increasing in v, so an organization is formed only if v is larger than a threshold v_n^*.*

Propositions 8 and 9 are interesting because they suggests why the two factors highlighted by Olson (size and individual incentives) matter for a group's effectiveness. It is not just that as n increases or v decreases, we have a more severe free rider problem that depresses the probability of success. If it were only for this, the probability of success would change very little. A more important point is that the group organizes only for $n \leq n_\kappa$ and this has important implications for effectiveness.[32] As n passes n_κ, effectiveness collapses to almost zero, since without an organization the probability of success is extremely small and insensitive of v and n. Proposition 8 also explains why we should expect a dichotomy of organizations: the small, organized, and effective groups on the one hand; and large, unorganized, and ineffective group on the other hand. What creates the dichotomy is the decision to organize that transform a continuous effect in a discrete drop in effectiveness.

7 Variations and discussions

7.1 On High-value environments ($v \geq 1$)

An assumption that we have maintained throughout the analysis is that $v < 1$, where 1 is the highest possible cost $\overline{c}$. This assumption is standard and implied by the stronger assumption

[31]Note that the fact that c_n^U and c_n^O converge to zero does not imply the terms in parenthesis converge to zero; indeed, as we know form Proposition 6, they both converge to zero if $m_n \succ n^{2/3}$ and to one if $m_n \prec n^{2/3}$.

[32]The size threshold, n_κ, as well as the value threshold, v_n^*, both vary with the returns to scale. In principle, n_κ could be quite large.

made, for example, by Mailath and Postlewaite [1990] requiring that the total benefit of success vn is less than the marginal cost αn in the worst case scenario in which all types have a cost of 1, so that $v < \alpha$. The case with $v \geq 1$, however, has an interesting peculiarity that is worth discussing. If $v \geq 1$, then, for weak organizations we obtain a result similar to Theorem 2, because randomly selecting a group of size m_n, regardless of individual costs, does not violate the obedience constraint for any type. Specifically:

Proposition 10. *If $v \geq 1$, MHRA is satisfied, and either $m_n \prec n$ or $m_n = \alpha n$ for some fixed $\alpha < 1$, then for all n the optimal direct mechanism satisfying (IC) and (IMH) is a random mechanism in which each g such that $|g| = \alpha n$ is activated with probability $1/\binom{n}{m_n}$ and each g such that $|g| \neq m_n$ is activated with probability 0. The probability of success equals 1.*

When $v \geq 1$, however, the limit probability of success in the symmetric equilibrium of an unorganized group remains 0 when $m \succ n^{2/3}$. An implication of Proposition 9, therefore, is that the limit equivalence of the probabilities of success in organized and unorganized groups is not valid anymore when $v \geq 1$. In this case, moral hazard is not a problem; the only strategic problem faced by the members is coordination. Coordination can be easily solved by a honest and obedient mechanism, but is unsolvable in a symmetric equilibrium without an organization.[33]

7.2 On the divisibility of tasks

In the previous analysis we have assumed that the decision to contribute is dichotomous: agent i either contributes at a cost c_i or not. For example, an agent participates in a rally or not; an agent signs a petition or not; joins a union or a committee or not. There are however cases in which the contribution can be split up. For example, suppose that an agent has up to one day to donate to a cause, say the organization of a charity. However, if the agent cannot donate one day, perhaps the agent can donate less, say one hour. It is easy to see that the analysis can be easily extended to this case, though the results are interesting only when we assume some economies of scale, if the task cannot be "atomized" too much relative to the cost of providing the effort: in this case the obedience constraint becomes moot (and the optimal mechanism becomes too powerful to generate plausible predictions).

To see this, assume that a contribution now can be divided in λ parts: when $\lambda = 1$ the contribution is, say, one day; when $\lambda = 24$, it is one hour, etc. Now the mechanism can ask each agent to contribute any discrete amount $x \in \{0, 1/\lambda, 2/\lambda, ..., 1\}$, say from 0 hours to 24 hours. Assume that we need a total of m_n contribution units to achieve the collective goal and recall that the costs c_is are distributed in $[0, \overline{c}]$: now we can require m_n agents providing one unit, or up to λm_n agents providing one hour. If we can choose λ so large that $\overline{c}/\lambda < v$ and $\lambda m \leq n$, then we can achieve the common goal with probability one with a mechanism equivalent to the optimal (IC) and (INTIR) mechanism of Theorem 4. In this case, we just ask λm_n agents at random to contribute $1/\lambda$ each. If $m_n \prec n$, then for any λ such that $\overline{c}/\lambda < v$, it will be true that $\lambda m_n \leq n$ for large n, so success with probability 1 is feasible

[33] Of course, we can design an asymmetric equilibrium that achieves success with probability 1 in an unorganized group, but such an equilibrium would implicitly assume a solution of the coordination problem by ex ante selecting the "volunteers".

for n large enough. In many plausible environments, however, economies of scale make it unrealistic to assume that divisibility is fine enough to guarantee that *all* types including the most extreme would be willing to obey it if asked. If we assume that there is a $\overline{\lambda}$ satisfying $\overline{c}/\overline{\lambda} > v$, then the obedience constraint will always be binding as in the analysis presented above. For instance, this is always true for any $\overline{\lambda}$, for $\overline{c}$ large enough.

The analysis in previous sections carries over to this case where contributions can be discretely divisible rather than just dichotomous. In the dichotomous case a (reduced form) mechanism specifies a probability of success $p(c)$ and a probability of contributing $a(c)$, where $a(c) \in [0,1]$ and non-increasing in c. As before, a mechanism specifies an interim probability of success $p(c)$ and an interim *expected contribution* $a(c)$, where again $a(c) \in [0,1]$ and non-increasing in c. The analysis is completely analogous. Indeed, the same logic as in Section 5 suggests that, for finite n, $a(c)$ will be non increasing and positive up to a threshold $c^* = \min\{c \leq vvp(c^*) - c^* a(c^*) \leq vp(\overline{c})\}$ and then $a(c) = 0$ for $c > c^*$, just as above. Moreover, $a(c)$ will become flat as $n \to \infty$, so a VBO with $a(c) > a$ for "volunteers" and zero for free riders with a higher c will be asymptotically optimal, just as in the previous analysis.

7.3 The Two-type case

The result that generalized VBO mechanisms are optimal binary mechanisms, together with the fact that generalized VBO mechanisms are also approximately optimal HO mechanisms for large groups, suggests a connection between organizational solutions to the free rider problem when members have continuous types and when there are only two possible types. In this section we analyze the two-type case, using similar techniques as with the optimal binary mechanism problem. A key difference however is that in the continuous case, the volunteer threshold, c_n^*, is endogenously determined as a function of the mechanism through the honest and obedience constraints, and also depends on n Here we assume there are two types: a low type $\underline{c} \in (0, v)$ with probability $\phi \in (0,1)$; and a high type $\overline{c} > \underline{c}$ with probability $1 - \phi$. With only two types, the volunteer threshold can only be one of two possible values, $\underline{c}$ or $\overline{c}$.[34]

We only consider the case where $\underline{c} < v < \widetilde{c}$.[35] Since $\overline{c} > v$, HO requires $a(\overline{c}) = 0$. Using the notation of Section 2.2, a direct mechanism μ in this case is described in its reduced form by 3 numbers: $a = a(\underline{c})$, $p_1 = p(\underline{c})$, and $p_2 = p(\overline{c})$.

We solve for the optimal HO mechanism by considering a relaxed problem and then proving that its solution satisfies all constraints of the original problem. The relaxed problem requires 3 necessary conditions for an (HO) mechanism: feasibility; the fact that $c > v \to a(c) = 0$; and (IC), which in this case with two types can be written as:

$$a\underline{c} \leq v(p_1 - p_2) \tag{23}$$

[34]See Ledyard and Palfrey (1994) for an generalized VBO characterization of the optimal mechanism with strong organizations (INTIR) for the two type case with side payments.

[35]If $\underline{c} > v$ obedience implies $a = 0$ for both types. If $\overline{c} < v$ the random mechanism identified in Proposition 10 is optimal.

We start with a preliminary observation that any mechanism that solves the relaxed problem must be non-wasteful: whenever a group is activated, there are exactly $m = \alpha n$ members in the activated group; and groups with fewer than αn members are never activated. We prove this by contradiction by supposing that μ is wasteful at a positive measure set of profiles and then showing that it can be improved. First, define a new mechanism, μ', that is exactly the same as μ for all coalitions of size m, but eliminates all waste by reducing all activated successful coalitions to a size m by randomly selecting agents to drop out, and do not activate coalitions that are smaller than m. The resulting mechanism is feasible. Moreover, it continues to satisfy the (IC) constraint. This leaves p_1^μ and p_2^μ unchanged and reduces a^μ to $a^{\mu\prime} < a^\mu$. This implies that $a^{\mu\prime}\underline{c} < v(p_1^\mu - p_2^\mu)$. Note that after the modification the mechanism must be characterized by a set of probabilities q_k such that with $k \geq \alpha n$ low types, the probability that a randomly selected coalition of size αn is activated is q_k. If $q_k = 1$ for all $k \geq \alpha n$, then we have proven the result. Assume therefore that $q_k < 1$ for some $k \geq \alpha n$. But then we can increase q_k to $q_k' > q_k$ in order to increase $a^{\mu\prime}$ to $a^{\mu\prime, q_k'}$ and still satisfy $a^{\mu\prime, q_k'}\underline{c} \leq v(p_1^\mu - p_2^\mu)$. After this change the probability of success and welfare are higher, a contradiction with the assumption that μ is optimal.

Now note that a non-wasteful mechanism that solves the relaxed problem must be characterized by a set of probabilities q_k such that with $k \geq \alpha n$ low types, the probability that a randomly selected coalition of size αn is activated is q_k. We can therefore easily write the feasibility conditions as follows:

$$p_1 = \sum\nolimits_{k=\alpha n-1}^{n-1} B(k, n-1, \phi) q_{k+1},\ p_2 = \sum\nolimits_{k=\alpha n}^{n-1} B(k, n-1, \phi) q_k \tag{F2}$$
$$\text{and } a = \sum\nolimits_{k=\alpha n-1}^{n-1} \frac{\alpha n}{1+k} B(k, n-1, \phi) q_{k+1}$$

And thus the (IC) can be written in terms of the q_k as:

$$\Phi(q) = \underline{c}\left[\sum\nolimits_{k=\alpha n-1}^{n-1} \frac{\alpha n}{1+k} B(k, n-1, \phi) q_{k+1}\right] - v\left[\begin{array}{c} B(n-1, n-1, \phi) q_n \\ + \sum_{k=\lfloor \alpha n \rfloor}^{n-1} \left[\begin{array}{c} B(k-1, n-1, \phi) \\ -B(k, n-1, \phi) \end{array}\right] q_k \end{array}\right] \leq 0 \tag{24}$$

The relaxed problem can therefore consists in maximizing expected welfare $\phi\,(vp_1 - \underline{c}a) + (1-\phi)\,vp_2$ subject to (24) and (F2). This problem can be written as:

$$\max_{\{q_k\}_{k=\alpha n}^{n}} \left\{ \begin{array}{c} \phi \sum_{k=\alpha n-1}^{n-1} B(k, n-1, \phi) q_{k+1} \left(v - \frac{\alpha n}{1+k}\right) \\ + (1-\phi)\, v \sum_{k=\alpha n}^{n-1} B(k, n-1, \phi) q_k \end{array} \right\}$$
$$\text{s.t. } (24)$$

We next show that the optimal honest and obedient mechanism is a generalized threshold mechanism, and that it is characterized by just one easily computable real number $\kappa \in [0, n]$.[36] To this goal note that, for any general VBO described by $\{q_k\}_{k=m_n}^{n}$, we can find a κ such that $q_k = 0$ if $k < \lfloor \kappa \rfloor$, $q_k = 1$ if $k < \lceil \kappa \rceil$, and $q_k = \lceil \kappa \rceil - \kappa$ if $k = \kappa$, where $\lceil \kappa \rceil$ and $\lfloor \kappa \rfloor$

[36] The definition of a generalized threshold mechanism in the two-type case is the same as defined earlier, with $c^* = \underline{c}$.

are respectively the least integer larger than κ and the greatest integer smaller than κ. We say that a threshold κ is consistent with (24) if the corresponding q_k satisfies it. We have:

Theorem 5. *If $\underline{c} < v < \overline{c}$, the optimal mechanism is the generalized VBO characterized by the smallest threshold κ consistent with (24).*

Proof: The proof that the optimal mechanism is a generalized VBO follows a similar logic to the proof of Theorem 1. The fact that the best mechanism corresponds to the smallest threshold follows from the fact that the objective function is decreasing in κ. See appendix for details. ■

With this characterization, we can also answer the question: When is the simple VBO *exactly* optimal for finite values of n (i.e. $q_{\alpha_n n} = 1$ or $\kappa = 0$)? The following proposition provides a sufficient condition:

Corollary 3. *If $0 < \underline{c} < v < \overline{c}$ and $\frac{\underline{c}}{v} < \frac{1-\frac{\phi}{\alpha}}{1-\phi}$, then a simple VBO is optimal.*

Corollary 3 only shows that $\frac{\underline{c}}{v} < \frac{1-\frac{\phi}{\alpha}}{1-\phi}$ is a sufficient condition for a simple VBO mechanism to be optimal; it does not show that $\frac{\underline{c}}{v} \geq \frac{1-\frac{\phi}{\alpha}}{1-\phi}$ is a necessary condition, since it is possible that $\Psi(q) > 0$ for the mechanism in which $k^* = m$ and $q_m = 1$, in which case the (IC) constraint would be slack under the simple VBO mechanism.

While the results above show that the two type and the continuum model share some properties, they differ in their asymptotic properties. Specifically, we show below that with two types, if $m_n \succ n^{1/2}$, the probability of success of the *every* HO mechanism converges 0 in the 2-type model, while it converges to 1 in the continuum model if $m_n \prec n^{2/3}$.

Proposition 11. *If $0 < \underline{c} < v < \overline{c}$ and $m_n \succ n^{1/2}$, then $\lim_{n\to\infty} P_n \to 0$*

The intuition for this result is that for an HO mechanism to result in a positive probability of success for arbitrarily large n, the low cost type, $\underline{c}$, cannot be strictly positive.[37] This also demonstrates, once again, the sharp contrast between strong (INTIR) and weak (HO) organizations, since weak organizations succeed with probability 1 in the limit, for $m_n \prec n$.

8 Conclusions

We have developed a model of collective action in which a group can organize by constructing communication mechanisms to elicit private information and coordinate the actions of its members. We have stressed the importance of requiring the mechanism to be obedient, besides the more familiar requirements of incentive compatibility and individual rationality. Mechanisms that are only incentive compatible and individually rational make sure that members are willing to join a group and reveal their types, but they require members to commit to carry out the mechanism's recommendations, thus assuming away a key aspect of

[37] We conjecture that the probability of success for HO mechanisms converges to 0 in the two-type model *for all* $\beta > 0$. However, our proof here only establishes the result for $\beta > 1/2$.

the moral hazard problem. Obedience is not generally included in classic mechanism design problems, since in these applications mechanisms map vectors of type profiles to allocations; in collective action problems, on the contrary, mechanisms only map vectors of type profiles to recommendations: allocations are the decentralized results of the members' individual actions.

Except when we assume constant returns to scale in the technology underlying the group's effectiveness, an incentive compatible and individually rational mechanism that omits the obedience constraint can fully solve a group's problem, achieving full success even with no side payments. Moreover, regardless of how fast m_n grows, success is always guaranteed when the total benefit of the group is larger than the cost in the worst case scenario (in which all members have maximal costs). The predicted success of optimal honest and obedient mechanisms is typically much more modest, but perhaps more realistic. We showed that when m_n grows slower than $n^{2/3}$, success is achievable with certainty, with or without an organization. When m_n grows faster than $n^{2/3}$, however, success is impossible in the limit even if total benefit is always larger than total cost (a result that generalizes previous negative results proven only for the constant returns to scale case and assuming that total benefits are smaller than maximal total cost). Still, we showed that endowing a group with an honest and obedient organization gives a group a key advantage. The real benefit of an organization is that even when the probability of success converges to zero, it does so at a much slower rate than without an organization (which indeed is exactly zero after a finite n). The rate of convergence of the probability of success when m_n grows faster than $n^{2/3}$ is always strictly slower than exponential.

There are numerous ways the theory might usefully be extended. In our analysis we have relied on a very simple base model of collective action, a classic threshold public good game. It may be possible to explore the themes described above in much more general economic environments in which the size of the common goal that can be chosen by the collectivity is a continuous variable: as, for example, when the group does not only choose to build a bridge, but also its quality and its capacity. In addition, we have studied a completely static model. Many collective action problems are dynamic. The ideas presented here could be embedded in dynamic environments to extend previous work that has studied contribution games in dynamic environments with no organizations (see, for instance, Matthews [2013] and Battaglini et al. [2014]).

An important dimension of the problem that we plan to explore in future work is the study of how groups faced with multiple collective action problems strategically interact with each other. Groups may strategically interact because their respective goals are substitutes, as when there is a budget constraint that allows only a subset of projects to be realized. Or they can interact in environments with complementarities, which leads to "a collective action problem in a collective action problem": that is, the groups need to solve a collective action problem among themselves in the face of common goals, but each group also needs to solve its own internal collective action problem in order for the group to make a contribution.

The theory presented here also provides inspiration for new empirical questions that can be studied with laboratory experiments and possibly field work. We mentioned a significant literature in experimental economics that has studied contribution games with structured and unstructured preplay communication. Most of this literature has focused on environments

with complete information, or with only a few players. We leave for future research an empirical investigation of the effectiveness of the VBOs characterized in this paper and the comparison of their performance with unorganized groups.

References

[1] Archetti, M., and Scheuring, I. (2012). "Game theory of public goods in one-shot social dilemmas without assortment." *Journal of Theoretical Biology*, 299, 9–20.

[2] Ash, R. (1990). *Information Theory*. Dover Pubblications, New York.

[3] d'Aspremont, Claude and L.-A. Gérard-Varet (1979), "Incentives and Incomplete Information", *Journal of Public Economics*, 11, 25-45.

[4] d'Aspremont C., J. Crémer, and L. A. Gérard-Varet (1990), "Incentives and the existence of Pareto-optimal revelation mechanisms", *Journal of Economic Theory*, 51(2): 233-254.

[5] Barrett, S. and Dannenberg, A. (2014). "Sensitivity of Collective Action to Uncertainty about Climate Tipping Points," *Nature Climate Change*, 4(1), 36–39.

[6] Battaglini M. and R. Benabou (2003): "Trust, Coordination, and The Industrial Organization of Political Activism," *Journal of the European Economic Association*, 1(4), 2003.

[7] Battaglini M. (2017), "Public Protests and Policy Making," *Quarterly Journal of Economics*, 2017, 132(1): 485-549.

[8] Battaglini M. and B. Harstad (2016), "Participation and Duration of Environmental Agreements", *Journal of Political Economy*, 2016, 124(1): 160-204.

[9] Battaglini M., S. Nunnari and T. Palfrey (2014), "Dynamic Free Riding with Irreversible Investments," *American Economic Review*, 104(9): 2858-71.

[10] Bierbrauer, F. and M. Hellwig (2016), "Robustly Coalition-Proof Incentive Mechanisms for Public Good Provision are Voting Mechanisms and Vice Versa", *Review of Economic Studies*, 83: 1440–1464.

[11] Braver, S. L., and L. A. Wilson (1986), "Choices in Social Dilemmas: Effects of Communication within Subgroups." *The Journal of Conflict Resolution*, 30(1):51–62.

[12] Chamberlin, J. R. (1974), "Provision of Collective Goods as a Function of Group Size", *American Political Science Review*, 68: 707.

[13] Cremer, J. and McLean, R.P. (1988), "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56, 1247-1257.

[14] Diekmann, A. (1985), "Volunteer's dilemma", *Journal of Conflict Resolution*, 29: 605-610.

[15] Dixit, A. and M. Olson (2000), "Does voluntary participation undermine the Coase Theorem?" *Journal of Public Economics*, 76 (3): 309-35.

[16] Dziuda W. A. Gitmez and M. Shadmeher (2021), "The Difficulty of Easy Projects," *American Economic Review: Insights*, 3(3): 285–302.

[17] Esteban J. and D. Ray (2001), "Collective action and the group paradox", *American Political Science Review*, 95(3):663–672.

[18] Goldlucke, Susanne and Thomas Troger (2020), The Multiple-Volunteers Dilemma. *Manuscript*, University of Mannheim.

[19] Green, Jerry R. and Jean-Jacques Laffont (1979), *Incentives in public decision making.* North-Holland,1979.

[20] Hellwig, Martin F (2003), "Public Good Provision with Many Participants", *The Review of Economic Studies*, 70: 589-614.

[21] van de Kragt A., J. Orbell and R. Dawes (1983), Organizing groups for collective action, *American Political Science Review*, 80(4), 1171-1185.

[22] Ledyard, John O. and Thomas R. Palfrey (1994), Voting and Lottery Drafts as Efficient Public Goods Mechanisms. *The Review of Economic Studies*, 61(2): 327-355.

[23] Ledyard, John O. and Thomas R. Palfrey (1999), A Characterization of Interim Efficiency with Public Goods, *Econometrica*, 67(2):435-48.

[24] Ledyard, John O. and Thomas R. Palfrey (2002), The Approximation of Efficient Public Good Mechanisms by Simple Voting Schemes. *Journal of Public Economics*, 83(2):153-72.

[25] Ledyard, John O. and Thomas R. Palfrey (2007), A General Characterization of Interim Efficient Mechanisms for Independent Linear Environments, *Journal of Economic Theory*, 133(1):441-66.

[26] Mailath, George J. and Andrew Postlewaite (1990), Asymmetric Information Bargaining Problems with Many Agents. *Review of Economic Studies*, 57:351-67.

[27] Matthews, S. (2013), "Achievable Outcomes of Dynamic Contribution Games," *Theoretical Economics*, 8(2): 365-403.

[28] Myerson, Roger B. (1982), Optimal Coordination Mechanisms in Generalized Principal-Agent Problems. *Journal of Mathematical Economics* 10: 67-81.

[29] Nöldeke G. and J. Peña (2020), "Group size and collective action in a binary contribution game," *Journal of Mathematical Economics*, 88: 42-51.

[30] Olson, Mancur (1965), *The Logic of Collective Action.* Harvard University Press: Cambridge.

[31] Ostrom E. (1990), *Governing the Commons*, Cambridge University Press.

[32] Ostrom E. and J. Walker (1990), "Communication in a Commons: Cooperation Without External Enforcement," *Contemporary Laboratory Research in Political Economy*, ed. Tom Palfrey, p. 287-322, University Of Michigan Press.

[33] Ostrom E., J. Walker and R. Gardner (1992), "Covenants With and Without a Sword: Self-Governance is Possible, *American Political Science Review*, 86(2): 404-417.

[34] Palfrey, T.R. and H. Rosenthal (1984), "Participation and the Provision of Public Goods: a Strategic Analysis." *Journal of Public Economics*, 24:171–193.

[35] Palfrey, T. R. and H. Rosenthal (1991). Testing for effects of cheap talk in a public goods game with private information, *Games and Economic Behavior*, 3:183-220.

[36] Palfrey, T. R., H. Rosenthal, and N. Roy (2017), How Cheap Talk Enhances Efficiency in Threshold Public Goods Games. *Games and Economic Behavior*, 101: 234-59.

[37] Passarelli F. and G. Tabellini (2017), "Emotions and Political Unrest," *Journal of Political Economy*, 125(3):903-946.

[38] Rob, R. (1989), "Pollution Claim Settlements with Private Information", *Journal of Economic Theory*, 47, 307- 33.

[39] Simmons R. (1980), "The Logic of Cooperation? An Investigation of the Free Rider Hypothesis". Ph.D. dissertation, University of Oregon, Eugene.

[40] Tavoni A., A. Dannenbergb, G. Kallisc, and A. Löschelb (2011), "Inequality, communication, and the avoidance of disastrous climate change in a public goods game," *Proceedings of the National Academy of Science*, 108(29): 11825-11-829.

9 Appendix

9.1 Proof of Theorem 1

In Step 1, we show that $m_n \prec n^{2/3}$ implies that c_n^U is sufficiently larger than α_n for all n and in the limit. This guarantees that $\lim_{n\to\infty} P_n^U = 1$. In Step 2, we show that if $m_n \succ n^{2/3}$, then $\lim_{n\to\infty} P_n^U = 0$.

Step 1. We proceed in three steps.

Step 1.1. We first show that $m_n \prec n^{2/3}$ implies that $c_n^U > \alpha_n$ for n sufficiently large. Recall that the threshold, c_n^U, is the largest solution for $c \in [0,1]$ to the following equation: $c = F(vB(\alpha_n - 1, n-1, c))$. Therefore, $c_n^U > \alpha_n$ for any n large if, for sufficiently large n we have:

$$\alpha_n < F(vB(\alpha_n - 1, n-1, \alpha_n))$$

This condition guarantees that there is an intersection on the right of α_n. See Figure 1. The following lemma will prove useful in the argument.

Lemma A1. *If* $m_n \succ n^{2/3}$ then $B(\alpha_n - 1, n-1, \alpha_n)/\alpha_n \to 0$.

Proof. We can approximate the binomial combinatorial term for large n using Stirling's formula: $\binom{n}{k} = \sqrt{\frac{n}{2\pi k(n-k)}}\frac{n^n}{k^k(n-k)^{n-k}}$. First note that:

$$\binom{n-1}{\alpha n - 1} = \frac{(n-1)!}{(\alpha_n n - 1)!(n - \alpha_n n)!} = \frac{\alpha_n n}{n}\frac{n(n-1)!}{\alpha_n n\,(\alpha_n n - 1)!(n - \alpha_n n)!} = \frac{\alpha_n n}{n}\binom{n}{\alpha_n n} = \alpha_n \binom{n}{\alpha_n n}$$

Applying Stirling's formula yields:

$$\binom{n-1}{\alpha_n n - 1} \simeq \alpha_n \sqrt{\frac{1}{2\pi\alpha_n(1-\alpha_n)n}}\left[\frac{1}{(\alpha_n)^{\alpha_n}\,(1-\alpha_n)^{(1-\alpha_n)}}\right]^n$$

where $\simeq$ means that the two sequences converge to zero at the same speed. So an approximation of $B(\alpha_n n - 1, n-1, \alpha_n)$ is given by:

$$B(\alpha_n n - 1, n-1, \alpha_n) = \binom{n-1}{\alpha_n n - 1}\frac{\left[(\alpha_n)^{\alpha_n}\,(1-\alpha_n)^{1-\alpha_n}\right]^n}{\alpha_n} \simeq \sqrt{\frac{1}{2\pi\alpha_n(1-\alpha_n)n}}$$

We therefore have:

$$\frac{B(\alpha_n n - 1, n-1, \alpha_n)}{\alpha_n} \simeq \frac{1}{\alpha_n}\sqrt{\frac{1}{2\pi\,(\alpha_n)\,(1-\alpha_n)n}}v \to z \geq 1 \Leftrightarrow \frac{1}{2\pi\,(\alpha_n)^3\,(1-\alpha_n)n} \geq 1/v^2$$

The last inequality can be written as: $\left(\frac{m_n}{n}\right)^3(1 - \frac{m_n}{n})n \leq \frac{v^2}{2\pi}$. As $n \to \infty$, this condition is satisfied if $m_n/n^{2/3} \to 0$, or $m_n \prec n^{2/3}$. ■

We can now prove that $m_n \prec n^{2/3}$ implies that $c_n^U > \alpha_n$ for all n. To this goal, note that

as $n \to \infty$, $vB(\alpha n-1, n-1, \alpha_n) \to 0$, so we can write

$$F\left[vB(\alpha n-1, n-1, \alpha_n)\right] = vf(0) \cdot B(\alpha n-1, n-1, \alpha_n) + o(B(\alpha n-1, n-1, \alpha_n))$$

It follow that:

$$\begin{aligned}
\lim_{n\to\infty} \frac{F\left[vB(\alpha n-1, n-1, \alpha_n)\right]}{\alpha_n} &= \lim_{n\to\infty} \left[vf(0) \cdot \frac{B(\alpha n-1, n-1, \alpha_n)}{\alpha_n} + \frac{o(B(\alpha n-1, n-1, \alpha_n))}{\alpha_n}\right] \\
&= \lim_{n\to\infty} \left[vf(0) + \frac{o(B(\alpha n-1, n-1, \alpha_n))}{B(\alpha n-1, n-1, \alpha_n)}\right] \frac{B(\alpha n-1, n-1, \alpha_n)}{\alpha_n} \\
&= vf(0) \lim_{n\to\infty} \frac{B(\alpha n-1, n-1, \alpha_n)}{\alpha_n}
\end{aligned}$$

We conclude that whenever $\frac{B(\alpha n-1,n-1,\alpha_n)}{\alpha_n}$ converges to 0 or diverges at ∞, so does $\frac{F[vB(\alpha n-1,n-1,\alpha_n)]}{\alpha_n}$.

Step 1.2. We now prove that if $m_n \prec n^{2/3}$ the ratio c_n^U/α_n becomes arbitrarily large. The following lemma will prove useful in the argument.

Lemma A2. *There is a sequence φ_n such that $\varphi_n \to \infty$, $\varphi_n \cdot \alpha_n \to 0$ and $\frac{m_n}{\varphi_n} \to \infty$.*

Proof. Let $\varphi_n = (m_n)^\epsilon$ for $\varepsilon \in (0, 1/3)$. Naturally, it must be that $\varphi_n \to \infty$ since $m_n \to \infty$, and that $\frac{m_n}{\varphi_n} = m_n^{1-\varepsilon} \to \infty$; moreover, we have $\varphi_n \cdot \alpha_n = \varphi_n \cdot \frac{m_n}{n} = \left(\frac{m_n}{n}\right)^\epsilon \cdot \frac{m_n}{n^{1-\epsilon}} \to 0$ since $m_n \prec n^{2/3} \prec n^{1-\varepsilon} \prec n$. ■

We now show that:

$$\alpha_n \cdot \varphi_n < vB(\alpha_n - 1, n-1, \alpha_n \cdot \varphi_n),$$

where φ_n has the properties desciebed in Lemma A2. Geometrically, this implies that at $\alpha_n \cdot \varphi_n$ we are on the left of the intersection, so $\alpha_n \cdot \varphi_n < c_n^U$: this implies $c_n^U/\alpha_n > \varphi_n \to \infty$. To this goal, we can write:

$$\begin{aligned}
vB(\alpha_n - 1, n-1, \alpha_n \cdot \varphi_n) &= v\binom{n-1}{\alpha_n n - 1} \frac{\left[(\varphi_n\alpha_n)^{\alpha_n}(1-\varphi_n\alpha_n)^{1-\alpha_n}\right]^n}{\varphi_n\alpha_n} \\
&\simeq v\alpha_n \sqrt{\frac{1}{2\pi\alpha_n(1-\alpha_n)n}} \left[\frac{1}{(\alpha_n)^{\alpha_n}(1-\alpha_n)^{(1-\alpha_n)}}\right]^n \\
&\cdot \frac{\left[(\varphi_n\alpha_n)^{\alpha_n}(1-\varphi_n\alpha_n)^{1-\alpha_n}\right]^n}{\varphi_n\alpha_n} \\
&= v\frac{1}{\varphi_n} \sqrt{\frac{1}{2\pi\alpha_n(1-\alpha_n)n}} \left[(\varphi_n)^{\alpha_n} \cdot \frac{(1-\varphi_n\alpha_n)^{1-\alpha_n}}{(1-\alpha_n)^{(1-\alpha_n)}}\right]^n
\end{aligned}$$

Consider first the term:

$$(1-\varphi_n\alpha_n)^{(1-\alpha_n)n} = \left[(1-\varphi_n\alpha_n)^{\frac{1}{\varphi_n\alpha_n}}\right]^{\varphi_n\alpha_n(1-\alpha_n)n} \simeq \left(\frac{1}{e}\right)^{\varphi_n m_n}$$

Consider now the term:

$$(1-\alpha_n)^{(1-\alpha_n)n} \simeq \left[(1-\alpha_n)^{\frac{1}{\alpha_n}}\right]^{\alpha_n(1-\alpha_n)n} \simeq \left[(1-\alpha_n)^{\frac{1}{\alpha_n}}\right]^{\alpha_n n} \simeq \left(\frac{1}{e}\right)^{m_n}$$

It follows that:

$$\frac{(1-\varphi_n\alpha_n)^{(1-\alpha_n)n}}{(1-\alpha_n)^{(1-\alpha_n)n}} \to \left(\frac{1}{e}\right)^{(\varphi_n-1)m_n}$$

We can now write:

$$\begin{aligned} vB(\alpha_n-1,n-1,\alpha_n\cdot\varphi_n) &= \frac{v}{\varphi_n}\sqrt{\frac{1}{2\pi\alpha_n(1-\alpha_n)n}}\left[(\varphi_n)^{\alpha_n}\cdot(e)^{(1-\varphi_n)\alpha_n}\right]^n \\ &= v\varphi_n\sqrt{\frac{1}{2\pi\alpha_n(1-\alpha_n)n}}\cdot\frac{1}{(\varphi_n)^2}\left[(\varphi_n)^{\alpha_n}\cdot(e)^{(1-\varphi_n)\alpha_n}\right]^n \end{aligned}$$

Consider now $\frac{1}{(\varphi_n)^2}\left[(\varphi_n)^{\alpha_n}\cdot(e)^{(1-\varphi_n)\alpha_n}\right]^n$. We can write:

$$\begin{aligned} \log\left(\frac{1}{(\varphi_n)^2}\left[(\varphi_n)^{\alpha_n}\cdot(e)^{(1-\varphi_n)\alpha_n}\right]^n\right) &= n\left[\begin{array}{c}\alpha_n\log\varphi_n+(1-\varphi_n)\,\alpha_n\log e\\ -2\log\varphi_n\end{array}\right] \\ &= \alpha_n n\log\varphi_n\left[\begin{array}{c}1+\frac{(1-\varphi_n)\alpha_n\log e}{\alpha_n\log\varphi_n}\\ -\frac{2}{\alpha_n n}\end{array}\right]\to\alpha_n\log\varphi_n>0 \end{aligned}$$

It follows that $\frac{1}{(\varphi_n)^2}\left[(\varphi_n)^{\alpha_n}\cdot(e)^{-\frac{\varphi_n}{n}}\right]^n > 1$ for n large, so:

$$vB(\alpha_n-1,n-1,\alpha_n\cdot\varphi_n) > v\varphi_n\sqrt{\frac{1}{2\pi\alpha_n(1-\alpha_n)n}}$$

But then the condition:

$$v\varphi_n\sqrt{\frac{1}{2\pi\alpha_n(1-\alpha_n)n}} > \varphi_n\alpha_n \Leftrightarrow v\sqrt{\frac{1}{2\pi\alpha_n(1-\alpha_n)n}} > \alpha_n \tag{25}$$

implies that for any (arbitrarily large) $L>1$, there is a n_L such that

$$vB(\alpha_n-1,n-1,\alpha_n\cdot\varphi_n)/(\varphi_n\alpha_n) > L \tag{26}$$

since $m_n \prec n^{2/3}$ implies (25). Since $vB(\alpha_n - 1, n-1, \alpha_n \cdot n^{\xi}) \to 0$, we have:

$$\begin{aligned}\frac{F\left(vB(\alpha_n - 1, n-1, \alpha_n \cdot \varphi_n)\right)}{\alpha_n \cdot \varphi_n} &= \frac{vf(0)B(\alpha_n - 1, n-1, \alpha_n \cdot \varphi_n)}{\alpha_n \varphi_n}\left(\begin{array}{c}1\\ +\frac{o(B(\alpha_n-1,n-1,\alpha_n\cdot\varphi_n)}{B(\alpha_n-1,n-1,\alpha_n\cdot\varphi_n)}\end{array}\right)\\ &> vf(0)L\left(\begin{array}{c}1\\ +\frac{o(B(\alpha_n-1,n-1,\alpha_n\cdot\varphi_n)}{B(\alpha_n-1,n-1,\alpha_n\cdot\varphi_n)}\end{array}\right) > 1\end{aligned}$$

for n sufficiently large. We conclude that for $m_n \prec n^{2/3}$, we have $F\left(vB(\alpha_n - 1, n-1, \alpha_n \cdot \varphi_n)\right)/(\alpha_n \cdot \varphi_n) > 1$ for n large, which implies that for all n sufficiently large we have $c_n^U > \varphi_n \cdot \alpha_n$.

Step 1.3. We now prove that $c_n^U > \alpha_n \cdot \varphi_n$ for all n sufficiently large is sufficient to guarantee that the probability of success converges to 1. First, note that $c_n^U > \alpha_n \cdot \varphi_n$ implies $c_n^U - \alpha_n > \alpha_n(\varphi_n - 1)$. Second, note that the probability of failing is equal to the probability that fewer than $\alpha_n n$ agents volunteer and thus can be written as:

$$\begin{aligned}&\Pr(k \le \alpha_n n) = \Pr(\frac{k}{n} \le \alpha_n) = \Pr(\frac{k}{n} \le c_n - (c_n - a_n))\\ &\quad\le \Pr\left[\left|\frac{k}{n} - c_n^U\right| \ge (c_n - a_n))\right] \le \Pr\left[\left|\frac{k}{n} - c_n^U\right| \ge \alpha_n(\varphi_n - 1)\right]\\ &\quad\le \Pr\left[\left|\frac{k}{n} - c_n^U\right| \ge c_n \cdot \frac{\alpha_n}{c_n}(\varphi_n - 1)\right] = \Pr\left[\left|\frac{k}{n} - c_n^U\right| \ge \frac{\sqrt{c_n(1-c_n)}}{\sqrt{n}}\frac{\sqrt{nc_n}\frac{\alpha_n}{c_n}(\varphi_n - 1)}{\sqrt{(1-c_n)}}\right]\\ &\quad= \Pr\left[\left|\frac{k}{n} - c_n^U\right| \ge \sigma_{c_n^U}(\frac{k}{n}) \cdot \frac{\sqrt{nc_n}\frac{\alpha_n}{c_n}(\varphi_n - 1)}{\sqrt{(1-c_n)}}\right] \le \left(\frac{\sqrt{(1-c_n)}}{\sqrt{nc_n}\frac{\alpha_n}{c_n}(\varphi_n - 1)}\right)^2 \to 0\end{aligned}$$

where in the 2nd line we used $(c_n - a_n)) \ge \alpha_n(\varphi_n - 1)$; in the 3rd and 4th lines we used Chebyshev's inequality and the fact that $\sigma_{c_n^U}(\frac{k}{n}) = \frac{\sqrt{c_n(1-c_n)}}{\sqrt{n}}$; in the last step of the 4th line ("$\to$"), we use the fact that $\varphi_n \to \infty$.

Step 2. We now show that if $m_n \succ n^{2/3}$, then $\lim_{n\to\infty} P_n^U = 0$. We first establish that $\lim_{n\to\infty} \frac{p_n^U}{\alpha_n} = 0$. Assume not, so $\lim_{n\to\infty} \frac{p_n^U}{\alpha_n} = l$ for some $l > 0$. From the equilibrium condition we must have:

$$p_n^U = F\left(vB(\alpha_n n - 1, n-1, p_n^U)\right)$$

Note however that:

$$\begin{aligned}B(\alpha_n n - 1, n-1, p_n) &= \binom{n-1}{\alpha_n n - 1}\frac{\left[(p_n)^{\alpha_n}(1-p_n)^{1-\alpha_n}\right]^n}{p_n}\\ &= \frac{\alpha_n}{p_n}\binom{n}{\alpha_n n}\left[(p_n)^{\alpha_n}(1-p_n)^{1-\alpha_n}\right]^n = l \cdot \sqrt{\frac{1}{2\pi\alpha_n(1-\alpha_n)n}}\left[\xi_n\right]^n\end{aligned}$$

with $\xi_n = \frac{(p_n)^{\alpha_n}(1-p_n)^{1-\alpha_n}}{(\alpha_n)^{\alpha_n}(1-\alpha_n)^{1-\alpha_n}} < 1$ for any n (since $p_n < \alpha_n$ implies $(p_n)^{\alpha_n}(1-p_n)^{1-\alpha_n} <$

$(\alpha_n)^{\alpha_n}(1-\alpha_n)^{1-\alpha_n})$. So, we have, for large n:

$$1 = \frac{F\left(vB(\alpha_n n-1, n-1, p_n)\right)}{p_n} \simeq \frac{vf(0)\left(\xi_n\right)^n}{\alpha_n}\sqrt{\frac{1}{2\pi\alpha_n(1-\alpha_n)n}} \to 0$$

where in the second step ("$\simeq$") we used the fact that $p_n = l\alpha_n$, and in the last step ("$\to$"), we use the fact that $\frac{1}{\alpha_n}\sqrt{\frac{1}{2\pi\alpha_n(1-\alpha_n)n}} \to 0$ since $m_n \succ n^{2/3}$ and $(\xi_n)^n \leq 1$. This is a contradiction, implying that $\lim_{n\to\infty}\frac{p_n^U}{\alpha_n} = 0$. We next use this fact to show that $p_n^U = 0$ for large n. By definition we have

$$p_n = F(vB(\alpha_n n-1, n-1, p_n)) = \Psi(\alpha_n, n, p_n) \tag{27}$$

Note that if we don't have an equilibrium $p_n > 0$ on the right of α_n, then we must have an equilibrium $\widetilde{p}_n > 0$ on the left of α_n with $\Psi'(\alpha_n, n, p_n) < 1$, where $\Psi'(\alpha_n, n, p_n)$ denotes the derivative of $\Psi(\alpha_n, n, p_n)$ with respect to p_n for a given α_n and n. Note that for any constant $\epsilon > 0$ arbitrarily small, there is a n_ϵ such that $\Psi'(\alpha_n, n, p_n) > vf(0)(1-\epsilon)B'(\alpha_n n-1, n-1, p_n)$, where $B'(\alpha_n n-1, n-1, c_n)$ denotes the derivative of $B'(\alpha_n n-1, n-1, p_n)$ with respect to p_n for a given α_n and n. The derivative of the right hand side of (27) is:

$$\begin{aligned}
B'(\alpha_n n-1, n-1, p_n) &= B(\alpha_n n-1, n-1, p_n)\left[\frac{\alpha_n n-1}{p_n} - \frac{n-\alpha_n n}{1-p_n}\right] \\
&= \frac{p_n}{v}\left[\frac{\alpha_n n-1}{p_n} - \frac{n-\alpha_n n}{1-p_n}\right] = \frac{(\alpha_n - p_n)\,n - 1 + p_n}{v\,(1-p_n)} \\
&= \frac{\left(1-\frac{p_n}{\alpha_n}\right)\alpha_n - \frac{1-p_n}{n}}{v\,(1-p_n)}n = \frac{\left[\left(1-\frac{p_n}{\alpha_n}\right) - \frac{1-p_n}{\alpha_n n}\right]}{v\,(1-p_n)}\alpha_n n \to \frac{\alpha_n n}{v} \to \infty
\end{aligned}$$

where the equilibrium condition (27) is used in the second equality; and the last line follows from the earlier result that $\frac{p_n}{\alpha_n} \to 0$ when $m_n \succ n^{2/3}$ and $\alpha_n n \to \infty$. This leads to a contradiction, since it implies that for n large at any positive intersection $\Psi'(\alpha_n, n, p_n)$ is arbitrarily large. We conclude that the only equilibrium is $p_n^U = 0$. ■

9.2 Proof of Theorem 2

For a given n, consider the relaxed problem:

$$\max_{p(\cdot),a(\cdot)}\left\{vp(0) - E\left[a(c)\cdot\frac{1-F(c)}{f(c)}\right]\right\} \tag{28}$$
$$\text{s.t. } a(c) \text{ is non-increasing with } a(c)\in[0,1]$$
$$\text{and } p, a \text{ feasible}$$

derived from (9) by eliminating the (IR) constraint, and let $\mu_n(\mathbf{c})$ with associated reduced form mechanism $a_n(c)$, $p_n(c)$ be its solution. We proceed in three steps. In Step 1, starting from $\mu_n(\mathbf{c})$ we present a perturbed mechanism $\mu_n^\gamma(\mathbf{c})$ and show it is incentive compatible and feasible. In Step 2 we show that such a perturbation strictly improves the relaxed problem

(28) if $a_n(c)$ is strictly decreasing. In Step 3, we show that the solution of the relaxed problem is $p_n(c) = 1$, $a(c) = \alpha_n$. Moreover, when $m_n \preceq n$ and $v > \alpha$, or when $m_n \prec n$ and n is large, then this solution is a solution of the full problem (9).

Step 1. Since the argument is true for any n, we omit here the subscript n for simplicity. Let μ be any feasible and incentive compatible mechanism. Consider the following "flattening" perturbation of the mechanism. After a profile of reports $\mathbf{c}$, the perturbed mechanism is defined by a new activity function that uses $\mu(\mathbf{c})$ with probability $1-\gamma$, and $\mu(\widetilde{\mathbf{c}})$ and with probability γ, where $\widetilde{c}$ is a vector in which all components $c_i > 0$ are replaced with i.i.d realizations in $(0,1]$ from $F(x)$ (and components $c_i = 0$ are left unchanged). Let $\overline{a} = \int_0^1 a(x)f(x)dx$ and $\overline{p} = \int_0^1 p(x)f(x)dx$ This new allocation generates a reduced form mechanism:

$$p^\gamma(c_i) = \gamma\overline{p} + (1-\gamma)p(c_i), \text{ and } a^\gamma(c_i) = \gamma\overline{a} + (1-\gamma)a(c_i)$$

for $c_i \in (0,1]$ and $p^\gamma(0) = p(0)$, $a^\gamma(0) = a(0)$. Note that since $a(0) \geq a(c_i)$ and $p(0) \geq p(c_i)$ for all $c_i \in [0,1]$, we must have that $a(0) \geq \overline{a} = \int_0^1 a(x)f(x)dx$ and similarly $p(0) \geq \overline{p}$.

The new reduced form allocation is clearly feasible since we have shown the feasible activity function that generates it. It also does not change $p(0)$. Note that after the change incentive compatibility is satisfied since $a^\gamma(c_i)$ is non increasing in $[0,1]$ and, after the change, we have:

$$\begin{aligned} U^\gamma(x) &= \gamma\left[v\overline{p} - \overline{a}x\right] + (1-\gamma)\left[vp(x) - a(x)x\right] \\ &= vp^\gamma(x) - a^\gamma(x)x \end{aligned}$$

For $c > 0$ and $c' \geq 0$, we have:

$$\begin{aligned} vp^\gamma(c) - a^\gamma(c)c &= \gamma\left[v\overline{p} - \overline{a}c\right] + (1-\gamma)\left[vp(c) - a(c)c\right] \\ &\geq \gamma\left[v\overline{p} - \overline{a}c\right] + (1-\gamma)\left[vp(c') - a(c')c\right] \\ &= vp^\gamma(c') - a^\gamma(c')c \end{aligned}$$

Moreover a type 0 does not want to imitate a type $c > 0$:

$$vp^\gamma(0) \geq \gamma\left[v\overline{p}\right] + (1-\gamma)\left[vp(0)\right] \geq \gamma\left[v\overline{p}\right] + (1-\gamma)\left[vp(c')\right] = vp^\gamma(c')$$

where the first inequality follows from the fact that $p(0) \geq \overline{p}$. Given IC, by the usual argument, we have:

$$\left[U^\gamma\right]'(x) = -\gamma\overline{a} - (1-\gamma)a'(x) = -a^\gamma(c_i)$$

So the change is feasible in the relaxed problem.

Step 2. To see that it increases the objective function, we need to show that $-\int_0^1 a^\gamma(x)(1-F(x))dx$ increases in γ, since $vp(0)$ is unchanged by the change. We can write it as:

$$\overline{a} \cdot \int_0^1 G(x)\frac{a^\gamma(x)}{\overline{a}}f(x)dx$$

where $G(x) = -\frac{1-F(x)}{f(x)}$ and $\overline{a}$ is $\int_0^1 a(x)f(x)dx$. Note that $\frac{a^\gamma(x)f(x)}{\overline{a}}$ is a density since

$\frac{a^\gamma(x)f(x)}{\overline{a}} \geq 0$ and $\int_0^1 \frac{a^\gamma(x)f(x)}{\overline{a}}dx = 1$. By MHRA, $G(x)$ is monotone non-decreasing in x, so the result is proven if we prove that an increase in γ implies a first order stochastic dominance improvement in $\frac{a^\gamma(x)}{\overline{a}}f(x)$. Define: $\Gamma^\gamma(t) = \int_0^t \frac{a^\gamma(x)}{\overline{a}}f(x)dx$. We prove the result if $\partial\Gamma^\gamma(t)/\partial\gamma < 1$ for all $t < 1$. We have:

$$\begin{aligned}\frac{\partial}{\partial\gamma}\Gamma^\gamma(t) &= \frac{\partial}{\partial\gamma}\left[\frac{1}{\overline{a}}\int_0^t \left[\gamma\overline{a} + (1-\gamma)a(x)\right]f(x)dx\right] \\ &= \int_0^t \left[1 - \frac{a(x)}{\overline{a}}\right]f(x)dx = F(t)\left[1 - \frac{E\left[a(x); x \leq t\right]}{E\left[a(x)\right]}\right] \leq 0\end{aligned}$$

where the last inequality follows from the fact that $a(x)$ non-increasing in x. It follows that increasing γ improves the relaxed problem, which is maximized at $\gamma = 1$. When $\gamma = 1$, feasibility and the (IC) are satisfied: so $\gamma = 1$ is optimal for the original problem as well.

Step 3. From Step 2 we know that the optimal mechanism solving the relaxed problem is independent of c: a_n^S, p_n^S. It is easy to see that this mechanism will always activate a coalition of size m. Assume not. Three cases are possible. First, the mechanism activates a coalition of size larger than m; second the mechanism selects a non empty coalition of size smaller than m. In the first case, just modify the mechanism by imposing that all activated coalitions are reduced to a size m by randomly selecting agents to drop; in the second, modify the mechanism by imposing that coalitions that are smaller than m are not selected and a coalition of size m is selected instead, with equal probability on all coalitions of size m. This leaves p_n^S unchanged and it and reduce a_n^S. No constraint is violated and the objective function is increased, a contradiction. The third case is that the mechanism always select a coalition of size m, but with probability $p_n^S < 1$. It is easy to see that this is not optimal since by increasing p_n^S we obtain a marginal improvement in utility equal to $1 - \alpha_n\frac{c}{v}$. We conclude that the optimal solution of the relaxed problem is $p_n^S = 1$ and $a_n^S = \alpha_n$. The fact that the optimal solution of the relaxed problem satisfied the (IR) constraint follows from the argument presented in Section 3.2.

9.3 Proof of Proposition 1

We first prove that for any α, n, $Y_n(p)$ has a unique fixed point p^O. We then prove that a simple VBO is incentive compatible if and only if the type parameter is $c^O = F^{-1}(p^O)$. The following lemma will prove useful.

Lemma A3. $\frac{B(\alpha n-1+j,n-1,p)}{B(\alpha n-1,n-1,p)} = \prod_{k=1}^{j} \frac{(n-\alpha n-1-k)}{\alpha n-1+k} \cdot \left(\frac{p}{1-p}\right)^j$.

Proof: We prove this by induction. The formula is correct for $j = 1$ since $\frac{B(m+1,n-1,p)}{B(m,n-1,p)} =$

$\frac{(n-m-1)}{m+1} \cdot \frac{p}{1-p}$, and for $j = 2$ since:

$$\begin{aligned}\frac{B(m+2,n-1,p)}{B(m,n-1,p)} &= \frac{B(m+2,n-1,p)}{B(m+1,n-1,p)} \cdot \frac{B(m+1,n-1,p)}{B(m,n-1,p)} \\ &= \frac{(n-m-2)}{m+2}\frac{(n-m-1)}{m+1} \cdot \left(\frac{p}{1-p}\right)^2\end{aligned}$$

The induction hypothesis is that the formula is correct for $j-1$:

$$\frac{B(m+j-1,n-1,p)}{B(m,n-1,p)} = \prod_{i=1}^{j-1} \frac{(n-m-i)}{m+i} \cdot \left(\frac{p}{1-p}\right)^{j-1}$$

This implies:

$$\begin{aligned}\frac{B(m+j,n-1,p)}{B(m,n-1,p)} &= \frac{B(m+j,n-1,p)}{B(m+j-1,n-1,p)} \cdot \frac{B(m+j-1,n-1,p)}{B(m,n-1,p)} \\ &= \frac{(n-m-j)}{m+j}\left(\frac{p}{1-p}\right)\prod_{l=1}^{j-1}\frac{(n-m-l)}{m+l} \cdot \left(\frac{p}{1-p}\right)^{j-1} \\ &= \prod_{i=1}^{j}\frac{(n-m-i)}{m+i} \cdot \left(\frac{p}{1-p}\right)^{j}\end{aligned}$$

so the formula is correct for j, which proves the claim. ■

We now proceed in two steps.

Step 1. For any α, n, $Y_n(p)$ is defined as:

$$Y_n(p) = F\left(\frac{vB(\alpha n-1,n-1,p)}{\sum_{j=\alpha n-1}^{n-1}\frac{\alpha n}{j+1}B(j,n-1,p)}\right).$$

We can rewrite it as:

$$Y_n(p) = F\left(\frac{v}{1+\sum_{j=\alpha n}^{n-1}\frac{\alpha n}{j+1}\frac{B(j,n-1,p)}{B(\alpha n-1,n-1,p)}}\right) = F\left(\frac{v}{1+\sum_{j=1}^{n-\alpha n}\frac{\alpha n}{j+\alpha n}\frac{B(\alpha n-1+j,n-1,p)}{B(\alpha n-1,n-1,p)}}\right)$$

We now show that:

$$\sum_{j=1}^{n-\alpha n}\frac{\alpha n}{j+\alpha n}\frac{B(\alpha n-1+j,n-1,p)}{B(\alpha n-1,n-1,p)}$$

is strictly increasing in p, so $Y_n(p)$ is continuous and strictly decreasing in p. By Lemma A3, we have:

$$\frac{B(\alpha n-1+j,n-1,p)}{B(\alpha n-1,n-1,p)} = \prod_{i=1}^{j}\frac{n-\alpha n+1-i}{\alpha n-1+i} \cdot \left(\frac{p}{1-p}\right)^{j}$$

It follows that:

$$\sum_{j=1}^{n-\alpha n} \frac{\alpha n}{j+\alpha n} \frac{B(\alpha n-1+j, n-1, p)}{B(\alpha n-1, n-1, p)} = \sum_{j=1}^{n-\alpha n} \frac{\alpha n}{j+\alpha n} \cdot \prod_{i=1}^{j} \frac{n-\alpha n+1-i}{\alpha n-1+i} \cdot \left(\frac{p}{1-p}\right)^{j}$$

which is increasing in p. Moreover, it is easy to see that $Y_n(0) = F(v) > 0$ and $Y_n(1) = 0$. Hence $Y_n(p)$ has a unique fixed point in $(0, 1)$.

Step 2. Incentive compatibility requires that $U(c) \geq vp(c') - ca(c')$ for all $c, c' \in [0, 1]$. It is straightforward to show that this is implied by $vp(c^O) - c^O a(c^O) = vp_2$. The unique fixed point of Y, denoted p^O, defines the volunteer threshold, c^O. Given this threshold, we let p_1 denote the (constant) interim probability of success for all types $c \leq c^O$, let p_2 denote the (constant) interim probability of success for all types $c > c^O$, and let a denote the (constant) interim probability of being activated for all types $c \leq c^O$. From the definition of a VBO, p_1, p_2, a are given by the following formulas:

$$a = \sum_{k=\alpha n-1}^{n-1} \frac{\alpha n}{k+1} B(k, n-1, F\left(c^O\right)) \tag{29}$$

$$p_1 = \sum_{k=\alpha n-1}^{n-1} B(k, n-1, F\left(c^O\right)) \tag{30}$$

$$p_2 = \sum_{k=\alpha n}^{n-1} B(k, n-1, F\left(c^O\right)) \tag{31}$$

and $p^O = F(c^O)$.

The equilibrium condition for IC is:

$$ac^O = v(p_1 - p_2) \tag{32}$$

By substituting equations (29), (30), and (31) into equation (32), for any n we can obtain an expression for $c_n^O(v)$, the VBO volunteer threshold cost as a function of group size and the value of success.

$$F^{-1}(p_n^O) = c_n^O(v) = v \frac{B(\alpha n-1, n-1, p_n^O)}{\sum_{k=\alpha n-1}^{n-1} \frac{\alpha n}{k+1} B(k, n-1, p_n^O)} \tag{33}$$

where the numerator on the right hand side is $p_{1n} - p_{2n}$ and the denominator is a_n, the probability a volunteer is activated. It is easy to see that (33) implies the statement in the Proposition. ■

9.4 Proposition 2

First observe that all group members whose recommended action is to free ride will obey the recommendation because they assume all other members are obedient and know that either 0 or exactly m of the other members have been requested to participate. Hence, their

participation will not affect success of failure, so they are better off free riding. Second, all members whose recommended action is to activate will obey the recommendation because they assume all other agents are obedient and know that exactly $m-1$ of the other agents have been recommended to activate. Hence their payoff is 0 if they disobey the request to activate and $v-c_i>0$ if they obey the request, since $c_i \leq c^O \in (0,v)$. It follows that, using the notation of Section 2.2.1, the following ex post moral hazard condition is verified:

$$u_g^i(c) \geq u_{\xi_i(g)}^i(c) \tag{ENA}$$

for any g such that $\mu_g(c_i, c_{-i}) > 0$ and for any function $\xi_i(g)$ mapping g to either $\{g, (g_{-i}, 0)\}$ if $g \in I_i$ or $\{g, (g_{-i}, 1)\}$ if $g \notin I_i$. Note that $\xi_i(g)$ is measurable with respect to the vector g, while $\delta_i(g_i)$ ignores g_{-i} and is measurable only with respect to g_i. We say that a mechanism is honest and ex post obedient (HEO) if it satisfies (IC) and (ENA).

It is now easy to see that (IC) and (ENA) imply (HO). To see this note that:

$$\begin{aligned} E_{c_{-i}}\left[\sum_{g\in I}\mu_g(c_i,c_{-i})\,u_g^i(c)\right] &\geq E_{c_{-i}}\left[\sum_{g\in I}\mu_g(\widetilde{c}_i,c_{-i})\,u_g^i(c)\right] \\ &\geq E_{c_{-i}}\left[\sum_{g\in I}\mu_g(\widetilde{c}_i,c_{-i})\,u_{\xi_i(g)}^i(c)\right] \end{aligned} \tag{34}$$

for any $i=1,...,n$, $c_i, \widetilde{c}_i \in [0,1]$ and any function $\xi_i(g)$ mapping g to either $\{g, g\backslash\{i\}\}$ if $g \in I_i$ or $\{g, g\cup\{i\}\}$ if $g \notin I_i$. ■

9.5 Proof of Proposition 3

The probabilities of success are: $P_n^O = \sum_{k=\alpha n}^n B(k,n,p_n^O)$ and $P_n^U = \sum_{k=\alpha n}^n B(k,n,p_n^U)$,which are respectively increasing in p_n^O and p_n^U. Hence we need to show that $p_n^O > p_n^U$ for any n. Recall that:

$$p_n^U = F\left(vB(\alpha n-1, n-1, p_n^U)\right), \text{ and } p_n^O = Y_n(p_n^O)$$

where $Y_n(\cdot)$ is defined by (10). Since p_n^O solves equation (33) and, as proven in Step 1 of the proof of Proposition 2, $Y_n(p)$ is strictly decreasing in p, we must have that if $p \in \left(p_n^O, v\right]$, then $p > Y_n(p)$. It follows that if $p_n^U > p_n^O$, we must have that:

$$p_n^U > F\left(v\frac{B(\alpha n-1,n-1,p_n^U)}{\sum_{j=\alpha n-1}^{n-1}\frac{\alpha n}{j+1}B(j,n-1,p_n^U)}\right) > F\left(vB(m-1,n-1,p_n^U)\right) = p_n^U$$

a contradiction. Therefore $P_n^O/P_n^U > 1$. ■

9.6 Proof of Proposition 4

We proceed in two steps.

Step 1. To prove the first bullet point, let $m_n = \alpha n$ for some $\alpha \in (0,1)$. From (33) in

Proposition 2, p_n^O is the unique solution to:

$$p_n^O = F\left[v\frac{1}{1+\sum_{j=1}^{n-\alpha n}\frac{\alpha n}{j+\alpha n}\frac{B(\alpha n-1+j,n-1,p_n^O)}{B(\alpha n-1,n-1,p_n^O)}}\right] \tag{35}$$

and from Lemma A3, we have:

$$\begin{aligned}\sum_{j=1}^{n-\alpha n}\frac{\alpha n}{j+\alpha n}\frac{B(\alpha n-1+j,n-1,p_n^O)}{B(\alpha n-1,n-1,p_n^O)} &= \sum_{j=1}^{n-\alpha n}\frac{\alpha n}{j+\alpha n}\cdot\prod_{i=1}^{j}\frac{n-\alpha n+1-i}{\alpha n-1+i}\left(\frac{p_n^O}{1-p_n^O}\right)^j \qquad (36)\\ &\leq \sum_{j=1}^{n-\alpha n}\frac{\alpha n}{j+\alpha n}\left(\frac{n-\alpha n}{\alpha n}\frac{p_n^O}{1-p_n^O}\right)^j\\ &\leq \sum_{j=1}^{n-\alpha n}\left(\frac{1-\alpha}{\alpha}\frac{p_n^O}{1-p_n^O}\right)^j\end{aligned}$$

Assume by contradiction that $\lim_{n\to\infty}p_n^O=0$, then:

$$\lim_{n\to\infty}\sum_{j=1}^{n-\alpha n}\left(\frac{1-\alpha}{\alpha}\frac{p_n^O}{1-p_n^O}\right)^j \leq \lim_{n\to\infty}\sum_{j=1}^{n}\left(\frac{1-\alpha}{\alpha}\frac{p_n^O}{1-p_n^O}\right)^j < \infty$$

since $\frac{1-\alpha}{\alpha}\frac{p_n^O}{1-p_n^O}\to 0$. Hence

$$\lim_{n\to\infty}\sum_{j=1}^{n-\alpha n}\frac{\alpha n}{j+\alpha n}\prod_{i=1}^{j}\frac{n-\alpha n+1-i}{\alpha n-1+i}\left(\frac{p_n^O}{1-p_n^O}\right)^j < \infty \Rightarrow \lim_{n\to\infty}p_n^O > 0$$

a contradiction.

Step 2. To prove the second bullet point, let $m_n \prec n$. From Proposition 3, p_n^O is again the unique positive solution to (35) and from Lemma A3, we have (36). Assume by contradiction that $\lim_{n\to\infty}\frac{p_n^O}{\alpha_n}<1$, so $\lim_{n\to\infty}\frac{1-\alpha_n}{\alpha_n}\frac{p_n^O}{1-p_n^O}=\theta<1$. In this case:

$$\lim_{n\to\infty}\sum_{j=1}^{n-\alpha_n n}\left(\frac{1-\alpha_n}{\alpha_n}\frac{p_n^O}{1-p_n^O}\right)^j \to \sum_{j=1}^{\infty}(\theta)^j = \frac{\theta}{1-\theta} < \infty$$

Hence

$$\lim_{n\to\infty}\sum_{j=1}^{n-\alpha_n n}\frac{\alpha_n n}{j+\alpha_n n}\prod_{i=1}^{j}\frac{n-\alpha_n n+1-i}{\alpha_n n-1+i}\left(\frac{p_n^O}{1-p_n^O}\right)^j < \infty$$

$$\Rightarrow \lim_{n\to\infty}p_n^O > 0 = \lim_{n\to\infty}\alpha_n \Rightarrow \lim_{n\to\infty}\frac{p_n^O}{\alpha_n} = \infty > 1$$

where $\lim_{n\to\infty}\alpha_n = 0$ follows from $m_n \prec n$. We have a contradiction and conclude that $\lim_{n\to\infty}\frac{p_n^O}{\alpha_n}\geq 1$. ■

9.7 Proof of Proposition 5

Consider the first bullet point of the proposition. From equations (30 and 31) we have:

$$p_1(c_n^O) - p_2(c_n^O) = B(\alpha n - 1, n - 1, F\left(c_n^O\right)) \to_n 0$$

since all single terms of the binomial expansion converge to 0 as $n \to \infty$. Hence, equation 33 implies that $\lim_{n\to\infty} c_n^O a(c_n^O) = 0$, which in turn implies $\lim_{n\to\infty} a(c_n^O) = 0$ since $\lim_{n\to\infty} c_n^O = c^O > 0$. This implies $\lim_{n\to\infty} P_n^O = 0$. To show $\lim_{n\to\infty} c_n^O \leq \alpha$, suppose instead that $\lim_{n\to\infty} c_n^O > \alpha$. In that case, the fraction of volunteers converges in probability to a fraction strictly greater than α, implying $\lim_{n\to\infty} P_n^O = 1$, a contradiction. Consider now the second bullet point. Theorem 1 implies that $P_n^U = 0$ for all $n > \overline{n}_U(\alpha, v)$, and Proposition 5 implies that $c_n^O > 0$ for all n, so $P_n^O > 0$ for all n. Hence, $P_n^U / P_n^O = 0$ for all $n > \overline{n}_U(\alpha, v)$. ■

9.8 Proof of Proposition 6

We proceed in two steps.

Step 1. The first part of Proposition 6 follows from Proposition 3 and part (1) of Theorem 1: for any n, the probability of success of an organized group using a VBO mechanism is greater than or equal to the probability of success of an unorganized group. Since the probability of success of an unorganized group converges to one when $m_n \prec n^{2/3}$, the same must be true for an organized group using a VBO mechanism.

Step 2. Suppose that $m_n \succ n^{2/3}$ and $\lim_{n\to\infty} P_n^O > 0$. For n large enough, it must be that:

$$\lim_{n\to\infty} \frac{1}{\alpha_n} \cdot F\left[\frac{vB(\alpha_n n - 1, n - 1, \alpha_n)}{\sum_{k=\alpha n-1}^{n-1} \frac{\alpha_n n}{1+k} B\left(k, n-1, \alpha_n\right)}\right] \geq 1 \tag{37}$$

If this inequality is not satisfied, then $\lim_{n\to\infty} \alpha_n > \lim_{n\to\infty} c_n$ and so $P_n^0 \to 0$, for the same reason as in the proof of Theorem 1.

We first prove that there is a constant $a_\infty > 0$ such that $\sum_{k=\alpha n-1}^{n-1} \frac{\alpha_n n}{1+k} B\left(k, n-1, \alpha_n\right) \to a_\infty$ as $n \to \infty$. To this goal, note that, for any $\eta > 0$ arbitrarily small:[38]

$$\begin{aligned}
\sum\nolimits_{k=\alpha_n n+\eta\sqrt{n}}^{n-1} B\left(k, n-1, \alpha_n\right) &= \Pr\left(k \geq \alpha_n n + \eta\sqrt{n}\right) = \Pr\left(\frac{k}{n} - \alpha_n \geq \frac{\eta}{\sqrt{n}}\right) \\
&\leq \Pr\left(\left|\frac{k}{n} - \alpha_n\right| \geq \sigma_{\alpha_n}(k/n) \frac{\eta}{\sqrt{\alpha_n(1-\alpha_n)}}\right) \\
&\leq \lim \frac{\alpha_n(1-\alpha_n)}{\eta^2} \simeq \frac{\delta}{\eta^2} \cdot \frac{m_n}{n} \to 0
\end{aligned}$$

where $\sigma_{\alpha_n}(k/n) = \sqrt{\frac{\alpha_n(1-\alpha_n)}{n}}$ and in the last step we use Chebyshev's inequality. So we

[38] The initial term of the summation below should be written as $\lceil \alpha_n n + \eta\sqrt{n} \rceil$, since $\alpha_n n + \eta\sqrt{n}$ may not be an integer. To keep the notation simple, and without loss of generality since irrelevant for the argument, in the following we ignore this issue.

have:

$$\begin{aligned}\lim_{n\to\infty}\sum\nolimits_{k=\alpha n-1}^{n-1}\frac{\alpha_n n}{1+k}B\left(k,n-1,\alpha_n\right) &= \lim_{n\to\infty}\left[\begin{array}{c}\sum_{k=\alpha n-1}^{\left(\alpha+\frac{\eta}{\sqrt{n}}\right)n-1}\frac{\alpha_n n}{1+k}B\left(k,n-1,\alpha_n\right)\\ +\sum_{k=\alpha_n n+\eta\sqrt{n}}^{n-1}\frac{\alpha_n n}{1+k}B\left(k,n-1,\alpha_n\right)\end{array}\right]\\ &= \lim_{n\to\infty}\sum\nolimits_{k=\alpha n-1}^{\left(\alpha+\frac{\eta}{\sqrt{n}}\right)n-1}\frac{\alpha_n n}{1+k}B\left(k,n-1,\alpha_n\right)\end{aligned}$$

and $\lim_{n\to\infty}\sum_{k=\alpha n-1}^{\left(\alpha+\frac{\eta}{\sqrt{n}}\right)n-1}B\left(k,n-1,\alpha_n\right)\geq\frac{1}{2}$. We can also write:

$$\begin{aligned}&\lim_{n\to\infty}\sum\nolimits_{k=\alpha n-1}^{\left(\alpha+\frac{\eta}{\sqrt{n}}\right)n-1}\frac{\alpha_n n}{1+k}B\left(k,n-1,\alpha_n\right)\\ \geq\ &\left(\lim_{n\to\infty}\frac{1}{1+\eta\cdot\frac{n^{1/2}}{m_N}}\right)\cdot\lim_{n\to\infty}\sum\nolimits_{k=\alpha n-1}^{\left(\alpha+\frac{\eta}{\sqrt{n}}\right)n-1}B\left(k,n-1,\alpha_n\right)\geq\frac{1}{2}\end{aligned}$$

where in the last line we use the fact that $\alpha_n n=m_n\succ n^{1/2}$, so $n^{1/2}/m_n\to 0$.

Given that $\lim_{n\to\infty}\sum_{k=\alpha n-1}^{n-1}\frac{\alpha_n n}{1+k}B\left(k,n-1,\alpha_n\right)=a_\infty>0$, we can rewrite (37) as:

$$\lim_{n\to\infty}\left[\frac{1}{\alpha_n}F\left(\frac{vB(\alpha n-1,n-1,\alpha_n)}{\sum_{k=\alpha n-1}^{n-1}\frac{\alpha_n n}{1+k}B\left(k,n-1,\alpha_n\right)}\right)\right]=\frac{vf(0)}{\alpha_\infty}\lim_{n\to\infty}\frac{B(\alpha n-1,n-1,\alpha_n)}{\alpha_n}$$

which again converges to 0 if $m_n\succ n^{2/3}$ and diverges at infinity if $m_n\prec n^{2/3}$. ■

9.9 Proof of Lemma 1

We will omit n here as a subscript for simplicity whenever it does not create confusion. Let μ^b be an optimal binary mechanism and let c^b be the volunteer cut-point associated with μ^b and q_k be the corresponding probability of success when there are k volunteers. We will prove the result that if $q_k>0$ for some $k\geq m$, then $q_{k+j}=1$ for $j>0$. This implies that there is a k^b such that $q_j=0$ for $j<k^b$ and $q_j=1$ for $j>k^b$ and at most at one k we have $q_{k^b}\in(0,1)$. So the optimal binary mechanism is a k^b-VBO except at most for an event with probability that converges to zero as $n\to 0$, i.e. when there are exactly k^b volunteers.

We proceed in two steps. In step 1, we establish that the optimal binary mechanism is non-wasteful, meaning that it does not ever activate more agents than necessary; step 2 shows that it is characterized by a threshold k^b.

Step 1. We first show that the optimal honest and obedient binary mechanism must be non-wasteful in the sense that whenever a group is activated, there are exactly m members in the activated group. We prove this by contradiction by supposing that μ^b is wasteful at a positive measure set of profiles and then showing that it can be improved. First, define a new mechanism, μ', that is exactly the same as μ for all coalitions of size m, but eliminates all waste by reducing all activated successful coalitions to a size m by randomly selecting agents to drop out, and by not activating unsuccessful coalitions that are smaller than m. This leaves c^b, p_1^μ and p_2^μ unchanged and reduces a^μ to $a^{\mu'}<a^\mu$. This implies

that $a^{\mu\prime}c^b < v(p_1^\mu - p_2^\mu)$, so some cost types in a neighborhood above c^b are strictly better off volunteering, which violates incentive compatibility. Now, for any $\widetilde{c} > c^b$, consider a modified version of μ', denoted by $\widetilde{\mu}'$ that has the same success probabilities, $\{q_k\}_{k=m}^n$ as μ' except that all members with $c < \widetilde{c}$ are volunteers, so there is a bigger pool of volunteers. This increases p_1^μ and p_2^μ to $p_1^{\widetilde{\mu}'} > p_1^\mu$ and $p_2^{\widetilde{\mu}'} > p_2^\mu$ and changes $a^{\mu\prime}$ to $\widetilde{a}^{\mu\prime}$. Denote by $\widetilde{c}^b > c^b$ the first such value of $\widetilde{c} > c^b$ such that $\widetilde{a}^{\widetilde{\mu}\prime}\widetilde{c}^b = v(p_1^{\widetilde{\mu}'} - p_2^{\widetilde{\mu}'})$. (Such a point exists by the intermediate value theorem.) Denote by $u(c; c^b)$ the interim expected utility of a member with cost c under μ, and denote by $u(c; \widetilde{c}^b)$ the interim expected utility of a member with cost c under the modified mechanism, $\widetilde{\mu}'$ with volunteer cutoff $\widetilde{c}^b > c^b$. Because $p_2^{\widetilde{\mu}'} > p_2^\mu$, we know that $u(c; \widetilde{c}^b) = u(\widetilde{c}^b; \widetilde{c}^b) > u(c; c^b)$ for all $c \geq \widetilde{c}^b$, so these members are strictly better off. For $c \in (c^b, \widetilde{c}^b)$ we have $u(c; \widetilde{c}^b) \geq u(\widetilde{c}^b; \widetilde{c}^b) > u(c; c^b)$, so these members are also better off. Finally, for all $c \in [0, c^b)$ (the volunteers under μ) are better off because for each $k \geq m$ for which $q_k > 0$ there are a positive measure of additional profiles $\mathbf{c}$ with exactly k volunteers, and for each such additional profile, the c-type member in $[0, \widetilde{c}^b)$ gets a conditional expected utility of $(v - \frac{m}{k}c)q_k > 0$. (Such members receive the same conditional expected utility for all other profiles.) Hence, $u(c; \widetilde{c}^b) > u(c; c^b)$ for all $c \leq c^b$. Hence all agents are better off under $\widetilde{\mu}'$ than under μ. All constraints are satisfied and the objective function is increased, a contradiction. Hence the optimal mechanism is non-wasteful

Step 2. If $q_k > 0$ for $k \geq m$ and $q_{k+j} < 1$ for $j > 1$, then there must be a k' such that $q_{k'} > 0$ for $k' \geq m$ and $q_{k'+1} < 1$, so we only need to prove the result for the case of $j = 1$.

Assume by contradiction that $q_k > 0$ for some $k \geq m$ and $q_{k+1} < 1$. Let c^b be the minimum cost above which an agent is activated with probability zero. Then incentive compatibility is binding at c^b if $ac^b = v(p_1^b - p_2^b)$, where:

$$p_1^b - p_2^b = B(n-1, n-1, F(c^b))q_n + \sum\nolimits_{k=m}^{n-1} [B(k-1, n-1, F(c^b)) - B(k, n-1, F(c^b))]q_k$$

and

$$a = \sum\nolimits_{k=m-1}^{n-1} \frac{m}{1+k} B(k, n-1, F(c^b))q_{k+1}.$$

We can marginally reduce q_k by $-dq_k < 0$ and marginally increase q_{k+1} by $dq_{k+1} > 0$ so that the (IC) constraint is unchanged, thus keeping c^b constant. This requires:

$$\begin{aligned} & c^b \left[-\frac{m}{k} B(k-1, n-1, F(c^b)) + \frac{m}{k+1} B(k, n-1, F(c^b)) \frac{dq_{k+1}}{dq_k} \right] dq_k \\ & = \left[\begin{array}{c} -\big(B(k-1, n-1, F(c^b)) - B(k, n-1, F(c^b))\big) \\ +B(k, n-1, F(c^b)) - B(k+1, n-1, F(c^b)) \frac{dq_{k+1}}{dq_k} \end{array} \right] dq_k \end{aligned} \tag{38}$$

Note we can write:

$$\begin{aligned} p_1^b - p_2^b &= B(n-1, n-1, F(c^b))q_n \\ &\quad + \sum\nolimits_{k=m}^{n-1} [B(k-1, n-1, F(c^b)) - B(k, n-1, F(c^b))]q_k = \sum\nolimits_{k=m}^{n} \Theta_k q_n \end{aligned} \tag{39}$$

where we denote:

$$\Theta_n = B(n-1, n-1, F(c^b))$$
$$\Theta_k = B(k-1, n-1, F(c^b)) - B(k, n-1, F(c^b)) \text{ for } k = n-1, .., m$$

We can rewrite the previous expression as:

$$\Theta_k = B(k-1, n-1, F(c^b)) - B(k, n-1, F(c^b)) = B(k, n-1, F(c^b)) \left[\frac{k}{n-k} \frac{1-F(c^b)}{F(c^b)} - 1 \right]$$

Similarly, we have:

$$\begin{aligned}
\Theta_{k+1} &= B(k, n-1, F(c^b)) - B(k+1, n-1, F(c^b)) \\
&= \frac{(n-1)!}{(k)!\,(n-k-1)!} \left(F(c^b)\right)^k (1-F(c^b))^{n-k-1} \\
&- \frac{(n-1)!}{(k+1)!\,(n-k-2)!} \left(F(c^b)\right)^{k+1} (1-F(c^b))^{n-k-2} \\
&= B(k+1, n-1, F(c^b)) \left[\frac{k+1}{n-k-1} \frac{1-F(c^b)}{F(c^b)} - 1 \right]
\end{aligned}$$

Substituting into (38) gives:

$$\begin{aligned}
& c^b \left[-\frac{m}{k} B(k-1, n-1, F(c^b)) + \frac{m}{k+1} B(k, n-1, F(c^b)) \frac{dq_{k+1}}{dq_k} \right] dq_k \\
&= \left[\begin{array}{c} -\left(B(k-1, n-1, F(c^b)) - B(k, n-1, F(c^b))\right) \\ +B(k, n-1, F(c^b)) - B(k+1, n-1, F(c^b)) \frac{dq_{k+1}}{dq_k} \end{array} \right] dq_k
\end{aligned}$$

or

$$\begin{aligned}
& c^b \left[\begin{array}{c} -\frac{m}{k} \frac{k}{n-k} \frac{1-F(c^b)}{F(c^b)} B(k, n-1, F(c^b)) \\ +\frac{m}{k+1} \frac{k+1}{n-k-1} \frac{1-F(c^b)}{F(c^b)} B(k+1, n-1, F(c^b)) \frac{dq_{k+1}}{dq_k} \end{array} \right] dq_k \\
= & \left[\begin{array}{c} -B(k, n-1, F(c^b)) \left[\frac{k}{n-k} \frac{1-F(c^b)}{F(c^b)} - 1 \right] \\ +B(k+1, n-1, F(c^b)) \left[\frac{k+1}{n-k-1} \frac{1-F(c^b)}{F(c^b)} - 1 \right] \frac{dq_{k+1}}{dq_k} \end{array} \right] dq_k
\end{aligned}$$

It follows that:

$$\begin{aligned}
& \frac{dq_{k+1}}{dq_k} \left[\left(c^b \frac{m}{k+1} - 1 \right) \frac{k+1}{n-k-1} \frac{1-F(c^b)}{F(c^b)} + 1 \right] B(k+1, n-1, F(c^b)) \\
&= \left[\left(c^b \frac{m}{k} - 1 \right) \frac{k}{n-k} \frac{1-F(c^b)}{F(c^b)} + 1 \right] B(k, n-1, F(c^b)) \\
& \Leftrightarrow \frac{dq_{k+j}}{dq_k} = \frac{1 - \left(1 - c^b \frac{m}{k}\right) \frac{k}{n-k} \frac{1-F(c^b)}{F(c^b)}}{1 - \left(1 - c^b \frac{m}{k+1}\right) \frac{k+1}{n-k-1} \frac{1-F(c^b)}{F(c^b)}} \frac{B(k, n-1, F(c^b))}{B(k+1, n-1, F(c^b))} = T_{n,k} \frac{B(k, n-1, F(c^b))}{B(k+1, n-1, F(c^b))}
\end{aligned}$$

where:

$$T_{n,k} = \frac{1-\left(1-c^b\frac{m}{k}\right)\frac{k}{n-k}\frac{1-F(c^b)}{F(c^b)}}{1-\left(1-c^b\frac{m}{k+1}\right)\frac{k+1}{n-k-1}\frac{1-F(c^b)}{F(c^b)}} > 1 \Leftrightarrow \frac{k-c^b m}{n-k} < \frac{k-c^b m+1}{n-k-1} \Leftrightarrow n > c^b m$$

After this change, the probability of success increases, indeed we have:

$$\begin{aligned} dP_n &= \left[-B(k,n-1,F(c^b)) + B(k+1,n-1,F(c^b))\frac{dq_{k+1}}{dq_k}\right] dq_k \\ &= \left[-B(k,n-1,F(c^b)) + B(k+1,n-1,F(c^b)) \cdot T_{n,k}\frac{B(k,n-1,F(c^b))}{B(k+1,n-1,F(c^b))}\right] dq_k \\ &> (T_{n,k}-1)\, B(k,n-1,F(c^b)) \cdot dq_k > 0 \end{aligned}$$

Since the probability of success increases but the average probability of participation remains constant (since c^b is unchanged), we must increase welfare, ceteris paribus. This implies that the original mechanism was not optimal, a contradiction. ■

9.10 Proof of Lemma 2

We show that for any ε, there is a n_ε such that for $n > n_\varepsilon$ we have:

$$p_n^*(0) - p_n^*(c_n^*) < \varepsilon$$

where $c_n^* = \sup\{c \in [0,1] \,|\, a_n^*(c) > 0\}$.

There are two cases to consider: (1) $\lim_{n\to\infty} c_n^* = 0$; (2) $\lim_{n\to\infty} c_n^* = c_\infty^* \in (0,v)$.

Case 1: $\lim_{n\to\infty} c_n^* = 0$

By IC, we have:

$$\begin{aligned} vp_n^*(c_n^*) - a_n^*(c_n^*)c_n^* &\geq vp_n^*(0) - a_n^*(0)c_n^* \\ &\Leftrightarrow v\left[p_n^*(0) - p_n^*(c_n^*)\right] \leq \left[a_n^*(0) - a_n^*(c_n^*)\right] c_n^* \to 0 \end{aligned}$$

It follows that for any $\varepsilon > 0$, there is a n_ε such that $p_n^*(0) - p_n^*(c_n^*) < \varepsilon$ for $n > n_\varepsilon$.

Case 2: $\lim_{n\to\infty} c_n^* = c_\infty^* \in (0,v)$.

We first show that for any $\varepsilon > 0$ and for every $\delta \in (0, c_\infty^*)$, there is a $n_{\varepsilon,\delta}$ such that $p_n^*(\delta) - p_n^*(c_\infty^*) < \frac{\varepsilon}{2}$ if $n > n_{\varepsilon,\delta}$.

Suppose by contradiction that this is not true, then, for some $\delta \in (0, c_\infty^*)$, it must be that $p_n^*(\delta) - p_n^*(c_\infty^*) > \varepsilon$ for all n. Define $\tau_{0,n-1}$ as the probability that there is at least one member out of $n-1$ with cost $c > c_n^*$; similarly let $\tau_{1,n-1}$ denote the probability that there is at least one member out of $n-1$ with cost $c < \delta$. Denote by P_0^n, P_1^n, P_B^n the probability of success conditioning on, respectively, the presence of a type $c \leq \delta$ and no type $c \geq c_n^*$; the presence of a type $c \geq c_n^*$ and no type $c \leq \delta$; and the presence of both a type $c \leq \delta$ and a

type $c \geq c_n^*$. Then, using this notation, we have:

$$E\left[p_n^*(c)|c \leq \delta\right] = \tau_{0,n-1}P_B^n + (1-\tau_{0,n-1})P_0^n$$
$$p_2 = \tau_{1,n-1}P_B^n + (1-\tau_{1,n-1})P_1^n$$

and furthermore we have: $p_n^*(\delta) \leq E\left[p_n^*(c)|c \leq \delta\right]$ and $p_n^*(c_\infty^*) \geq p_2$. So we have:

$$\begin{aligned} 0 &\leq p_n^*(\delta) - p_n^*(c_\infty^*) \leq E\left[p_n^*(c); c \leq \delta\right] - p_2 \\ &\leq \left(\tau_{0,n-1} - \tau_{1,n-1}\right) P_B^n + (1-\tau_{0,n-1})P_0^n - (1-\tau_{1,n-1})P_1^n \end{aligned}$$

As $n \to \infty$, both $\tau_{0,n-1}$ and $\tau_{1,n-1}$ converge to 1. Since P_0^n, P_1^n and P_B^n are all bounded, we have that for any $\varepsilon > 0$ and $\delta > 0$, there is a $n_{\varepsilon,\delta}$ such that $p_n^*(\delta) - p_n^*(c_\infty^*) < \frac{\varepsilon}{2}$ for all $n > n_{\varepsilon,\delta}$.

Finally, by incentive compatibility: $v\left[p_n^*(0) - p_n^*(\delta)\right] \leq \left[a_n^*(0) - a_n^*(\delta)\right]\delta \leq \delta$. Hence, if we set $\delta/v = \varepsilon/2$:

$$p_n^*(0) - p_n^*(v\frac{\varepsilon}{2}) \leq \frac{\varepsilon}{2}. \tag{40}$$

Furthermore, for any ε, there is a n_ε such that:

$$n > n_\varepsilon \Longrightarrow p_n^*(v\frac{\varepsilon}{2}) - p_n^*(c_\infty^*) \leq \frac{\varepsilon}{2} \tag{41}$$

Combining (40) and (41) implies that for any ε, there is a n_ε such that for $n > n_\varepsilon$: $p_n^*(0) - p_n^*(c_\infty^*) \leq \varepsilon$. ■

9.11 Proof of Theorem 3

As discussed in Section 5.1, we derive the objective function in the relaxed problem (17) using the (IC) constraint in a similar way as in (9). The constraint in the second line of (17) is just the implication of IC, also present in (9). The constraint in the third line of (17) follows from the following lemma:

Lemma A4: *If (IC) and (1) hold, then $a(c) = 0$ for $c > c^*$, where*

$$c^* = \min\left\{c \leq v \,|vp(c) - ca(c) \leq vp_2\right\}.$$

Proof: By (1), $c > v$ implies $a(c) = 0$. Consider any $c \in [c^*, v]$. By the definition of c^*, $U(c) \leq U(v)$; by (IC), and monotonicity implies that $U(c) \geq U(v)$, thus $U(c) - U(v) = 0$ for $c \in [c^*, v]$. This implies: $\int_{c^*}^{v} a(x)dx = \int_{v}^{c^*} U'(x)dx = U(c^*) - U(v) = 0$. Since $a(c)$ is monotonic, we get $a(c) = 0$ for $c > c^*$. ■

Let $V_n^{**}(c)$ denote the expected value for a type c in an optimal mechanism that solves the relaxed problem (17) and $V_n^{**} = E_c\{V_n^{**}(c)\}$ be the value of the objective function in (17). Note that $V_n^{**} \geq V_n^*$, where V_n^* is the optimal mechanism in an honest and obedient mechanism, since (17) is a relaxed version of (16).

Define an *ε-bounded mechanism*, $\widetilde{\mu}_n^{\varepsilon}(\mathbf{c})$, and associated reduced form mechanism $\widetilde{a}_n^{\varepsilon}(c)$, $\widetilde{p}_n^{\varepsilon}(c)$ as follows. It is solves the problem for the optimal mechanism (17), but with an additional condition:

$$\widetilde{a}_n(c) > 0 \Rightarrow p_n(0) - \widetilde{p}_n(c) < \varepsilon$$

The value for a type c and the expected values of this mechanism are $\widetilde{V}_n^{\varepsilon}(c)$ and $\widetilde{V}_n^{\varepsilon}$, respectively When $\varepsilon = 1$ (or larger), the additional constraint is slack, so $\widetilde{V}_n^{\varepsilon} = V_n^{**}$. When $\varepsilon = 0$, $\widetilde{\mu}_n^{\varepsilon}(\mathbf{c})$ is a binary mechanism, that is there is a $\widetilde{c}_n^{\varepsilon}$ such that $\widetilde{p}_n^{\varepsilon}(c) = \widetilde{p}_n^{\varepsilon}(0)$ for $c \leq \widetilde{c}_n^{\varepsilon}$ and $\widetilde{p}_n^{\varepsilon}(c) = \widetilde{p}_n^{\varepsilon}(1)$ for $c > \widetilde{c}_n^{\varepsilon}$. Moreover, incentive compatibility implies that $\widetilde{a}_n^{\varepsilon}(c) = \widetilde{a}_n^{\varepsilon}(0)$ for $c \leq \widetilde{c}_n^{\varepsilon}$ and $\widetilde{a}_n^{\varepsilon}(c) = 0$ for $c > \widetilde{c}_n^{\varepsilon}$. We denote a binary mechanism as follows: $\mu_n^b(\mathbf{c})$ and associated $a_n^b(c)$, $p_n^b(c)$ with values $V_n^b(c)$ and V_n^b.

We proceed in two steps:

Step 1: For any η, there exists n_η such that $n > n_\eta \Longrightarrow V_n^b \geq V_n^{**} - \eta$, and hence $V_n^b \geq V_n^* - \eta$.

Consider $D_n = V_n^{**} - \widetilde{V}_n^0 = V_n^{**} - V_n^b$. There are two possibilities:

(1) $\lim_{n\to\infty} D_n = 0$. In thhis case:

$$\lim_{n\to\infty} D_n = \lim_{n\to\infty} \left(V_n^{**} - V_n^b\right) = \lim_{n\to\infty} \left(V_n^{**} - V_n^G\right) = 0$$

where V_n^G is the value in the optimal generalized VBO and we are done.

(2) $\lim_{n\to\infty} D_n = D > 0$. In this case let $\eta \in (0, D)$, and for any such η and any n, define $\varepsilon(n, \eta)$ as follows:

$$V_n^b = \widetilde{V}_n^{\varepsilon(n,\eta)} - \eta/2$$

Note that $\varepsilon(n, \eta) \in (0, 1)$ for any n. It follows that $\lim_{n\to\infty} \varepsilon(n, \eta)$ exists and $\lim_{n\to\infty} \varepsilon(n, \eta) = \varepsilon(\eta) \in [0, 1]$.

Suppose that $\varepsilon(\eta) > 0$. Then for any $\varepsilon' \leq \varepsilon(\eta)$, there is a n_η^1 such that for $n > n_\eta^1$ we have $V_n^b \geq \widetilde{V}_n^{\varepsilon'} - \eta/2$. From Lemma 2, we know that there is a n_η^2 such that for $n > n_\eta^2$ we have $\widetilde{V}_n^{\varepsilon} = V_n^{**}$, we conclude that for $n > \max\{n_\eta^1, n_\eta^2\}$, $V_n^b \geq V_n^{**} - \eta$, a contradiction with the assumption that $\lim_{n\to\infty} D_n = D > 0$.

Therefore $\varepsilon(\eta) = 0$. The rest of the proof of Step 1 relies on the following lemma:

Lemma A5. *If $\lim_{n\to\infty} \varepsilon(n, \eta) = 0$, then for any arbitrarily small $\epsilon \in (0, \eta/2)$, there is an n_ϵ such that for $n > n_\epsilon$, we have $\widetilde{V}_n^{\varepsilon(n,\eta)} \leq \widetilde{V}_n^0 + \epsilon$.*

Proof. We first prove that, for any sequence ε_m such that $\varepsilon_m \to 0$ as $m \to \infty$ and $\varepsilon_m \leq \varepsilon_{m-1}$, we have:

$$\lim_{m\to\infty} \lim_{n\to\infty} \widetilde{V}_n^{\varepsilon_m} = \lim_{n\to\infty} \lim_{m\to\infty} \widetilde{V}_n^{\varepsilon_m}$$

To this goal define for convenience $\widetilde{V}_n^j = \widetilde{V}_n^{\varepsilon_j}$ and:

$$a_n^j = \widetilde{V}_n^{j-1} - \widetilde{V}_n^j \text{ with } a_n^0 = 0,$$

noting that by construction $a_n^j \geq 0$. We can write:

$$\lim_{n\to\infty}\lim_{m\to\infty}\widetilde{V}_n^m = -\lim_{n\to\infty}\lim_{m\to\infty}\sum\nolimits_{j=1}^{m} a_n^j$$

since $\widetilde{V}_n^j = -\sum_{j=1}^{m} a_n^j$. So we have:

$$\lim_{n\to\infty}\lim_{m\to\infty}\widetilde{V}_n^j = -\lim_{n\to\infty}\lim_{m\to\infty}\sum\nolimits_{j=1}^{m} a_n^j = -\lim_{n\to\infty}\sum\nolimits_{j=1}^{\infty} a_n^j = -\lim_{n\to\infty}\int a_n$$

where the $\int a_n = \sum_{j=1}^{\infty} a_n^j = \lim_{m\to\infty}\sum_{j=1}^{m} a_n^j$ is the Lebesgue integral of a_n with respect to the counting measure.

Define now the sequence $(g^j)_{j=0}^{\infty}$ such that $g^j = \sup_n a_n^j$. Note that $g^j \geq a_n^j \geq 0$ for all n by definition. Moreover, g^j is integrable with respect to the counting measure since a_n^j is integrable and integrability is passed by the *sup* operator. Finally, we have

$$\int g_n = \lim_{m\to\infty}\sum_{j=1}^{m} g^j = \lim_{m\to\infty}\sum_{j=1}^{m}\sup_n a_n^j = \lim_{m\to\infty}\sup_n\sum_{j=1}^{m} a_n^j = \lim_{m\to\infty}\sup_n \widetilde{V}_n^m \leq v$$

where v is the value of the collective good. We conclude that g^j is an integrable function that dominates a_n^j for any n. We can therefore apply the Dominated Convergence Theorem as follows:

$$\begin{aligned}\lim_{n\to\infty}\lim_{m\to\infty}\widetilde{V}_n^m &= -\lim_{n\to\infty}\int a_n = -\int \lim_{n\to\infty} a_n \\ &= -\lim_{m\to\infty}\sum\nolimits_{j=1}^{m}\lim_{n\to\infty} a_n^j = -\lim_{m\to\infty}\lim_{n\to\infty}\sum\nolimits_{j=1}^{m} a_n^j = \lim_{m\to\infty}\lim_{n\to\infty}\widetilde{V}_n^m\end{aligned}$$

So we have $\lim_{n\to\infty}\lim_{m\to\infty}\widetilde{V}_n^m = \lim_{m\to\infty}\lim_{n\to\infty}\widetilde{V}_n^m$.

Now note that:

$$\begin{aligned}\lim_{n\to\infty}\widetilde{V}_n^{\varepsilon_m} &= \lim_{n\to\infty} V_n^{**} \Leftrightarrow \lim_{m\to\infty}\lim_{n\to\infty}\widetilde{V}_n^{\varepsilon_m} = \lim_{m\to\infty}\lim_{n\to\infty} V_n^{**} \\ &\Leftrightarrow \lim_{m\to\infty}\lim_{n\to\infty}\widetilde{V}_n^{\varepsilon_m} = \lim_{n\to\infty} V_n^{**} \geq \lim_{n\to\infty}\widetilde{V}_n^{\varepsilon(n,\eta)}\end{aligned}$$

since we have proven in Lemma 2 that the optimal (HO) mechanism becomes approximately flat for large n, so the constraint becomes slack for n large enough. Moreover, $\widetilde{V}_n^{\varepsilon(n,\eta)} \geq \lim_{m\to\infty}\widetilde{V}_n^{\varepsilon_m} = \widetilde{V}_n^0$. It follows that:

$$\begin{aligned}\lim_{n\to\infty}\lim_{m\to\infty}\widetilde{V}_n^{\varepsilon_m} &\leq \lim_{n\to\infty}\widetilde{V}_n^{\varepsilon(n,\eta)} \leq \lim_{m\to\infty}\lim_{n\to\infty}\widetilde{V}_n^{\varepsilon_m} = \lim_{n\to\infty}\lim_{m\to\infty}\widetilde{V}_n^{\varepsilon_m} \\ &\Leftrightarrow \lim_{n\to\infty}\widetilde{V}_n^{\varepsilon(n,\eta)} = \lim_{n\to\infty}\lim_{m\to\infty}\widetilde{V}_n^{\varepsilon_m} = \lim_{n\to\infty}\widetilde{V}_n^0\end{aligned}$$

It follows that for any $\epsilon > 0$, there is an n_ϵ such that for $n > n_\epsilon$, we have $\widetilde{V}_n^{\varepsilon(n,\eta)} \leq \widetilde{V}_n^0 + \epsilon$, where $\epsilon - \eta/2 < 0$ ■

From Lemma A5, we have that, for $n > n_\epsilon$, $V_n^b = \widetilde{V}_n^{\varepsilon(n,\eta)} - \eta \leq \widetilde{V}_n^b + \epsilon - \eta/2 < V_n^b$, a contradiction. Form the fact that we obtain a contradiction for any $\varepsilon(\eta) \geq 0$, we conclude that $\lim_{n\to\infty}\left(V_n^{**} - V_n^G\right) = 0$. ■

Step 2. We can now put together Step 1 and Lemma 1 to argue that a VBO is approximately optimal for n large. Since $V_n^b \leq V_n^*$, Step 1 implies that $\left|V_n^b - V_n^*\right| \to 0$ as $n \to \infty$. Lemma 1, moreover, shows that the optimal binary mechanism is a generalized VBO with threshold k_n^*. Let $VBO(k_n^*)$ be a VBO with threshold $k = k_n^*$ *and no mixing* for $k = k_n^*$. The $VBO(k_n^*)$ generates utility that converges to the utility of the generalized VBO with threshold k_n^*, since the probability of exactly $k = k_n^*$ volunteers converges to zero. Since the generalized VBO is equivalent to an optimal binary mechanism that generates utility V_n^b, we have $\left|V_n^b - V_n^{VBO(k_n^*)}\right| \to 0$. Hence, we have that for any η there is a n_η such that for $n > n_\eta$ $V_n^{VBO(k_n^*)} \geq V_n^* - \eta$ for some threshold k_n^*, which implies the result. ■

9.12 Proof of Theorem 4

We complete here the second step of the proof of the second bullet point of Theorem 4 by showing that if the probability of success in the best binary mechanism (which is a general VBO) converges to zero, then the probability of success in the optimal mechanism converges to zero as well. This implies that the expected welfare in the two mechanisms converge to the same value. For this we use Theorem 3, that proves that the general VBO is approximately optimal.

Suppose by contradiction that the probability of success in the optimal mechanism P_n^* (non necessarily binary or VBO) converges to some positive value $P^* > 0$, but the probability of success in the optimal generalized VBO mechanism P_n^G converges to zero. Let W_n^* and W_n^G be the expected per capita welfare in the optimal mechanism and in the optimal VBO. Note that for any ε, there is a $n_{1,\varepsilon}$ such that for $n > n_{1,\varepsilon}$:

$$W_n^G = vP_n^G \left(1 - E\left(\frac{a^G(c)}{P_n^G} \cdot \frac{c}{v}\right)\right) \leq vP_n^G \leq \varepsilon/2$$

since by assumption $P_n^G \to 0$. Moreover for any ε, there is a $n_{2,\varepsilon}$ such that for $n > n_{2,\varepsilon}$

$$W_n^* = vP_n^* \left(1 - E\left(\frac{a_n^*(c)}{P_n^*} \cdot \frac{c}{v}\right)\right) \geq vP^* - \varepsilon/2 > 0$$

since: a. for all $c \leq v$, $a_n^*(c)\frac{c}{v} \leq p_n^*(c) - p_n^*(v) \to 0$, as proven in Theorem 4; and b. $P_n^* \to P^* > 0$. It follows that for any ε, there is a $n_\varepsilon = \max\{n_{1,\varepsilon}, n_{2,\varepsilon}\}$ such that for $n > n_\varepsilon$, $W_n^* - W_n^G > vP^* - \varepsilon$.

By Theorem 3, for any arbitrarily small $\eta > 0$, there is a n_η such that for $n > n_\eta$, $\left|W_n^* - W_n^G\right| < \eta$, where W_n^* and W_n^G are the expected per capita welfare in the optimal mechanism and in the optimal VBO. It follows that for n large, $\eta + \varepsilon > \left|W_n^* - W_n^G\right| \geq vP^*$, which is a contradiction since η and ε are both arbitrarily small, and vP^* is bounded away from zero. ■

9.13 Proof of Proposition 7

We can bound the probability of success in an optimal honest and obedient mechanism as follows. Let $D\left(\frac{k}{n}\,\|p\right) = \frac{k}{n}\log\frac{k/n}{p} + (1-\frac{k}{n})\log\frac{1-k/n}{1-p}$ be the Kullback–Leibler divergence, or relative entropy. We can write:

$$\begin{aligned}\lim_{n\to\infty} P_n^* &= \lim_{n\to\infty}\sum\nolimits_{k=m_n}^{n} B\left(k,n,p_n^*\right) \geq \lim_{n\to\infty} P_n^O = \sum\nolimits_{k=m_n}^{n} B\left(k,n,p_n^O\right)\\ &\geq \lim_{n\to\infty}\frac{1}{\sqrt{8m_n(1-\frac{m_n}{n})}}\exp\left(-nD\left(\frac{m_n}{n}\,\|p_n^O\right)\right) \simeq \lim_{n\to\infty}\frac{1}{\sqrt{8m_n}}\exp\left(-nD\left(\frac{m_n}{n}\,\|p_n^O\right)\right)\end{aligned}$$

where in the last line we used the lower bound on the tail of a Binomial distribution (Lemma 4.7.2 in Ash [1990]). Since $\frac{m_n}{n}\to p_n^O$, we have that for any $\epsilon>0$, there exists n_ϵ such that for $n>n_\epsilon$, we have $D\left(\frac{m_n}{n}\,\|p_n^*\right)\leq\epsilon/2$. So

$$\begin{aligned}\lim_{n\to\infty}\frac{P_n^*}{m_n^{-1/2}e^{-\epsilon n}} &\geq \lim_{n\to\infty}\frac{\frac{1}{\sqrt{8m_n}}\exp\left(-nD\left(\frac{m_n}{n}\,\|p_n^*\right)\right)}{e^{-\epsilon n}}\\ &= \lim_{n\to\infty}\frac{e^{\left[\epsilon-\exp\left(-nD\left(\frac{m_n}{n}\|p_n^*\right)\right)\right]n}}{\sqrt{8m_n}} \geq \lim_{n\to\infty}\frac{e^{\frac{\epsilon}{2}n}}{\sqrt{8m_n}} = \infty\end{aligned}$$

So for any $\epsilon>0$, P_n^* converges strictly faster than $e^{-\epsilon n}$. We conclude that P_n^* converges to zero at a rate that is strictly lower than exponential. ■

9.14 Proof of Proposition 8

If $m_n \succ n^{2/3}$, $c_n^O>0$ for all n and $c_n^U=0$ for n sufficiently large, so ΔV_n^* is proportional to v, and thus clearly increasing in v. If instead $m_n \prec n^{2/3}$, then $F(c_n^O)>m_n/n$ and $F(c_n^U)>m_n/n$, so the first terms in the square parenthesis of (21) and (22) converge to zero faster than the second terms, which can therefore be ignored for large n. Hence, for large enough n:

$$EU_O(c_n^*) - EU_U(c_n^U) \simeq v\left[\sum_{j=\alpha n}^{n-1} B(j,n-1,c_n^*) - \sum_{j=\alpha n}^{n-1} B(j,n-1,c_n^U)\right]$$

which is strictly increasing in v since $\sum_{j=\alpha n}^{n-1} B(j,n-1,c)$ is strictly increasing in c and $c_n^O>c_n^U$ by Proposition 4. ■

9.15 Proof of Theorem 5

In the optimal honest and obedient mechanism the high type is never activated, so the mechanism, in reduced form, is characterized by a probability of activation a^μ for $\underline{c}$, an expected probability of success p_1^μ when a low type reports to be a low type and a probability of success when the low type reports to be a high type, p_2^μ. Incentive compatibility implies

that:

$$a^{\mu}\underline{c} \leq v(p_1^{\mu} - p_2^{\mu})$$

We now prove that the mechanism must looks as follows: (1) It elicits the types truthfully; and (2) Let k be the number of low types. There is a $k^* \geq \alpha n$ such that for $k < k^*$, no agent is activated. For $k > k^*$, a coalition of αn agents among the reported low types is selected and activated, triggering success. If $k = k^*$, there is a probability π such that a coalition of $m = \alpha n$ agents among the low types is selected and activated, triggering success. When a coalition of αn agents is activated, the self-reported low types have equal probability of being selected.

We proceed in three steps.

Step 1. We first show that the optimal mechanism must be non-wasteful in the sense that whenever a group is activated, there are exactly $m = \alpha n$ members in the activated group. We prove this by contradiction.

Suppose that the optimal mechanism μ is wasteful at a positive measure set of profiles. First, define a new mechanism, μ', that is exactly the same as μ for all coalitions of size m, but eliminates all waste by reducing all activated successful coalitions to a size m by randomly selecting agents to drop out, and do not activate coalitions that are smaller than m. This leaves p_1^{μ} and p_2^{μ} unchanged and reduces a^{μ} to $a^{\mu\prime} < a^{\mu}$. This implies that $a^{\mu\prime}\underline{c} < v(p_1^{\mu} - p_2^{\mu})$. Note that after the modification the mechanism must be characterized by a set of probabilities q_k such that with $k \geq \alpha n$ low types, the probability that a randomly selected coalition of size αn is activated is q_k. If $q_k = 1$ for all $k \geq \alpha n$, then we have proven the result. Assume therefore that $q_k < 1$ for some $k \geq \alpha n$. But then we can increase q_k to $q_k' > q_k$ in order to increase $a^{\mu\prime}$ to $a^{\mu\prime, q_k'}$ and still satisfy $a^{\mu\prime, q_k'}\underline{c} \leq v(p_1^{\mu} - p_2^{\mu})$. After this change the probability of success and welfare are higher, which contradicts the assumption that μ is optimal.

Step 2. Incentive compatibility is satisfied if $a^{\mu}\underline{c} \leq v(p_1^{\mu} - p_2^{\mu})$. Note that incentive compatibility would be satisfied even if the inequality were strict (in this case the low type would strictly find it optimal to report to be a low type). Optimality however implies that the inequality is satisfied as an equality. If not, then $q_k = 1$ for all $k \geq \alpha n$ and $q_k = 0$ for all $k < \alpha n$; or we could profitably increase some q_k as in Step 1. We conclude that $a^{\mu}\underline{c} = v(p_1^{\mu} - p_2^{\mu})$.

Step 3. Note again that a mechanism must be characterized by a set of probabilities q_k such that with $k \geq \alpha n$ low types, the probability that a randomly selected coalition of size αn is activated is q_k. We will prove the result that if $q_k > 0$ for some $k \geq \alpha_n n$, then $q_{k+j} = 1$ for $j > 0$. Note that if $q_k > 0$ for $k \geq \alpha_n n$ and $q_{k+j} < 1$ for $j > 1$, then there must be a k' such that $q_{k'} > 0$ for $k' \geq \alpha_n n$ and $q_{k'+1} < 1$, so we only need to prove the result for the case of $j = 1$.

Assume by contradiction that $q_k > 0$ for some $k \geq \alpha_n n$ and $q_{k+1} < 1$. Let a be the expected probability of participation for a low cost $\underline{c}$; and p_1, p_2 the expected probabilities of success for a low type conditioning on reporting truthfully or not. By Step 2 IC implies

$a^\mu \underline{c} = v(p_1^\mu - p_2^\mu)$ where:

$$\begin{aligned} p_1^\mu - p_2^\mu &= B(n-1, n-1, \phi) q_n & (42) \\ &+ \sum\nolimits_{k=\alpha n}^{n-1} [B(k-1, n-1, \phi) - B(k, n-1, \phi)] q_k, & (43) \\ \text{and } a &= \sum\nolimits_{k=\alpha n-1}^{n-1} \frac{\alpha n}{1+k} B(k, n-1, \theta) q_{k+1} & (44) \end{aligned}$$

These are exactly the same formulas as presented in step 2 of Lemma 1, except that now the probability of a "low type" is not c^b, but θ. We can marginally reduce q_k by $-dq_k < 0$ and marginally increase q_{k+1} by $dq_{k+1} > 0$ so that the (IC) constraint is unchanged. Following exactly the same steps as in step 2 of Lemma 1, we can show that this requires:

$$\frac{dq_{k+j}}{dq_k} = T_{n,k} \frac{B(k, n-1, \phi)}{B(k+1, n-1, \phi)}$$

where:

$$T_{n,k} = \frac{1 - \left(1 - \underline{c}\frac{\alpha n}{k}\right) \frac{k}{n-k} \frac{1-\phi}{\phi}}{1 - \left(1 - \underline{c}\frac{\alpha n}{k+1}\right) \frac{k+1}{n-k-1} \frac{1-\phi}{\phi}} > 1 \Leftrightarrow \frac{k - \underline{c}m}{n-k} < \frac{k - \underline{c}m + 1}{n-k-1} \Leftrightarrow n > \underline{c}m$$

After this change, the probability of success increases, indeed we have:

$$\begin{aligned} dP_n &= \left[-B(k, n-1, \phi) + B(k+1, n-1, \phi) \frac{dq_{k+1}}{dq_k} \right] dq_k \\ &= \left[-B(k, n-1, \phi) + B(k+1, n-1, \phi) \cdot T_{n,k} \frac{B(k, n-1, \phi)}{B(k+1, n-1, \phi)} \right] dq_k \\ &> (T_{n,k} - 1)\, B(k, n-1, \theta) \cdot dq_k > 0 \end{aligned}$$

Since the probability of success increases, welfare must increase. This implies that the original mechanism was not optimal, a contradiction.

It follows from the argument above that the optimal mechanism is characterized by a threshold $k^* \geq \alpha n$ and a probability $\pi \in (0, 1]$ such that $q_k = 1$ for $k \geq k^*$, $q_k = 0$ for $k < k^*$ and $q_k = \pi$ for $k = k^*$. Any such k^* and π uniquely defines $p_1^{k^*,\pi} - p_2^{k^*,\pi}, a^{k^*,\pi}$ from equations (42) and (44) above.[39] Since the objective function is decreasing in k^* it follows that the optimal mechanism is then characterized by the least integer $k^* \geq \alpha n$ for which there exists $\pi \in (0, 1]$ such that: $a^{k^*,\pi} \underline{c} \leq v(p_1^{k^*,\pi} - p_2^{k^*,\pi})$, where:

$$p_1^{k^*,\pi} - p_2^{k^*,\pi} = B(n-1, n-1, \phi) + \pi \begin{bmatrix} B(k^*-1, n-1, \phi) \\ -B(k^*, n-1, \phi) \end{bmatrix} + \sum\nolimits_{k=k^*+1}^{n-1} \begin{bmatrix} B(k-1, n-1, \phi) \\ -B(k, n-1, \phi) \end{bmatrix}$$

and $a^{k^*,\pi} = \pi B(k^*-1, n-1, \phi) + \sum_{k=k^*}^{n-1} \frac{\alpha n}{1+k} B(k, n-1, \phi)$. ■

[39] Notice that $q_n = 1$ ($k^* < n$) since IC is not binding for $k^* = n$.

9.16 Proof of Corollary 3

For ease of notation, instead of writing m_n, simply write m as the minimum number of activated volunteers that is required for success, since the argument below does not depend on n. Let μ be an optimal $2-type$ mechanism. From Theorem 5, an optimal two-type mechanism is characterized by threshold $k^* \geq m$ such that and $q_k = 0$ for $k < k^*$ and $q_k = 1$ for $k > k^*$.Suppose by contradiction that $\frac{\underline{c}}{v} < \frac{1-\frac{\phi}{\alpha}}{1-\phi}$ and $k^* > m$, so that $q_m = 0$. Since μ is optimal, it must be the case that increasing q_m while holding q_{-m} fixed (which increases the objective function) must violate the incentive constraint: $v(p_1 - p_2) - a\underline{c} \geq 0$. Furthermore, the fact that increasing q_m increases the objective function and $q_m = 0$ in the optimal mechanism implies that the (IC) constraint must hold with equality. Otherwise, one could increase q_m slightly without violating the incentive constraint, so:

$$v(p_1 - p_2) - a\underline{c} = 0. \tag{45}$$

Recall that the exact formulas for p_1, p_2, a in terms of the q_k's are given in equations (F2). Substituting these three expressions into the incentive constraint (45) gives:

$$\Psi(q) \equiv v \left[\begin{array}{c} \sum_{k=m-1}^{n-1} B(k, n-1, \phi) q_{k+1} \\ -\sum_{k=m}^{n-1} B(k, n-1, \phi) q_k \end{array} \right] - \underline{c} \sum_{k=m-1}^{n-1} \frac{m}{1+k} B(k, n-1, \phi) q_{k+1} = 0$$

Since $q_m = 0$ in the optimal mechanism, it must be that $\frac{\partial \Psi}{\partial q_m} \leq 0$. Simple calculation gives:

$$\frac{\partial \Psi}{\partial q_m} = v(B(m-1, n-1, \phi) - B(m, n-1, \phi)) - \underline{c} B(m-1, n-1, \phi)$$

But we have the binomial identity: $B(m, n-1, \phi) = \frac{n-m}{m} \frac{\phi}{1-\phi} B(m-1, n-1, \phi)$. So $\frac{\partial \Psi}{\partial q_m}$ reduces to:

$$\begin{aligned} \frac{\partial \Psi}{\partial q_m} &= vB(m-1, n-1, \phi)(1 - \frac{n-m}{m} \frac{\phi}{1-\phi}) - \underline{c} B(m-1, n-1, \phi) \\ &= B(m-1, n-1, \phi) \left[v(1 - \frac{n-m}{m} \frac{\phi}{1-\phi}) - \underline{c} \right] \end{aligned}$$

Hence $\frac{\partial \Psi}{\partial q_m} \leq 0$ if and only if: $\frac{\underline{c}}{v} \geq 1 - \frac{n-m}{m} \frac{\phi}{1-\phi}$. This implies that if $\frac{\underline{c}}{v} < \frac{1-\frac{\phi}{\alpha}}{1-\phi}$ and $k^* > m$ we have a contradiction. We conclude that if $\frac{\underline{c}}{v} < \frac{1-\frac{\phi}{\alpha}}{1-\phi}$, then $k^* = m$ and $q_k = 1$ for all $k \geq m$. ■

9.17 Proof of Proposition 11

Let θ_n be the threshold chosen in the general VBO with n agents. If $\lim_{n\to\infty} \theta_n > \phi$, then we have that $P \to 0$ and the result is proven. Assume therefore that $\lim_{n\to\infty} \theta_n \leq \phi$. Note that the (IC) constraint requires:

$$a\frac{\underline{c}}{v} \leq p_{1,n} - p_{2,n} \tag{46}$$

where $p_{1,n}, p_{2,n}$ are the success probability conditioning on a low and high type. The right hand side converges to zero at the speed of $B(\lfloor \theta n \rfloor - 1, n-1, \phi)$. Since $\theta_n < \phi$, we must have:

$$B(\lfloor \theta_n n \rfloor - 1, n-1, \phi) \leq B(\lfloor \phi n \rfloor - 1, n-1, \phi) \simeq \frac{1}{\sqrt{n}}$$

Consider now the left hand side of (46). Note that for any $\eta, \varepsilon > 0$:

$$\begin{aligned}
\Pr\left(k \geq \phi n + \eta n^{1/2+\varepsilon}\right) &= \Pr\left(\frac{k}{n} - \phi \geq \frac{\eta}{n^{1/2-\varepsilon}}\right) = \Pr\left(\frac{k}{n} - \phi \geq \sqrt{\frac{\phi(1-\phi)}{n}} \frac{n^{\varepsilon}\eta}{\sqrt{\phi(1-\phi)}}\right) \\
&= \Pr\left(\frac{k}{n} - \phi \geq \sigma_{\phi}(k/n) \frac{n^{\varepsilon}\eta}{\sqrt{\phi(1-\phi)}}\right) \leq \frac{\phi(1-\phi)}{\eta^2 n^{2\varepsilon}} \to 0
\end{aligned}$$

where $\sigma_{\phi}(k/n) = \sqrt{\phi(1-\phi)/n}$. Therefore, we can write:

$$\begin{aligned}
\lim_{n\to\infty} a_n &= \lim_{n\to\infty} \sum\nolimits_{k=\theta_n n-1}^{\left(1+\frac{\eta}{\phi n^{1/2-\varepsilon}}\right)\phi n-1} \frac{\alpha_n n}{1+k} B\left(k, n-1, \phi\right) \\
&\geq \lim_{n\to\infty} \frac{\alpha_n n}{\left(1+\frac{\eta}{\phi n^{1/2-\varepsilon}}\right)\phi n} \cdot \lim_{n\to\infty} \sum\nolimits_{k=\theta_n n-1}^{\left(1+\frac{\eta}{\phi n^{1/2-\varepsilon}}\right)\phi n-1} B\left(k, n-1, \phi\right) \\
&= \lim_{n\to\infty} \frac{\alpha_n n}{\left(1+\frac{\eta}{\phi n^{1/2-\varepsilon}}\right)\phi n} \cdot \lim_{n\to\infty} \sum\nolimits_{k=\theta_n n-1}^{n-1} B\left(k, n-1, \phi\right) \\
&= \lim_{n\to\infty} \frac{\alpha_n}{\phi} \cdot \lim_{n\to\infty} p_{1,n}
\end{aligned}$$

where $p_{1,n}$ is the conditional probability of success for a low type with n agents. Two things are possible. Suppose first $\lim_{n\to\infty} p_{1,n} = 0$, then the probability of success is also converging to zero and the result is proven. To see this formally note that:

$$P_n = \phi p_{1,n} + (1-\phi)\, p_{2,n} \leq p_{1,n}$$

since $p_{1,n} \geq p_{2,n}$. Assume therefore $\lim_{n\to\infty} p_{1,n} = \widehat{p} > 0$. Then we have:

$$\lim_{n\to\infty} a_n \frac{\underline{c}}{v} = \lim_{n\to\infty} \frac{\alpha_n}{\phi} \cdot \widehat{p} \cdot \frac{\underline{c}}{v} = \left(\frac{\widehat{p}}{\phi} \cdot \frac{\underline{c}}{v}\right) \cdot \lim_{n\to\infty} \frac{m_n}{n}$$

It follows that $\frac{p_{1,n}-p_{2,n}}{a_m} \simeq \frac{n^{1/2}}{m_n} \to 0$, since $m_n \succ n^{1/2}$. It follows that $\frac{p_{1,n}-p_{2,n}}{a_m} \to 0$, implying $\frac{\underline{c}}{v} \leq 0$, a contradiction. ■