INSTITUTO DE ECONOMÍA

DOCUMENTO de TRABAJO

538
2020

Should the government provide public goods if it cannot commit?

Francisco Silva.

www.economia.puc.cl • ISSN (edición impresa) 0716-7334 • ISSN (edición electrónica) 0717-7593

Should the government provide public goods if it cannot commit? *

Francisco Silva †

July 20, 2019

Abstract

I compare two different systems of provision of binary public goods: a centralized system, ruled by a benevolent dictator who has limited commitment power; and a decentralized system, based on voluntary contributions, where agents can communicate but cannot write contracts. I show that any ex-post individually rational allocation that is implementable by the centralized system is also implementable by the decentralized system. This suggests that when the public good provision problem is merely an informational one, as is the case with binary public goods, a decentralized system may perform better.

JEL classification: D82, H41

Keywords: communication, free-riding, commitment power, voluntary provision of public goods

*I would like to thank Martin Besfamille, Mustafa Dogan, Nicolas Figueroa, Tzuo-Hann Law, George Mailath, Samir Mamadehussene, Steven Matthews, Daniel Neuhann, Qiusha Peng, Andrew Postlewaite, Rakesh Vohra, Yanhao Wei and Felipe Zurita as well as the seminar participants at PUC Chile, the game theory international conference at Stony Brook and the Latin American workshop in economic theory at Bogota for their helpful comments.

†Department of Economics, Pontificia Universidad Catolica de Chile, Vicuña Mackenna 4860, Piso 3. Macul, Santiago, Chile. (email: franciscosilva@uc.cl).

1 Introduction

The question of who should provide public goods is an old one in the economic literature. Is it better to have a centralized system, where some type of ruler determines the quantity of public good provision and who should pay for it? Or is it better instead to have a decentralized system, where agents determine their own contributions? The latter system seems better if the incentives of the ruler are not aligned with the population's, i.e. if the ruler is selfish, evil, etc.[1] What if the ruler is a benevolent dictator (BD)? In that case, the centralized system becomes much more appealing. In particular, if the agents' preferences are public, it is known since Samuelson (1954) that, unlike the centralized system that is ruled by a BD, decentralized alternatives of provision of public goods will typically fail to be efficient, due to what I refer to as the "classical" free riding problem: each agent disregards the positive impact that his private contribution to the provision of public goods has on other agents.

When preferences are private, both systems face the additional difficulty of eliciting that information from the agents. However, it seems clear that, at least if the informational problem is sufficiently irrelevant, the centralized alternative will still do better, because of the classical free-riding problem. But, I would argue, it is less clear which system does better if the informational problem is the dominant one. In particular, in this paper, I argue that there is a trade-off between the two systems. While centralized systems do better when the classical free-riding problem is more important, decentralized systems might do better when it is the "informational" free-riding that is more important.

In order to make my argument, I choose a particular class of public goods precisely because there is no classical free-riding problem: binary public goods. I say that there is no classical free-riding problem because, if there is complete information, there are

[1] See Buchanan and Tullock (1962) and the ensuing literature for a discussion on the incentives of the ruler.

many Nash equilibria of the voluntary contribution game that are efficient.[2]

At first glance, it would appear as though centralized systems would be preferred even when the public good is binary, because of the "revelation principle". Recall that the revelation principle (see, e.g. Myerson (1979)) states that, when designing a mechanism in order to maximize some objective function, one can restrict attention to revelation mechanisms, where agents simply report their private type to a mediator, who then maps those reports into outcomes. If one thinks of the mediator as the BD, it follows that anything that a decentralized system can do, the BD can also do through a revelation mechanism.

There is, however, a key assumption that is made in this argument: that the BD has commitment power. In the typical mechanism design approach to the problem of public good provision, the BD maps the agents' truthful reports to units of the public good to be provided and to transfers each individual must make. The assumption that the BD has commitment power precludes any changes to this promised mapping once the agents make their reports. In particular, even if, after some set of reports, the BD becomes convinced that there is a better alternative than the one that the mechanism imposes (given the particular objective of the BD), she is assumed not to be able to take it. I would argue that this is, at least, a debatable assumption. Who is there to stop the BD from taking the superior alternative? If being able to commit means to be able to write enforceable contracts, who would enforce the contracts written by the BD?

In the paper, I compare two sets of allocations (mappings between the agents' private types and outcomes) in a context where the agents' types are independent: those that can be implemented by a centralized system governed by a BD with limited

[2]To see why that is, consider the following example: say that there are only two people in the community and that it is known that each person values the public good in 5, while the total cost of the public good is 8. It is easy to see that there are many Nash Equilibrium outcomes where the public good ends up being provided.

commitment power and those that can be implemented by a decentralized system, based on the voluntary contributions of the agents. The decentralized system I consider is a variation of what is usually called the "voluntary contribution mechanism" (VCM) (see, for example, Anderson and Putterman (2006) or Masclet et al. (2003)). In the standard VCM, agents simultaneously choose a contribution to the public good provision. If the total contributions exceed the cost of providing the public good, the public good is provided; if not, those contributions are lost.[3] Agastya et al. (2007) study this system and show that there is gain in adding an initial stage to it where agents simultaneously send a public cheap talk message, which can be used to convey information about their preferences, before choosing their contributions.[4] I call this system an extended VCM and compare it to its centralized counterpart: in the BD system, the first stage is the same (agents send a public message), but, in the second stage, the agents' reports simply get mapped by the BD into a decision about whether the public good is provided and who should pay for it.

By assuming that the dictator of the centralized system is benevolent and inequality averse, I find that *any ex-post individually rational allocation* (an allocation for which the ex-post utility of each agent is not negative) *that is implementable by a BD can also be implemented in the extended VCM.* Furthermore, I show by way of example, that there are allocations with appealing properties that can be implemented by the extended VCM but cannot be implemented by the BD.

While at first glance, the result might sound unintuitive, recall that there is no classic free-riding problem so, a priori, there is no reason for the centralized system to perform better. In fact, if anything, one would expect the decentralized system to do better; notice that in the decentralized system, agents are the ones who choose their

[3]The main result also follows under the following alternative assumption: if the total contributions are not enough to fund the provision of the public good, then those contributions are returned to the agents.

[4]In the related literature, I discuss the relation between Agastya et. al. (2007) and this paper.

own transfers, which, one would think, would make them less reluctant to announce their preferences when compared to the centralized alternative, when it is the BD who decides what to do with that information. It turns out that the argument is not as simple, as the fact that the dictator is benevolent and inequality averse has a crucial importance.

The reason why the decentralized system ends up outperforming the BD system is that the set of outcomes that can happen in the second stage in an equilibrium in the extended VCM is larger than in the BD system. For example, in the BD system, second stage outcomes have an efficiency requirement that is not there in the extended VCM; in the extended VCM, it is perfectly possible that it is known by the agents that there is a transfer vector that guarantees that every agent has a positive utility should the public good be provided, and yet each agent chooses not to make a positive contribution. That is not possible in the BD system, because of the BD's lack of commitment power. In a way, the fact that the dictator is benevolent (and inequality averse) harms her ability to elicit the agents' private information given that she cannot commit.

The paper proceeds as follows. I start by describing the related literature. In section 2, I present the model. In section 3, I present and prove the main result. In section 4, I discuss the implications of the main result and some of the assumptions. All proofs are in the appendix.

1.1 Related Literature

I believe this paper makes contributions to three different areas of the economic literature. First, to the literature on public goods. The classic literature on public good provision with incomplete information, which includes Groves (1973), d'Aspremont and Gerard-Varet (1979), Laffont and Maskin (1979) among others, typically assumes the mediator/BD has complete contracting ability. This paper relates more closely to the literature that reduces the commitment power of the BD. Schmidt (1996) provides an argument for the privatization of public firms. The idea is that, if the government is directly responsible for the firm and is unable to commit, it will receive private information that will make it less able to provide incentives for the agents employed by the firm to exert effort. Hence, the author argues, privatization (and subsequent regulation) can be seen as a useful commitment device by the government. The main difference from Schmidt (1996) to this paper is that the former focuses on the moral hazard problem rather than on the adverse selection problem the BD faces.

Second, this paper may be interpreted in light of the literature on the decentralization of the government. The classical analysis of this problem is due to Oates (1972), where the author argues that decentralization will be preferred as long as the provision of the public good in a given region does not generate large enough positive spillovers on the other regions. Besley and Coate (2003) and Lockwood (2002) relax the assumption made in Oates (1972) that a centralized government selects a uniform policy for all its regions but still assume complete information. There are also several papers that analyze the same question under incomplete information but do not allow for communication among the regions (for example Kessler (2014) or Cho (2013)), which limits the benefits of decentralization. Klibanoff and Poitevin (2013) is an exception in that the authors do allow for some bargaining to occur between the regions. However, when modelling the decentralized system, it is assumed that the regions are able to write contracts among themselves. But if it is possible for the regions to write contracts among themselves, then it seems reasonable to also allow the government to write contracts with the regions, which, by the revelation principle would be (at least weakly) preferred. For this reason, in my analysis, the regions (agents) are not allowed to write contracts.

Finally, my analysis of the extended VCM system builds on the notion that allowing agents to communicate enhances considerably the set of allocations that can be implemented through an equilibrium. There is a vast literature that shows that allowing agents to communicate might allow for the implementation of all incentive compatible allocations, which eliminates the need for a mediator (Barany 1992, Ben-Porath 1998 and 2003, Forges 1990, Gerardi 2004). However, typically, these papers present very complicated communication procedures that are not practical, and, as a result, appear highly unrealistic. More similar to this paper, Matthews and Postlewaite (1989) show that, in a bilateral trade setting, the introduction of a single cheap talk stage, prior to having the traders participate in a double action, allows the implementation of a much larger set of allocations. Agastya et al. (2007) explore the same idea in the context of a binary public good provision problem and show that allowing agents to exchange "simple" messages ("yes" or "no" messages) in a voluntary contribution game enlarges the set of implementable allocations.

2 Model

2.1 Fundamentals and definitions

I consider a community with $N > 1$ risk neutral agents. Each agent n is endowed with a private type $v_n \in [\underline{v}_n, \overline{v}_n] \equiv V_n$, where $\overline{v}_n > \underline{v}_n \geq 0$, independent across n. Each type v_n is drawn from a continuous CDF F_n, while each PDF is denoted by f_n. I assume that the public good $g \in \{0, 1\}$. The agent's utility is given by $u_n = v_n g - t_n$, where t_n represents the transfer paid by agent n. The cost of providing the public good is given by $c > 0$, i.e., the public good can only be provided if the sum of transfers received is not lower than c. I assume that $\overline{v}_n < c$ for all n, so that no agent is willing to provide the public good by himself.

An allocation is a function $(\rho, \tau) : V_1 \times ... \times V_N \equiv V \rightarrow \{0, 1\} \times \mathbb{R}^N$, where $\rho(v)$ and $\tau_n(v)$ respectively represent the quantity of the public good that is provided and agent n's transfer for some type vector $v \in V$.[5]

Definition 1 *An allocation (ρ, τ) is ex-post individually rational (IR) if and only if, for all $v \in V$ and for all n,*

$$\rho(v) v_n - \tau_n(v) \geq 0.$$

The set of all ex-post IR allocations is denoted by Ψ^{IR}.

2.2 Extended VCM

In the extended VCM, there are two periods. In the first period, there is a cheap talk round where every agent n simultaneously chooses a public message $m_n \in M_n$, where M_n is some arbitrary message set; in the second period, after observing message vector $m = (m_1, ..., m_N)$, agents simultaneously choose what transfer to make. If the sum of transfers (weakly) exceeds c, the public good is provided.[6]

The crucial assumption is that agents cannot write contracts, so they must stick to cheap talk in the first period. If one was to assume that they were able to write contracts, one would implicitly assume that there is someone, like a BD, who would be willing and able to enforce them. But, if that was the case, one would think that the BD could write the contracts herself in a centralized system. Recall that, by the revelation principle, if the BD is able to commit, a centralized system is preferred, so it is the fact that the BD is unable to commit to enforcing all contracts that makes this comparison relevant.

[5] While this paper focuses only on deterministic allocations, it can be shown that the result of section 3 also holds when random allocations are allowed (in a random allocation, a distribution over (ρ, τ) follows each type vector v).

[6] While, in order to complete the description of the extended VCM, I assume that once an agent has selected a positive transfer, he immediately loses it, this is not necessary for the main result of section 3. An alternative assumption could be that if the sum of transfers is not enough to provide the public good, then every contributing agent receives their transfer back.

Each agent n's strategy is made of two parts: a message $m_n = \sigma_n(v_n) \in M_n$ for each type $v_n \in V_n$ and a transfer $t_n = \gamma_n(v_n, m) \in \mathbb{R}_+$ for each type $v_n \in V_n$ and observed message vector m. Let $\sigma = (\sigma_1, ..., \sigma_N)$ and $\gamma = (\gamma_1, ..., \gamma_N)$, and let the set of all perfect Bayesian equilibria (PBE) (σ, γ) of the extended VCM be denoted by Γ^{VCM}.

Definition 2 *An allocation (ρ, τ) is implementable by the extended VCM if there is a profile $(\sigma, \gamma) \in \Gamma^{VCM}$ such that, for all $v \in V$ and for all n,*

$$\gamma_n(v_n, \sigma(v)) = \tau_n(v)$$

and

$$\mathbf{1}\left\{\sum_{n=1}^{N} \gamma_n(v_n, \sigma(v)) \geq c\right\} = \rho(v).$$

Let all allocations that are implementable by the extended VCM be denoted by Ψ^{VCM}.[7]

In general, as is standard with cheap talk games, there are many equilibria, so that set Ψ^{VCM} can be quite large. In particular, a babbling equilibrium always exists, where agents simply forego the opportunity to talk in the first period and act as if the mechanism was the simple VCM. But, as the following example illustrates, cheap talk allows for other equilibria.

Example: Assume that $N = 2$, $v_n \in [0, 1]$ for $n = 1, 2$ and $c = \frac{3}{2}$. Consider the following strategy profile (σ^*, γ^*) with $M_n = \{L, H\}$, where

$$\sigma_n^*(v_n)(m_n) = \begin{cases} 1 \text{ if } \left\{m_n = L \text{ and } v_n < \frac{3}{4}\right\} \text{ or if } \left\{m_n = H \text{ and } v_n \geq \frac{3}{4}\right\} \\ 0 \text{ otherwise} \end{cases}$$

[7]Notice that, because I only consider deterministic allocations, there is no loss of generality in not allowing the agents to randomize.

for $n = 1, 2$, and

$$\gamma_n^* (v_n, m) = \begin{cases} \frac{3}{4} \text{ if } m = (H_1, H_2) \text{ and } v_n \geq \frac{3}{4} \\ 0 \text{ otherwise} \end{cases}$$

for all $m \in \{L_1, H_1\} \times \{L_2, H_2\}$, $v_n \in [0, 1]$ and $n = 1, 2$. In words, each agent n reports $m_n = H$ if his type $v_n \geq \frac{3}{4}$ and reports $m_n = L$ otherwise. Then, each agent n makes a transfer of $t_n = \frac{3}{4}$ if both reports are H (so that the public good is provided) but makes no contribution otherwise. Notice that (σ^*, γ^*) is a PBE of the extended VCM and implements (ex-post IR) allocation (ρ^*, τ^*), where

$$\rho^* (v) = \begin{cases} 1 \text{ if } v_1 \geq \frac{3}{4} \text{ and } v_2 \geq \frac{3}{4} \\ 0 \text{ otherwise} \end{cases}$$

and

$$\tau^* (v) = \begin{cases} \left(\frac{3}{4}, \frac{3}{4}\right) \text{ if } v_1 \geq \frac{3}{4} \text{ and } v_2 \geq \frac{3}{4} \\ (0, 0) \text{ otherwise} \end{cases},$$

which would not be implemented had the agents not been allowed to cheap talk.

2.3 Benevolent Dictator System

The BD system is very similar to the extended VCM except that, in the second stage, it is the BD who decides what transfer each agent makes, i.e., in the first stage, each agent sends a public message $m_n \in M_n$; in the second stage, the BD chooses whether to provide the public good - $p(m) \in \{0, 1\}$ - and demands a transfer of $t_n(m) \in \mathbb{R}$ of each agent n. Mapping (p, t) can be thought of as the mechanism the BD imposes. It is *feasible* if

$$p(m) = 1 \Rightarrow \sum_{n=1}^{N} t_n(m) \geq c$$

for all $m \in M$.

A strategy for each agent is a choice of a message $m_n = \sigma_n(v_n) \in M_n$ for each type $v_n \in V_n$. Strategy profile σ is a *Bayes-Nash equilibrium* (BNE) of the game induced by mechanism (p, t) if and only if, for all n,

$$
\begin{aligned}
& v_n E_{v_{-n}}\left(p\left(\sigma\left(v_n, v_{-n}\right)\right)\right) - E_{v_{-n}}\left(t_n\left(\sigma\left(v_n, v_{-n}\right)\right)\right) \\
\geq \; & v_n E_{v_{-n}}\left(p\left(\sigma\left(v_n', v_{-n}\right)\right)\right) - E_{v_{-n}}\left(t_n\left(\sigma\left(v_n', v_{-n}\right)\right)\right)
\end{aligned}
$$

for all $v \in V$ and $v_n' \in V_n$.

Lemma 1 *Consider any profile (σ, p, t) for which σ is a BNE of the game induced by mechanism (p, t) and assume that there is some agent n and two types $v_n' \in V_n$ and $v_n'' \in V_n$ such that*

$$E_{v_{-n}}\left(p\left(\sigma\left(v_n', v_{-n}\right)\right)\right) = E_{v_{-n}}\left(p\left(\sigma\left(v_n'', v_{-n}\right)\right)\right) = k_0.$$

for some $k_0 \in [0, 1]$. It follows that, for all $v_n \in [v_n', v_n'']$,

$$E_{v_{-n}}\left(p\left(\sigma\left(v_n, v_{-n}\right)\right)\right) = k_0 \text{ and } E_{v_{-n}}\left(t_n\left(\sigma\left(v_n, v_{-n}\right)\right)\right) = k_1$$

for some $k_1 \in \mathbb{R}$.

Lemma 1 implies that if there are two types v_n' and v_n'' who send two messages that are (interim) equivalent (in the sense that they lead to the same probability that the public good is provided and the same expected transfer), any type in between those two types also sends an (interim) equivalent message. Put differently, lower types choose messages that lead to a smaller probability that the public good is provided but also to a smaller expected transfer, while larger types choose messages that lead to a larger probability that the public good is provided but also a larger transfer. Larger types prefer these messages as they are the ones who value the public good the most. Figure 1 displays a simple 3 message example. By lemma 1, it has to be that message $m_n = A$ has the lowest interim probability that the public good is provided and the lowest expected transfer, followed by message $m_n = B$ and then by message $m_n = C$.

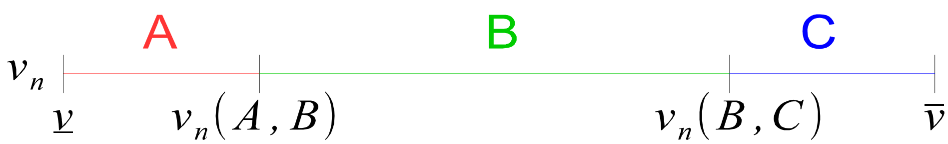

Figure 1: Only three messages are sent in equilibrium: $M_n = \{A, B, C\}$. Types in the red line report $m_n = A$, types in the green line report $m_n = B$, types in the blue line report $m_n = C$.

The typical mechanism design approach, which assumes that the BD can commit, is to choose from all feasible profiles (σ, p, t) for which σ is a BNE of the game induced by mechanism (p, t) the one that maximizes some objective function (presumably some measure of welfare). In order to model the BD's limited commitment power, I limit further the set of profiles (σ, p, t) that the BD can select. In particular, I assume that the BD can only choose profiles (σ, p, t) that have the two properties described below: the "posterior Pareto efficiency" (PPE) property and the "monotonicity of transfers" (MT) property. As I show in section 5, both assumptions are necessary for the result of section 3 to hold in the sense that if only one of them holds, one can find examples where the result fails.

Before describing the two properties, it is convenient to introduce some additional notation. For all $m \in M$, let $\pi^{\sigma}(m) = (\pi_1^{\sigma}(m_1), ..., \pi_N^{\sigma}(m_N))$ denote the public beliefs generated after message m has been sent and given strategy profile σ. More rigorously,

$$\pi_n^{\sigma}(m_n)(v_n) = \frac{f_n(v_n)\,\mathbf{1}\{\sigma_n(v_n) = m_n\}}{\int\limits_{\underline{v}_n}^{\overline{v}_n} f_n(v_n)\,\mathbf{1}\{\sigma_n(v_n) = m_n\}\,dv_n}$$

for all m_n such that

$$\int\limits_{\underline{v}_n}^{\overline{v}_n} f_n(v_n)\,\mathbf{1}\{\sigma_n(v_n) = m_n\}\,dv_n > 0.$$

Definition 3 *Profile (σ, p, t) has the posterior Pareto efficiency (PPE) property if and only if, for all $m \in M$,*

$$\sum_{n=1}^{N} \inf\left[supp\left(\pi_n^{\sigma}(m_n)\right)\right] \geq c \Rightarrow p(m) = 1$$

and

$$\sum_{n=1}^{N} t_n(m) = \begin{cases} c \text{ if } p(m) = 1 \\ 0 \text{ otherwise} \end{cases}.$$

Notice that $\inf\left[supp\left(\pi_n^{\sigma}(m_n)\right)\right]$ represents the lowest valuation agent n might have, given the public beliefs following message m_n. Therefore, the BD knows that, if she was to demand a transfer of $\inf\left[supp\left(\pi_n^{\sigma}(m_n)\right)\right]$ or lower from agent n, his utility would certainly be positive, as long as the public good is provided. The first part of the PPE property states that, if it is possible for the BD to fund the provision of the public good while guaranteeing that no agent makes a contribution that is larger than their valuation, the public good must be provided.

The second part of the PPE property imposes that there be no wasted transfers; if the public good is provided, then the sum of the contributions must not exceed the total cost, while if the good is not provided, no transfers are requested.

In words, the PPE property simply states the following: if the BD *knows* that she can make everyone better off by taking a certain action, she must take that action. Interpreted in this way, I would argue that the PPE property is a minimum condition to be able to refer to the dictator as benevolent.

Definition 4 *Profile (σ, p, t) has the monotonicity of transfers (MT) property if and only if for all $m_n' \in M_n$ and $m_n'' \in M_n$, either*

$$p(m_n', m_{-n}) = p(m_n'', m_{-n}) = 1 \Rightarrow t_n(m_n', m_{-n}) \geq t_n(m_n'', m_{-n})$$

for all m_{-n} or

$$p(m_n', m_{-n}) = p(m_n'', m_{-n}) = 1 \Rightarrow t_n(m_n', m_{-n}) \leq t_n(m_n'', m_{-n})$$

for all m_{-n}.

The MT property states that if one message - say m_n' - leads to a larger transfer than some other message - say m_n'' - for some set of reports of the other agents (m_{-n}) for which the public good is provided, then it leads to a larger transfer for any such reports. In other words, if one was to order agent n's messages according to the transfers he is asked to make, that order would be independent of the other agents' reports. The motivation for this property is to make the BD concerned with "fairness". Because the agents' types are independent, the public beliefs about agent n's type only depend on agent n's report. What the MT property imposes is that the *order* of transfers depends only on those beliefs. In particular, if the BD has "fairness" concerns, the BD ranks the messages in terms of how much the agent is believed to value the public good and then charges a larger transfer if the agent is believed to value the public good a lot.

Figure 1 is helpful in interpreting the MT property; a BD who is concerned with being fair would choose

$$t_n(C, m_{-n}) \geq t_n(B, m_{-n}) \geq t_n(A, m_{-n})$$

for all m_{-n} (which would satisfy the MT property), because she would infer that the agent values the public good the most if he sends message $m_n = C$, then $m_n = B$ and then $m_n = A$.

Let Γ^{BD} denote the set of all profiles (σ, p, t) that have the PPE and MT properties and are such that σ is a BNE of the game induced by feasible mechanism (p, t).

Definition 5 *An allocation (ρ, τ) is implementable by the BD system if there is a profile $(\sigma, p, t) \in \Gamma^{BD}$ such that, for all $v \in V$ and for all n,*

$$t_n(\sigma(v)) = \tau_n(v)$$

and

$$p(\sigma(v)) = \rho(v).$$

Let all allocations that are implementable by the BD system be denoted by Ψ^{BD}.

3 Main result

The main result of the paper is that, if the public good is binary, every allocation that is implementable by the BD system and is ex-post individually rational is also implementable by the extended VCM.

Theorem 1 $\Psi^{BD} \cap \Psi^{IR} \subseteq \Psi^{VCM}$.

In what follows, I provide an explanation for the result (the complete proof can be found in the appendix).

Proof (Sketch). Consider some allocation $(\rho, \tau) \in \Psi^{BD} \cap \Psi^{IR}$ and let $(\sigma, p, t) \in \Gamma^{BD}$ represent the profile that implements allocation (ρ, τ) in the BD system. By definition, (σ, p, t) has the PPE and MT properties and σ is a BNE of the game induced by mechanism (p, t). For expositional reasons, let us say that σ_n is the one represented in figure 1 of the previous section. I show that, in the extended VCM, agent n would like to report as he did in the BD system and then select the same transfers that the BD would have selected.

One possible deviation for agent n would be to deviate on his report but then not deviate on the ensuing transfer. However, these type of deviations can be immediately ruled out because, if such a deviation was worthwhile in the extended VCM it would have also been worthwhile in the BD system. So, the only deviations that are worth considering are those that (also) involve a deviation on the transfer choice. In fact, a deviation on the report is more appealing in the extended VCM because they can be followed by deviations on the transfers, i.e., agent n might find it worthwhile to deviate on the report because he anticipates he will deviate on the transfer as well. There are four parts to arguing that these deviations will not be strictly beneficial.

1) *Agents do not deviate on their transfer on the path of play.* More precisely, if some agent n of type v_n reports the same message as he would have reported in the BD system (for example, a green type v_n reporting B), he finds it optimal to choose the transfer the BD would have chosen for him. To see why that is, notice that, in a BD system, only one of two things can happen. The first possibility is that the reports that are made are such that the BD would have chosen not to provide the public good, in which case, she would have requested no transfers from any of the agents. Clearly, in this case, choosing to make a positive transfer would not be desirable for

agent n (because $\overline{v}_n < c$). The second possibility is that, after observing the agents' reports, the BD would have chosen to provide the public good and requested a total sum of transfers equal to c. In this case, a deviation would still not be desirable as an increase in the transfer would not increase the public good level, and a reduction would lead to the public good not being provided. As a result, the largest expected utility the agent could obtain from deviating would be 0, which is (weakly) worse than not deviating (because the allocation that gets implemented is ex-post IR). In this part of the argument, it is crucial that the public good is binary. Essentially, the agent does not deviate in the second stage from what the BD would have done, because the classical free riding problem is assumed away.

2) *Agents do not deviate "down".* As discussed above, deviating on the report in extended VCM is, in general, more appealing for the agents than in the BD system precisely because they can always deviate on their transfer as well. However, the only time that an agent would like to deviate on their transfer is when that transfer is larger than their valuation, which never happens on the path of play by 1). It also never happens if the agent reports "down", i.e., if the agent chooses a message that, in equilibrium, is chosen by a lower type. For example, imagine that some green type v_n was to deviate to reporting A. The largest transfer he would be asked, if he was in the BD system, would not be larger than $v_n(A, B)$, because the allocation that gets implemented is ex-post IR. Seeing as the agent's type $v_n \geq v_n(A, B)$, he would never deviate on the transfer, so that a deviation to A in the extended VCM is equal to a deviation to A in the BD system, and, therefore, not strictly beneficial.

3) *Deviating to the next "largest" message is the best deviation "up".* Consider some red type v_n. He has two possible deviations: reporting $m_n = B$ or reporting $m_n = C$. I argue that deviating to $m_n = B$ is better than deviating to $m_n = C$. Notice that the PPE property and the fact that the allocation is ex-post IR imply

that $p(C, m_{-n}) \geq p(B, m_{-n})$ for all m_{-n}. Furthermore, the MT property implies that $t_n(C, m_{-n}) \geq t_n(B, m_{-n})$ for any $m_{-n} \in M_{-n}$. Therefore, if $m_n = C$ is more appealing than $m_n = B$ for type v_n there must be some set of reports $m'_{-n} \in M_{-n}$ for which the public good is not provided with message $m_n = B$ but is provided with message $m_n = C$, i.e., $p\left(B, m'_{-n}\right) = 0$ and $p\left(C, m'_{-n}\right) = 1$. Consider any such set of reports $m'_{-n} \in M_{-n}$. Because of the PPE property and ex-post individually rationality, it follows that the reason why $p\left(B, m'_{-n}\right) = 0$ is because it is not possible for the BD to fund the public good while being certain that every agent has a positive utility. But if $m_n = C$, the BD would be certain, which is why the public good is provided. As a result, it must be that $t_n(C, m_{-n}) > v_n(A, B) \geq v_n$; whenever the public good provision goes from 0 to 1 as a consequence of the agent reporting $m_n = C$ rather than $m_n = B$, the transfer that the agent would be expected to make would be larger than his valuation. As a result, there is no reason for type v_n to prefer C over B.

4) *Deviating to the next "largest" message is not strictly preferred.* In figure 1, this means that no red type v_n wants to report $m_n = B$. This step follows by continuity of the agent's valuation. Notice that type $\underline{v}_n$ does not want to report $m_n = B$ by 3), while type $v_n(A, B)$ is indifferent between A and B. Imagine that the probability that the public good is provided in the extended VCM should type v_n report $m_n = B$ is larger than if he was to report $m_n = A$ and suppose that type v_n strictly prefers to report $m_n = B$. In that case, type $v_n(A, B)$, who values the provision of the public good even more, would also strictly prefer to report $m_n = B$, which we know he does not. Likewise, if that probability was smaller, type $\underline{v}_n$, who values the public good less than type v_n, would also strictly prefer to report $m_n = B$, which, again, we know he does not. So, it must be that type v_n does not strictly prefer to report $m_n = B$. ■

4 Discussion

4.1 Are the two systems equivalent?

Proposition 1 states that the extended VCM can be used to implement all the allocations that are implementable by the BD system, provided they are ex-post IR. The question that remains is whether the extended VCM can implement allocations that are *not* implementable by the BD system. I argue that, not only is the answer "yes", but the "extra" allocations might actually be quite appealing.

The difference between the BD system and the extended VCM is the freedom that exists in the latter system for agents to choose their own transfers. On the one hand, this is a bad thing, as it gives the agents more opportunities to deviate from the behavior the BD would like to impose. However, the whole exercise of the previous section was exactly to show that the freedom that agents have to choose their own transfers in the extended VCM does not prevent any of the allocations which are implementable by a BD to be implementable also by the extended VCM, provided they are ex-post IR. Basically, when the public good is binary, the agents are never in a situation where they would want to choose a different transfer than the one that the BD would have chosen.

However, the fact that agents have the freedom to choose their own transfers also has a positive side to it in that what happens in the second stage becomes less restricted. In the BD system, if it is known that it is possible to provide the public good while ensuring that all agents have a positive ex-post utility, then the good must be provided by the BD. However, in the extended VCM, it does not, as there is always a Bayes-Nash equilibrium of the second stage voluntary contribution subgame where no agent provides a positive transfer. This is why there are some (possibly appealing) allocations that can only be implemented by the extended VCM. To illustrate I consider the following example from Agastya et al. (2007).

Example: There are only two agents, $v_n \sim U(\underline{v}, \overline{v})$ for $n = 1, 2$ and $\overline{v} + \underline{v} \leq c$. Let Ω denote the set of all allocations (ρ, τ) which are incentive compatible, interim IR, and ex-ante budget balanced. Notice that, if an allocation is ex-post IR, it is also interim IR; and if an allocation is ex-post budget balanced, it is also ex-ante budget balanced. So, it follows that $\Psi^{BD} \cap \Psi^{IR} \subseteq \Omega$.

For all $\lambda \in [0, 1]$, let $\omega(\lambda) \in \Omega$ be such that

$$\omega(\lambda) = \arg \max_{(\rho,\tau) \in \Omega} \lambda E(u_1 | \rho, \tau) + (1 - \lambda) E(u_2 | \rho, \tau)$$

where $E(u_n | \rho, \tau)$ represents the expected utility of agent n, given allocation (ρ, τ), for $n = 1, 2$. In words, λ represents the weight one gives to each agent and $\omega(\lambda)$ represents the set of allocations that maximize the weighted sum of utilities. Notice that, for an allocation to be efficient within set Ω, it must belong to some $\omega(\lambda)$, for some λ.

In Agastya et al. (2007), the authors first characterize, for all $\lambda \in [0, 1]$, set $\omega(\lambda)$. In particular, they show that, if $(\rho, \tau) \in \omega(\lambda)$, then,

$$\rho(v) = \rho^\lambda(v) \equiv \begin{cases} 1 \text{ if } v_2 \geq m^\lambda v_1 + b^\lambda \\ 0 \text{ otherwise} \end{cases}$$

for all $v \in [\underline{v}, \overline{v}]^2$, where $m^\lambda < 0$ and $b^\lambda \in \mathbb{R}$. Given this characterization, the authors go on to show that, for each $\lambda \in [0, 1]$, there is an allocation $(\rho^\lambda, \tau^\lambda) \in \omega(\lambda)$ that can be implemented by the extended VCM.

Proposition 2 *For any $\lambda \in [0, 1]$, there is no allocation $(\rho, \tau) \in \omega(\lambda)$ that is implementable in the BD system.*

What proposition 2 shows is that, while some of the "best" allocations of set Ω, a set that is larger than $\Psi^{BD} \cap \Psi^{IR}$, can be implemented by the extended VCM, none of the "best" allocations can be implemented by the BD system.

Figure 2 is helpful in understanding the argument. Notice that $\rho^{\lambda}(v) = 1$ if and only if v belongs to the space above the red line. The argument (which is detailed in the appendix) has two parts. First, because the red line is negatively sloped, it follows that each agent $n = 1, 2$ must report truthfully in interval $[\widehat{v}_n, \overline{v}]$ in order for allocation $(\rho, \tau) \in \omega(\lambda)$ to be implementable by the BD. Second, whenever the agents report truthfully, the BD is "forced" to provide to provide good whenever $v_1 + v_2 \geq c$ by the PPE property, i.e., the BD must provide the public good whenever v lies in the space above the green line in figure 2. The contradiction is that the green and red lines do not coincide, because $m^{\lambda} \neq -1$ and $b^{\lambda} \neq c$.

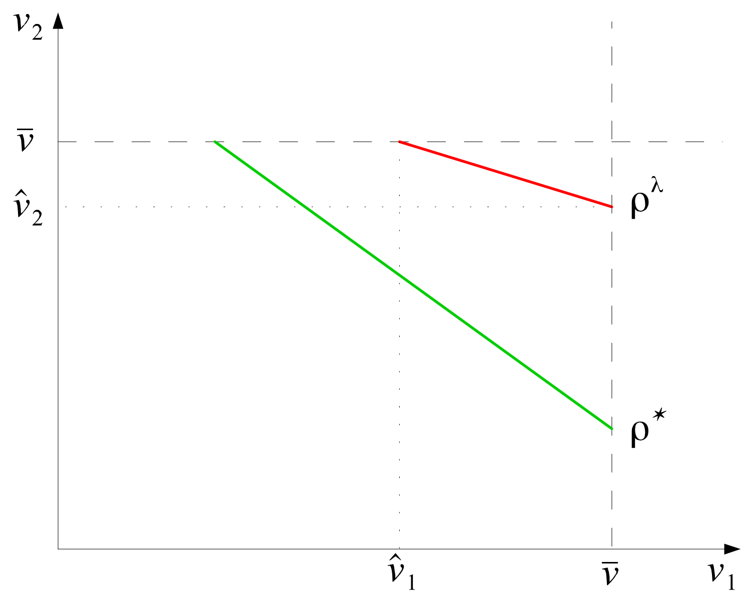

Figure 2: Comparison between ρ^{λ} and ρ^{*}, where $\rho^{*}(v) = 1$ if and only if $v_1 + v_2 \geq c$.

4.2 Non-binary Public Goods

As discussed in the introduction, one of the advantages of centralized systems is that they are able to more effectively deal with the "classical" free-riding problem. By assuming that the public good is binary, I am essentially "shuting-down" that advantage of centralized systems. I do this purposefully to focus exclusively on the problem of eliciting the agents' private information.

If the public good is not binary, the classical free-riding problem re-emerges and there is a trade-off between the two systems. To see this more clearly, let us assume that $g \in \left\{0, \frac{1}{k}, \frac{2}{k}, ..., \frac{k}{k}\right\}$ for some $k \geq 1$ and that the cost of providing g units of the public good is given by cg. Accordingly, an allocation $(\rho, \tau) : V \rightarrow \left\{0, \frac{1}{k}, \frac{2}{k}, ..., \frac{k}{k}\right\} \times \mathbb{R}_{+}^{N}$ maps the agents' types to units of the public good to be provided and to a transfer vector.

Consider the BD system. Notice that both properties - PPE and MT - still continue to make sense in this context. In particular, in what concerns the PPE property, if the BD knows that the sum of the agents' valuations is larger than c, she must provide $g = 1$ units of the public good, because anything less would be Pareto dominated. Therefore, the set of allocations (ρ, τ) that are implementable by a BD system and are ex-post IR are the same regardless of k.

However, in the extended VCM, things are different, because, if $k > 1$, the classical free-riding problem arises. In particular, the maximum transfer each agent n makes in the extended VCM is $\frac{1}{k}\overline{v}_n$. To see why that is, imagine that, after some message $m \in M$, there is some agent n who makes a transfer of $\frac{1}{k}\overline{v}_n + \varepsilon$ for some $\varepsilon > 0$ and that this leads to the provision of $\frac{j}{k}$ units of the public good, for some $j > 0$. If agent n was to reduce his transfer by $\left(\frac{1}{k}\overline{v}_n + \delta\right)$, where

$$\delta \in \left(0, \min\left\{\frac{1}{k}\left(c - \overline{v}_n\right), \varepsilon\right\}\right),$$

the units of the public good that would be provided would be $\frac{j-1}{k}$. Therefore, the benefit for the agent of this deviation would be

$$\frac{1}{k}\left(\overline{v}_n - v_n\right) + \delta > 0$$

Given that each agent never transfers more than $\frac{1}{k}\overline{v}_n$, it follows that the maximum amount of units of the public good that can be provided by the extended VCM is $\frac{N}{k}\overline{v}$, where $\overline{v} \equiv \max_n \overline{v}_n$, which becomes smaller and smaller as k increases, i.e., the extended VCM becomes increasingly unappealing.[8]

In a way, parameter k represents how important the classical free-riding problem is relative to the underlying informational problem: if k is small (large), the classical free-riding problem is less (more) important than the informational free-riding problem.

4.3 Explicitly modelling the BD

When designing the BD system, I followed a "reduced form" approach in that I did not directly model the BD. Instead, I have tried to be as least restrictive as possible so that any allocation that could conceivably be thought of as being implementable by some centralized system would be in set Ψ^{BD}.

A different approach would have been to endow the BD with preferences and make her a player in the cheap talk game. While my approach is more general, directly endowing the BD with preferences that she maximizes is likely to be more familiar to the reader. To be more specific, consider the following game: first, each one of the N agents simultaneously choose a public message; second, the BD observes the message vector m and forms beliefs $\pi^{\sigma}(m) \in \Delta(V)$ about the N agents' types; based on those beliefs, the BD chooses $p(m)$ and $t(m)$ *in order to maximize her own utility function.* Therefore, for any set of beliefs $\pi \in \Delta(V)$, the BD maximizes some function W_{π}. A

[8] In a similar context, Barbieri (2012) also highlights some of the problems that the voluntary provision of public goods with pre-play communication may have with non-binary goods.

reasonable and relevant question is what type of utility functions W_π generate a set of PBE (σ, p, t) of this cheap talk game that have the two properties that were assumed in the text: the PPE and the MT properties. In this section, I provide three examples of such functions.

Example A: *The BD is risk neutral and inequality averse.*

For any belief distribution $\pi \in \Delta(V)$, the BD chooses $x \in \{0, 1\}$ and $y \in [0, c]^N$ to maximize

$$W_\pi(x, y) = E\left[\left(\prod_{n=1}^{N}(v_n x - y_n + c)\right)|\pi\right].$$

Because the agents' types are independent, it follows that

$$W_\pi(x, y) = \prod_{n=1}^{N}\left(E(v_n|\pi_n)\, x - y_n + c\right).$$

The BD is risk neutral because she only cares about the expected utility of each agent and is inequality averse because she prefers to equate all of the agents' expected utilities. The assumption of risk neutrality seems particularly fitting because the agents are also assumed to be risk neutral.

Proposition 3 *If the BD's preferences are as described in Example A, then any PBE (σ, p, t) has the PPE and MT properties.*

Proof. The PPE property follows directly from function $W_\pi(x, y)$ being strictly decreasing with each y_n for any $\pi \in \Delta(V)$ and $x \in \{0, 1\}$ and from the fact that $E(v_n|\pi_n) \geq \inf[supp(\pi_n)]$. The MT property is shown in the appendix. ■

Example B: *The BD is risk neutral and inequality averse (but not an expected utility maximizer).*

For any belief distribution $\pi \in \Delta(V)$, the BD chooses $x \in \{0, 1\}$ and $y \in [0, c]^N$ to maximize

$$W_{\pi}(x,y) = \sum_{n=1}^{N} w_n \left(E\left(v_n | \pi_n\right) x - y_n \right),$$

where each $w_n : \mathbb{R} \to \mathbb{R}$ is a function of the expected utility of the agent. Because the belief π only enters the BD's utility through the implied expected utilities of the agents, the BD is risk neutral. In order to assume that the BD is inequality averse, i.e., in order to assume that she prefers those who value the public good the most to pay for it the most, I assume that each w_n function is strictly concave.

Proposition 4 *If the BD's preferences are as described in Example B and each w_n is such that $w_n'(u) > 0$ and $w_n''(u) < 0$ for all $u \in \mathbb{R}$, then any PBE (σ, p, t) has the PPE and the MT properties.*

Proof. The PPE property follows directly from $w_n'(u) > 0$ for all $u \in \mathbb{R}$. The MT property is shown in the appendix. ■

Example C: *The BD is risk averse and inequality averse.*

For any belief distribution $\pi \in \Delta(V)$, the BD chooses $x \in \{0,1\}$ and $y \in [0,c]^N$ to maximize

$$W_{\pi}(x,y) = E\left[\left(\sum_{n=1}^{N} w_n \left(v_n x - y_n \right) \right) | \pi \right],$$

where $w_n : \mathbb{R} \to \mathbb{R}$ represents the ex-post utility of the BD relative to agent n, which only depends on agent n's ex-post utility. By assuming that each w_n is strictly concave, one is assuming that the BD is risk averse and inequality averse.

Proposition 5 *If the BD's preferences are as described in Example C and each w_n is such that $w_n'(u) > 0$, $w_n''(u) < 0$ and $w_n'''(u) = 0$ for all $u \in \mathbb{R}$, then any PBE (σ, p, t) has the PPE and the MT properties.*

Proof. The PPE follows directly from $w_n'(u) > 0$ for all $u \in \mathbb{R}$. The proof for the MT property is on the appendix. ■

These examples confirm what is stated in the introduction and in the text: that the PPE property only depends on the BD being "Paretian", i.e., benefiting whenever the agents are made better off; and that the MT property relies on the assumption that the BD has fairness/inequality concerns.

4.4 How tight are the PPE and MT properties?

In this section, I provide two examples which show that if either property is dropped, it is possible to find ex-post IR allocations that could only be implemented by the BD system.

Example 1 (when the MT property fails) *Suppose there are* 2 *agents with* $v_n \in [0.3, 0.9]$ *for* $n = 1, 2$ *and* $c = 1$. *Consider the following strategy profile: agent* n *reports* L_n *if* $v_n < x_n$, M_n *if* $x_n \leq v_n < y_n$ *and* H_n *if* $v_n \geq y_n$, *where* $x_1 = 0.6$, $y_1 = 0.75$, $x_2 = 0.5$ *and* $y_2 = 0.8$. *The agents' reports are mapped according to* (p, t), *where* p *is given by the following table:*

	H_2	M_2	L_2
H_1	1	1	1
M_1	1	1	0
L_1	1	0	0

while t *is given by*

	t_1	t_2
H_1H_2	0.75	0.25
H_1M_2	0.7	0.3
H_1L_2	0.7	0.3
M_1H_2	0.4	0.6
M_1M_2	0.5	0.5
L_1H_2	0.3	0.7

Notice that the PPE property is satisfied, because, for every message vector m for which it is known that $v_1 + v_2 \geq c$, the public good is being provided. But, the MT property is not, because $t_2(H_1, H_2) < t_2(H_1, M_2)$ but $t_2(M_1, H_2) > t_2(M_1, M_2)$. It is also easy to confirm that the allocation is ex-post IR and forms an equilibrium of the BD system (i.e., the agents want to play as described) provided

$$\Pr\{v_1 \geq y_1\} = \Pr\{v_1 \in [x_1, y_1)\} = \frac{2}{5}$$

and

$$\Pr\{v_2 \geq y_2\} = \Pr\{v_1 \in [x_2, y_2)\} = \frac{1}{13}.$$

However, the corresponding allocation cannot be implemented by the extended VCM, because if agent 2's type is $v_2 = x_2 = 0.5$, a deviation to H_2 would return a payoff of

$$\frac{2}{5} * 0.25 > \frac{2}{5} * 0.2$$

where the right hand side is the agent's payoff if he reports M_2.

Example 2 (when the PPE property fails) *Suppose there are three agents, that $v_1 \in [0.7, 2]$, $v_2 \in [0.9, 2]$, $v_3 \in [0.2, 2]$ and $c = 3$. Consider the following reporting profile of a BD system: agent 1 reports L_1 if $v_1 \leq 1.1286$, M_1 if $v_1 \in (1.1286, 1.3176]$ and H_1 if $v_1 > 1.3176$; agent 2 sends message L_2 if $v_2 \leq 1.0667$, M_2 if $v_2 \in (1.0667, 1.5105]$ and H_2 if $v_2 \geq 1.5105$; and agent 3 sends message L_3 if $v_3 \leq 1.314$ and H_3 if $v_3 > 1.314$.*

Function p is given by

$m_3 = H_3$	H_2	M_2	L_2
H_1	1	1	1
M_1	1	1	0
L_1	1	0	0

and

$m_3 = L_3$	H_2	M_2	L_2
H_1	1	0	0
M_1	0	0	0
L_1	0	0	0

Function t_1 is given by

$m_3 = H_3$	H_2	M_2	L_2		$m_3 = L_3$	H_2	M_2	L_2
H_1	0.9	1.3	0.8	*and*	H_1	1.3	0	0
M_1	0.9	0.9	0		M_1	0	0	0
L_1	0.7	0	0		L_1	0	0	0

Function t_2 is given by

$m_3 = H_3$	H_2	M_2	L_2		$m_3 = L_3$	H_2	M_2	L_2
H_1	1.1	1	0.9	*and*	H_1	1.5	0	0
M_1	0.9	0.9	0		M_1	0	0	0
L_1	1.3	0	0		L_1	0	0	0

Function t_3 is given by

$m_3 = H_3$	H_2	M_2	L_2		$m_3 = L_3$	H_2	M_2	L_2
H_1	1	0.7	1.3	*and*	H_1	0.2	0	0
M_1	1.2	1.2	0		M_1	0	0	0
L_1	1	0	0		L_1	0	0	0

Notice that the MT property is satisfied but the PPE property is not (for example, if $m = (L_1, M_2, H_3)$, the BD would be able to fund the public good while making everyone better off and yet $p((L_1, M_2, H_3)) = 0$). Furthermore, the agents' reporting profile is an equilibrium in the BD system if

$$\Pr\{v_1 \leq 1.1286\} = 0.2,\ \Pr\{v_1 \in (1.1286, 1.3176]\} = 0.3,\ \Pr\{v_1 > 1.3176\} = 0.5$$

$$\Pr\{v_2 \leq 1.0667\} = 0.25,\ \Pr\{v_2 \in (1.1286, 1.5105]\} = 0.35,\ \Pr\{v_2 > 1.5105\} = 0.4$$

and

$$\Pr\{v_3 \leq 1.314\} = 0.6,\ \Pr\{v_3 > 1.314\} = 0.4$$

and the allocation which is implemented is ex-post IR.

However, in the extended VCM, agent 1 *has an incentive to deviate on his report if he is of type* $v_1 = 1.1286$ *(his type is such that, in a BD system, he would be indifferent between reporting* L_1 *and* M_1*). By reporting* L_1 *(by not deviating), the agent receives an expected payoff of*

$$0.4 * 0.4 * (1.1286 - 0.7) = 0.0686$$

while, by deviating to H_1*, the agent receives an expected payoff of*

$$0.4 * 0.4 * (1.1286 - 0.9) + 0.25 * 0.4 * (1.1286 - 0.8) = 0.0694$$

because, if the realized messages are either (H_1, M_2, H_3) *or* (H_1, H_2, L_3)*, the agent chooses a transfer that is different than the one the BD would have chosen for him.*

4.5 Multiplicity of Equilibria

In the extended VCM, there are, in general, multiple equilibria, which is why Ψ^{VCM} is large. So large that it is larger than $\Psi^{BD} \cap \Psi^{IR}$. So far, I have interpreted Ψ^{VCM} being large as a good thing: everything that can be sustained by a BD system can also be sustained by the extended VCM. But, one might argue that this multiplicity is a bad thing: why would the players play one equilibrium instead of another? In fact, one could even make the case that, in centralized systems, this multiplicity will not exist, seeing as, presumably, the BD would just select whatever allocation she prefers. This type of argument can be the basis of another benefit for centralization which I have not considered in the model: that it limits the multiplicity of equilibria. However, I believe it is important to qualify this statement further.

There are two sources of multiplicity in the extended VCM. First, there is multiplicity on the second stage choices of the agents. To be more concrete, imagine that the first stage has already occurred and that agents have formed beliefs about each other's types. At that point, there are, in general, many equilibria of the ensuing subgame. For example, there is always an equilibrium where no agent makes a contribution. And, if there is an equilibrium where the public good is provided, there are, in general, many others where the only thing that changes is who pays for the public good the most.

A second type of multiplicity comes from the cheap talk of the first stage. And this multiplicity is the typical multiplicity one associates with cheap talk models.[9] In particular, as in any cheap talk model, there is always a babbling equilibrium where agents essentially forego the opportunity to communicate in the first period.

While the second type of multiplicity is exclusive of the extended VCM, I would argue that the first one is shared by the BD system as well. In the BD system, the agents are still the ones who choose how to report, not the BD. In particular, there is always a babbling equilibrium where the agents do not share any information just like in the extended VCM. So, set Ψ^{BD} is, in general, not made of a single allocation and can have allocations that are "bad", like those that are implemented by babbling equilibria.

5 Conclusion

The issue of who should provide public goods is naturally far more complex than what it is made out to be in my model. In fact, my intention when building the model was not necessarily to make it as realistic as possible but rather to make it so that my

[9] See Crawford and Sobel (1982).

argument was as clear as possible to the reader. Therefore, the result of this paper should not be interpreted as proving the superiority of decentralized systems. Instead, it contributes to identifying what are the possible trade-offs between the two systems.

In my model, it is clear that the extended VCM is better. But, as I mentioned in the introduction, I have purposely eliminated some of the benefits of centralized provision. One clear example is the assumption that the public good is binary. As I show in section 4.2, if one considers non-binary public goods, the comparison is less clear and it might be that either system is preferred. But, even though binary public goods are not common, the analysis of the binary public good case is important precisely because it identifies that there are two different forces at play: the classical free-riding that is better dealt with by centralized systems and the informational free-riding that is better dealt with decentralized systems.

Another potential benefit of centralized systems that has (purposefully) been shut down from the analysis is the issue of coordination. In a decentralized system, there are typically many equilibria so it might be that agents coordinate on a bad equilibrium. By contrast, in a centralized system, the central figure presumably has more control. Things are not as straightforward when the agents have private information, as I discuss in section 4.5., but, regardless, a point could still be made that centralized systems are better because they are less risky in this sense. That, however, does not negate the relevance of this paper's result - it still presents one more dimension to add to the trade-off between the two systems.

What is key to the argument is the set of assumptions about how the BD behaves, in particular, the PPE and MT properties. As I show in section 4.4, none of the assumptions can be eliminated for otherwise examples can be found of ex-post IR allocations that could be implemented by the BD system that would not be implementable by the extended VCM. I argue in the text and further in section 4.3. that these assumptions follow naturally from having a BD who has limited commitment power, is risk neutral (just like the agents) and is inequality averse. While these assumptions seem to me to be quite natural, the fact that none can be disregarded can also be interpreted as saying that for decentralized systems to be (weakly) better than centralized systems, it must be that the dictator is concerned with fairness and not just efficiency.

Finally, the assumption of ex-post individual rationality is perhaps the most important assumption in that any allocation that is not ex-post IR is not implementable by the extended VCM, because, no agent would ever choose a transfer larger than his valuation. Allocations that are not ex-post IR have a debatable appeal; they might present practical difficulties (for example, it could be that if an agent anticipates that he will be made worse off from the provision of the public good, he might be tempted to move to a region where the public good is not provided) and/or might be undesirable on ethical grounds, given that they depend on coercion. However, regardless of how unappealing these other allocations might be for the reader, the result still proves useful in demonstrating that those are the only allocations that decentralized systems cannot replicate, i.e., the benefit of centralized systems comes not from being more able to elicit the agents' private preferences but rather from coercion.

6 Appendix

6.1 Proof of lemma 1

First, notice that

$$E_{v_{-n}}\left(t_n\left(\sigma\left(v_n',v_{-n}\right)\right)\right)=E_{v_{-n}}\left(t_n\left(\sigma\left(v_n'',v_{-n}\right)\right)\right)=k_1$$

for some $k_1 \in \mathbb{R}$, because, if not, one of the types would have a strict incentive to mimic the other. Take any $v_n \in [v_n', v_n'']$ and notice that

$$v_n \left(E_{v_{-n}}\left(p\left(\sigma\left(v_n, v_{-n}\right)\right)\right) - k_0\right) \geq E_{v_{-n}}\left(t_n\left(\sigma\left(v_n, v_{-n}\right)\right)\right) - k_1$$

If $E_{v_{-n}}\left(p\left(\sigma\left(v_n, v_{-n}\right)\right)\right) > k_0$, then it follows that

$$v_n'' \left(E_{v_{-n}}\left(p\left(\sigma\left(v_n, v_{-n}\right)\right)\right) - k_0\right) > E_{v_{-n}}\left(t_n\left(\sigma\left(v_n, v_{-n}\right)\right)\right) - k_1$$

because $v_n'' > v_n$, which is a contradiction. If, one the other hand, $E_{v_{-n}}\left(p\left(\sigma\left(v_n, v_{-n}\right)\right)\right) < k_0$, it follows that

$$v_n' \left(E_{v_{-n}}\left(p\left(\sigma\left(v_n, v_{-n}\right)\right)\right) - k_0\right) \geq E_{v_{-n}}\left(t_n\left(\sigma\left(v_n, v_{-n}\right)\right)\right) - k_1$$

because $v_n'' < v_n$, which is a contradiction.

Proof. Finally, because $E_{v_{-n}}\left(p\left(\sigma\left(v_n, v_{-n}\right)\right)\right) = k_0$ for all $v_n \in [v_n', v_n'']$, it follows that $E_{v_{-n}}\left(t_n\left(\sigma\left(v_n, v_{-n}\right)\right)\right) = k_1$ for all $v_n \in [v_n', v_n'']$. ■

6.2 Proof of Proposition 1

Take any $(\widehat{\rho}, \widehat{\tau}) \in \Psi^{BD} \cup \Psi^{IR}$. Let $\left(\widehat{\sigma}, \widehat{p}, \widehat{t}\right) \in \Gamma^{BD}$ be the profile that implements allocation $(\widehat{\rho}, \widehat{\tau})$ in the BD system. For each n, let

$$\widehat{\gamma}_n\left(v_n, m\right) = \begin{cases} \widehat{t}_n\left(m\right) \text{ if } \widehat{t}_n\left(m\right) \leq v_n \\ 0 \text{ otherwise} \end{cases}$$

for all $m \in M$ and $v_n \in V_n$.

Notice that, because $(\widehat{\rho}, \widehat{\tau})$ is ex-post IR, if $(\widehat{\sigma}, \widehat{\gamma}) \in \Gamma^{VCM}$, then allocation $(\widehat{\rho}, \widehat{\tau})$ is implemented in the extended VCM. In what follows, I show precisely that. Consider some arbitrary agent n of type v_n and assume all other agents play according to $(\widehat{\sigma}, \widehat{\gamma})$.

Lemma 2 *Following any message vector $m \in M$, agent n prefers to choose transfer $t_n = \widehat{\gamma}_n (v_n, m)$ for all $m \in M$.*

Proof. If m is such that $\widehat{p}(m) = 0$, the transfer that each of the other agents will make is 0 by the PPE property. Seeing as $\overline{v}_n < c$, it is in the best interest of agent n to choose $t_n = \widehat{\gamma}_n (v_n, m) = 0$. If m is such that $\widehat{p}(m) = 1$, then $\sum_{\widehat{n}=1}^{N} \widehat{t}_{\widehat{n}} (m) = c$ by the PPE property. In this case, the best possible deviation for agent n would be to choose $t_n = 0$, which would return a payoff of 0. Because $(\widehat{\rho}, \widehat{\tau})$ is ex-post individually rational, this is not a strictly beneficial deviation. ■

Given Lemma 2, what is left to show is that agent n does not want to deviate on the report.

Lemma 3 *Agent n does not strictly prefer to report $m_n = \widehat{\sigma}_n (v'_n)$ for which*

$$E_{v_{-n}} \left(\widehat{p} \left(\widehat{\sigma}_n (v_n), \sigma_{-n} (v_{-n})\right)\right) = E_{v_{-n}} \left(\widehat{p} \left(\widehat{\sigma}_n (v'_n), \widehat{\sigma}_{-n} (v_{-n})\right)\right).$$

Proof. Take any $v'_n \in V_n$ such that

$$E_{v_{-n}} \left(\widehat{p} \left(\widehat{\sigma}_n (v_n), \sigma_{-n} (v_{-n})\right)\right) = E_{v_{-n}} \left(\widehat{p} \left(\widehat{\sigma}_n (v'_n), \widehat{\sigma}_{-n} (v_{-n})\right)\right).$$

Notice that, because of the PPE property and the fact that (ρ, τ) is ex-post individually rational, it follows that either

$$\widehat{p} \left(\widehat{\sigma}_n (v_n), \sigma_{-n} (v_{-n})\right) \geq \widehat{p} \left(\widehat{\sigma}_n (v'_n), \sigma_{-n} (v_{-n})\right)$$

for all v_{-n} or

$$\widehat{p} \left(\widehat{\sigma}_n (v_n), \sigma_{-n} (v_{-n})\right) \leq \widehat{p} \left(\widehat{\sigma}_n (v'_n), \sigma_{-n} (v_{-n})\right)$$

for all v_{-n}. Therefore,

$$\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right), \sigma_{-n}\left(v_{-n}\right)\right) = \widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right), \sigma_{-n}\left(v_{-n}\right)\right)$$

for almost all v_{-n}. Analogously, the MT property implies that

$$\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n\right), \sigma_{-n}\left(v_{-n}\right)\right) = \widehat{t}_n\left(\widehat{\sigma}_n\left(v_n'\right), \sigma_{-n}\left(v_{-n}\right)\right)$$

for almost all v_{-n}. Therefore, the expected payoff for type v_n of sending message $m_n = \widehat{\sigma}_n\left(v_n\right)$ and $m_n = \widehat{\sigma}_n\left(v_n'\right)$ is equal, so that there is no strict incentive to deviation. ■

Lemma 4 *Agent n does not strictly prefer to report $m_n = \widehat{\sigma}_n\left(v_n'\right)$ for which*

$$E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right), \sigma_{-n}\left(v_{-n}\right)\right)\right) > E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right).$$

Proof. Take any $v_n' \in V_n$ such that

$$E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right), \sigma_{-n}\left(v_{-n}\right)\right)\right) > E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right).$$

Notice that $v_n > v_n'$ (by the same logic as in Lemma 1). Because (ρ, τ) is ex-post individually rational, it follows that $\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n'\right), m_{-n}\right) \leq v_n' < v_n$ for all m_{-n}. This means that

$$\widehat{\gamma}_n\left(v_n, \widehat{\sigma}_n\left(v_n'\right), m_{-n}\right) = \widehat{t}_n\left(\widehat{\sigma}_n\left(v_n'\right), m_{-n}\right)$$

for all m_{-n}, which would make the deviation to reporting $m_n = \widehat{\sigma}_n\left(v_n'\right)$ equivalent to a similar deviation in the BD system. ■

Lemma 5 *If there are types* $v_n' \in V_n$ *and* $v_n'' \in V_n$ *such that*

$$E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) > E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n''\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) > E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right),$$

then agent n *prefers to report* $m_n = \widehat{\sigma}_n\left(v_n''\right)$ *over* $m_n = \widehat{\sigma}_n\left(v_n'\right)$*.*

Proof. Recall that, because the PPE and MT properties and the fact that $(\widehat{\rho}, \widehat{\tau})$ is ex-post IR, it follows that

$$\widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right) \geq \widehat{p}\left(\widehat{\sigma}_n\left(v_n''\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right) \geq \widehat{p}\left(\widehat{\sigma}_n\left(v_n\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right)$$

and

$$\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right) \geq \widehat{t}_n\left(\widehat{\sigma}_n\left(v_n''\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right) \geq \widehat{t}_n\left(\widehat{\sigma}_n\left(v_n\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right)$$

for all v_{-n}. The proof is complete by showing that for all v_{-n} such that

$$\widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right) = 1 > \widehat{p}\left(\widehat{\sigma}_n\left(v_n''\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right) = 0,$$

it follows that

$$\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right) > v_n$$

Take any such v_{-n} and notice that the PPE property implies that

$$\sum_{\widehat{n} \neq n} \inf\left(supp\left(\pi_{\widehat{n}}^{\widehat{\sigma}}\left(\widehat{\sigma}_{\widehat{n}}\left(v_{\widehat{n}}\right)\right)\right)\right) + \inf\left(supp\left(\pi_n^{\widehat{\sigma}}\left(\widehat{\sigma}_n\left(v_n''\right)\right)\right)\right) < c.$$

Because

$$\inf\left(supp\left(\pi_n^{\widehat{\sigma}}\left(\widehat{\sigma}_n\left(v_n''\right)\right)\right)\right) \geq \sup\left(supp\left(\pi_n^{\widehat{\sigma}}\left(\widehat{\sigma}_n\left(v_n\right)\right)\right)\right) \geq v_n,$$

it follows that

$$\sum_{\widehat{n}\neq n} \inf\left(supp\left(\pi_{\widehat{n}}^{\widehat{\sigma}}\left(\widehat{\sigma}_{\widehat{n}}\left(v_{\widehat{n}}\right)\right)\right)\right)+v_n<c.$$

Hence, because

$$\widehat{t}_{\widehat{n}}\left(\widehat{\sigma}_{\widehat{n}}\left(v_{\widehat{n}}\right),\widehat{\sigma}_{-\widehat{n}}\left(v_{-\widehat{n}}\right)\right)\leq \inf\left(supp\left(\pi_{\widehat{n}}^{\widehat{\sigma}}\left(\widehat{\sigma}_{\widehat{n}}\left(v_{\widehat{n}}\right)\right)\right)\right)$$

for all $\widehat{n}\neq n$, it must be that

$$\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)>v_n.$$

■

Lemma 6 *The agent does not want to report $m_n=\widehat{\sigma}_n\left(v_n'\right)$, where $v_n'\in V_n$ is such that i)*

$$E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)>E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)$$

and ii) there is no $v_n''\in V_n$ such that

$$E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)>E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n''\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)>E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right).$$

Proof. Case 1: $v_n=\inf\left(supp\left(\pi_n^{\widehat{\sigma}}\left(\widehat{\sigma}_n\left(v_n\right)\right)\right)\right)$. In this case, the agent does not deviate to $m_n=\widehat{\sigma}_n\left(v_n'\right)$ by the same argument as in the previous lemma: whenever the public good is provided as a result of the agent reporting $m_n=\widehat{\sigma}_n\left(v_n'\right)$ as opposed to reporting $m_n=\widehat{\sigma}_n\left(v_n\right)$, the requested transfer on the BD system would exceed the agent's valuation.

Case 2: $v_n > \inf\left(supp\left(\pi_n^{\widehat{\sigma}}\left(\widehat{\sigma}_n\left(v_n\right)\right)\right)\right) \equiv x_n$. Let

$$\widetilde{p}_n\left(m, v_n\right) = \begin{cases} \widehat{p}\left(m\right) \text{ if } \widehat{t}_n\left(m\right) \leq v_n \\ 0 \text{ otherwise} \end{cases}$$

and

$$\widetilde{t}_n\left(v_n, m\right) = \widehat{\gamma}_n\left(v_n, m\right).$$

Notice that $\widetilde{p}_n\left(m, v_n\right)$ and $\widetilde{t}_n\left(m, v_n\right)$ represent, respectively, the probability that the public good is provided and agent n's transfer after message m, when the agent's type is v_n.

Suppose that type v_n strictly prefers to deviate to $m_n = \widehat{\sigma}_n\left(v_n'\right)$ so that

$$\begin{aligned} & v_n E_{v_{-n}}\left(\widetilde{p}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right), v_n\right)\right) - E_{v_{-n}}\left(\widetilde{t}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right), v_n\right)\right) \\ > \; & v_n E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) - E_{v_{-n}}\left(\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) \end{aligned}$$

Case 2a:

$$E_{v_{-n}}\left(\widetilde{p}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right), v_n\right)\right) \leq E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right).$$

In this case, it follows that

$$\begin{aligned} & x_n E_{v_{-n}}\left(\widetilde{p}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right), x_n\right)\right) - E_{v_{-n}}\left(\widetilde{t}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right), x_n\right)\right) \\ \geq \; & x_n E_{v_{-n}}\left(\widetilde{p}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right), v_n\right)\right) - E_{v_{-n}}\left(\widetilde{t}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right), v_n\right)\right) \\ > \; & x_n E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) - E_{v_{-n}}\left(\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n\right), \widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right), \end{aligned}$$

so that, if the agent's type is $v_n = x_n$, he also strictly prefers to deviate to $m_n = \widehat{\sigma}_n\left(v_n'\right)$, which is a contradiction to case 1.

Case 2b:

$$E_{v_{-n}}\left(\widetilde{p}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right),v_n\right)\right)>E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right).$$

Let $\widehat{v}_n \in V_n$ be such that

$$\begin{aligned} & \widehat{v}_n E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)-E_{v_{-n}}\left(\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) \\ = & \widehat{v}_n E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)-E_{v_{-n}}\left(\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)\end{aligned}$$

Notice that it must be that

$$\widehat{v}_n=\sup\left(supp\left(\pi_n^{\widehat{\sigma}}\left(\widehat{\sigma}_n\left(v_n\right)\right)\right)\right)=\inf\left(supp\left(\pi_n^{\widehat{\sigma}}\left(\widehat{\sigma}_n\left(v_n'\right)\right)\right)\right)$$

so that

$$\begin{aligned} & \widehat{v}_n E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)-E_{v_{-n}}\left(\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) \\ = & \widehat{v}_n E_{v_{-n}}\left(\widetilde{p}_n\left(\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right),\widehat{v}_n\right)\right)-E_{v_{-n}}\left(\widetilde{t}_n\left(\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right),\widehat{v}_n\right)\right)\end{aligned}$$

However, given that $\widehat{v}_n \geq v_n$, it must be that

$$\begin{aligned} & \widehat{v}_n E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)-E_{v_{-n}}\left(\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) \\ < & \widehat{v}_n E_{v_{-n}}\left(\widetilde{p}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right),v_n\right)\right)-E_{v_{-n}}\left(\widetilde{t}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right),v_n\right)\right) \\ \leq & \widehat{v}_n E_{v_{-n}}\left(\widetilde{p}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right),\widehat{v}_n\right)\right)-E_{v_{-n}}\left(\widetilde{t}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right),\widehat{v}_n\right)\right),\end{aligned}$$

which is a contradiction. ■

The only thing left to show is that type v_n does not want to misreport "up" when there is no "next largest" message.

Lemma 7 *If, for all types $v_n' \in V_n$ such that*

$$E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) > E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)$$

there is some type $v_n'' \in V_n$ such that

$$E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) > E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n''\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) > E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right),$$

then the agent does not strictly prefer to report any $m_n = \widehat{\sigma}_n\left(v_n'\right)$.

Proof. Take any type $v_n' \in V_n$ such that

$$E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) > E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)$$

and suppose that type v_n strictly prefers to report message $m_n = \widehat{\sigma}_n\left(v_n'\right)$. Let

$$\begin{aligned}\varepsilon &= v_n E_{v_{-n}}\left(\widetilde{p}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right),v_n\right)\right) - E_{v_{-n}}\left(\widetilde{t}_n\left(\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right),v_n\right)\right)\\ &\quad - v_n E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) + E_{v_{-n}}\left(\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)\end{aligned}$$

Notice that there is some $v_n'' \in V_n$ such that

$$E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n'\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) > E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n''\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) > E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right)$$

that has the property that

$$\begin{aligned}\varepsilon &> v_n E_{v_{-n}}\left(\widetilde{p}_n\left(\left(\widehat{\sigma}_n\left(v_n''\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right),v_n\right)\right) - E_{v_{-n}}\left(\widetilde{t}_n\left(\left(\widehat{\sigma}_n\left(v_n''\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right),v_n\right)\right)\\ &\quad - v_n E_{v_{-n}}\left(\widehat{p}\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right) + E_{v_{-n}}\left(\widehat{t}_n\left(\widehat{\sigma}_n\left(v_n\right),\widehat{\sigma}_{-n}\left(v_{-n}\right)\right)\right),\end{aligned}$$

which contradicts Lemma 6. ■

6.3 Proof of proposition 2

In figure 2, the red line represents function $m^\lambda v_1 + b^\lambda$, while the green line represents function $-v_1 + c$. Let $\widehat{v}_1$ and $\widehat{v}_2$ be defined as in figure 2, i.e, $\widehat{v}_1 = \frac{\overline{v} - b^\lambda}{m^\lambda}$ and $\widehat{v}_2 = m^\lambda \overline{v} + b^\lambda$.

First, I show that for some allocation $(\rho, \tau) \in \omega(\lambda)$ to be implemented by the BD system, it must be that agent 1 reports truthfully when $v_1 \in (\widehat{v}_1, \overline{v})$ and agent 2 reports truthfully when $v_2 \in (\widehat{v}_2, \overline{v})$.

Suppose not, i.e., say that some type $v_1' \in (\widehat{v}_1, \overline{v})$ sends some message m_1' that is also sent by some other type $v_1'' \neq v_1'$. Then, it would have to $\rho^\lambda(v_1', v_2) = \rho^\lambda(v_1'', v_2)$ for all $v_2 \in (\underline{v}, \overline{v})$, because the BD would not be able to distinguish the two types, which is false as one can see by the fact that m^λ and that $v_1' \in (\widehat{v}_1, \overline{v})$. The same argument can be made for agent 2.

Consider the case where the agents' types (and their reports, because they report truthfully) are $(v_1', v_2') \in (\widehat{v}_1, \overline{v}) \times (\widehat{v}_2, \overline{v})$. If $v_1' + v_2' \geq c$, then, in the BD system, the public good has to be provided by the PPE property because there is a way to fund the public good which makes all agents better off with certainty. And if $v_1' + v_2' < c$, the public good cannot be provided because of the ex-post individual rationality assumption, as it would be impossible to make sure that both agents had a positive payoff. So, one can conclude that, for any report $(v_1, v_2) \in (\widehat{v}_1, \overline{v}) \times (\widehat{v}_2, \overline{v})$ in the BD system, the public good would be provided if and only if $\rho^*(v_1, v_2) = 1$, where

$$\rho^*(v) = \begin{cases} 1 \text{ if } v_2 \geq -v_1 + c \\ 0 \text{ otherwise} \end{cases}.$$

Simple algebra based on Agastya et al. (2007) shows that, for all $\lambda \in [0, 1]$, $(m^\lambda, b^\lambda) \neq (-1, c)$, which implies $\rho^\lambda \neq \rho^*$, a contradiction. Therefore, (ρ, τ) is not implementable by the BD system.

6.4 Proof of Proposition 3

Consider a PBE (σ, p, t) and take any agent n, and any two messages $m'_n \in M_n$ and $m''_n \in M_n$. Consider any m_{-n} for which

$$p(m'_n, m_{-n}) = p(m''_n, m_{-n}) = 1.$$

By definition,

$$t(m) \in \arg\max_{\widehat{t} \in \mathbb{R}^N_+} \prod_{n=1}^{N} \left(E(v_n | m_n) - \widehat{t}_n + c\right)$$

Given that there is a unique solution to the problem, if $E(v_n|m'_n) = E(v_n|m''_n)$, then it follows that $t_n(m'_n, m_{-n}) = t_n(m''_n, m_{-n})$. WLOG, suppose that $E(v_n|m'_n) > E(v_n|m''_n)$. I show that $t_n(m'_n, m_{-n}) \geq t_n(m''_n, m_{-n})$, which completes the proof.

If $t_n(m''_n, m_{-n}) = 0$, the statement follows trivially. Suppose that $t_n(m''_n, m_{-n}) > 0$. In that case, it must be that

$$E(v_n|m''_n) - t_n(m''_n, m_{-n}) \geq E(v_{\widehat{n}}|m_{\widehat{n}}) - t_{\widehat{n}}(m''_n, m_{-n})$$

for all $\widehat{n}$, with an equality whenever $t_{\widehat{n}}(m''_n, m_{-n}) > 0$. This is because, if not, one could reduce $t_n(m''_n, m_{-n})$ and increase the transfer of the agent whose expected utility was larger and make the BD better off.

By way of contradiction, suppose that $t_n(m'_n, m_{-n}) < t_n(m''_n, m_{-n})$, so that

$$E(v_n|m'_n) - t_n(m'_n, m_{-n}) > E(v_n|m''_n) - t_n(m''_n, m_{-n}).$$

It would have to be that, for all $\widehat{n} \neq n$,

$$E\left(v_{\widehat{n}}|m_{\widehat{n}}\right) - t_{\widehat{n}}\left(m_n', m_{-n}\right) \geq E\left(v_{\widehat{n}}|m_{\widehat{n}}\right) - t_{\widehat{n}}\left(m_n'', m_{-n}\right),$$

which would imply that

$$c = \sum_{\widehat{n}=1}^{N} t_{\widehat{n}}\left(m_n'', m_{-n}\right) > \sum_{\widehat{n}=1}^{N} t_{\widehat{n}}\left(m_n', m_{-n}\right),$$

which is a contradiction.

6.5 Proof of Proposition 4

Consider a PBE (σ, p, t) and take any agent n, and any two messages $m_n' \in M_n$ and $m_n'' \in M_n$. Consider any m_{-n} for which

$$p\left(m_n', m_{-n}\right) = p\left(m_n'', m_{-n}\right) = 1.$$

By definition,

$$t\left(m\right) \in \arg\max_{\widehat{t} \in \mathbb{R}_+^N} \sum_{n=1}^{N} w_n\left(E\left(v_n|m_n\right) - \widehat{t}_n\right)$$

Because each w_n is strictly concave, it follows that there is a unique solution, so that if $E\left(v_n|m_n'\right) = E\left(v_n|m_n''\right)$, it follows that $t_n\left(m_n', m_{-n}\right) = t_n\left(m_n'', m_{-n}\right)$. WLOG, suppose that $E\left(v_n|m_n'\right) > E\left(v_n|m_n''\right)$. I show that $t_n\left(m_n', m_{-n}\right) \geq t_n\left(m_n'', m_{-n}\right)$, which completes the proof.

If $t_n\left(m_n'', m_{-n}\right) = 0$, the statement follows trivially. Suppose that $t_n\left(m_n'', m_{-n}\right) > 0$, which implies that

$$w_n'\left(E\left(v_n|m_n''\right) - t_n\left(m_n'', m_{-n}\right)\right) \leq w_{\widehat{n}}'\left(E\left(v_{\widehat{n}}|m_{\widehat{n}}\right) - t_{\widehat{n}}\left(m_n'', m_{-n}\right)\right)$$

for all $\widehat{n}$, with an equality whenever $t_{\widehat{n}}(m''_n, m_{-n}) > 0$. Assume, by way of contradiction that $t_n(m''_n, m_{-n}) > t_n(m'_n, m_{-n})$. Then, it follows that

$$w'_n(E(v_n|m''_n) - t_n(m''_n, m_{-n})) > w'_n(E(v_n|m'_n) - t_n(m'_n, m_{-n})).$$

In turn, this implies that

$$w'_{\widehat{n}}(E(v_{\widehat{n}}|m_{\widehat{n}}) - t_{\widehat{n}}(m'_n, m_{-n})) \leq w'_{\widehat{n}}(E(v_{\widehat{n}}|m_{\widehat{n}}) - t_{\widehat{n}}(m''_n, m_{-n}))$$

for all $\widehat{n} \neq n$, which means that $t_{\widehat{n}}(m''_n, m_{-n}) \geq t_{\widehat{n}}(m'_n, m_{-n})$ for all $\widehat{n} \neq n$. But then, this means that

$$\sum_{\widehat{n}=1}^{N} t_{\widehat{n}}(m''_n, m_{-n}) > \sum_{\widehat{n}=1}^{N} t_{\widehat{n}}(m'_n, m_{-n}) = c,$$

which is a contradiction.

6.6 Proof of Proposition 5

Consider a PBE (σ, p, t) and take any agent n, and any two messages $m'_n \in M_n$ and $m''_n \in M_n$. Consider any m_{-n} for which

$$p(m'_n, m_{-n}) = p(m''_n, m_{-n}) = 1.$$

By definition,

$$t(m) \in \arg\max_{\widehat{t} \in \mathbb{R}_+^N} \sum_{n=1}^{N} E\left[w_n\left(v_n - \widehat{t}_n\right)|m_n\right]$$

Because $w'''_n = 0$, it follows that either

$$E\left[w'_n\left(v_n - \widehat{t}_n\right)|m'_n\right] \geq E\left[w'_n\left(v_n - \widehat{t}_n\right)|m''_n\right]$$

for all $\widehat{t}_n \in \mathbb{R}$ or

$$E\left[w_n'\left(v_n - \widehat{t}_n\right)|m_n'\right] \leq E\left[w_n'\left(v_n - \widehat{t}_n\right)|m_n''\right]$$

for all $\widehat{t}_n \in \mathbb{R}$. WLOG, say it is the latter. I claim that $t_n(m_n', m_{-n}) \geq t_n(m_n'', m_{-n})$.

If $t_n(m_n'', m_{-n}) = 0$, the statement follows trivially, so consider the case that $t(m_n'', m_{-n}) > 0$. In that case, it follows that

$$E\left[w_n'\left(v_n - t_n\left(m_n'', m_{-n}\right)\right)|m_n''\right] \leq E\left[w_{\widehat{n}}'\left(v_{\widehat{n}} - t_{\widehat{n}}\left(m_n'', m_{-n}\right)\right)|m_{\widehat{n}}\right]$$

for all $\widehat{n}$. Suppose, by way of contradiction, that $t_n(m_n'', m_{-n}) > t_n(m_n', m_{-n})$. It follows that

$$\begin{aligned} E\left[w_n'\left(v_n - t_n\left(m_n'', m_{-n}\right)\right)|m_n''\right] &\geq E\left[w_n'\left(v_n - t_n\left(m_n'', m_{-n}\right)\right)|m_n'\right] \\ &> E\left[w_n'\left(v_n - t_n\left(m_n', m_{-n}\right)\right)|m_n'\right]. \end{aligned}$$

Therefore, it must be that for all $\widehat{n} \neq n$,

$$E\left[w_{\widehat{n}}'\left(v_{\widehat{n}} - t_{\widehat{n}}\left(m_n', m_{-n}\right)\right)|m_{\widehat{n}}\right] \leq E\left[w_{\widehat{n}}'\left(v_{\widehat{n}} - t_{\widehat{n}}\left(m_n'', m_{-n}\right)\right)|m_{\widehat{n}}\right],$$

which implies that $t_{\widehat{n}}(m_n'', m_{-n}) \geq t_{\widehat{n}}(m_n', m_{-n})$ for all $\widehat{n} \neq n$. But then it follows that

$$\sum_{\widehat{n}=1}^{N} t_{\widehat{n}}\left(m_{\widehat{n}}'', m_{-\widehat{n}}\right) > \sum_{\widehat{n}=1}^{N} t_{\widehat{n}}\left(m_{\widehat{n}}', m_{-\widehat{n}}\right) = c,$$

which is a contradiction.

References

[1] Agastya, Murali, Flavio Menezes, and Kunal Sengupta. "Cheap talk, efficiency and egalitarian cost sharing in joint projects." Games and Economic Behavior 60.1 (2007): 1-19.

[2] Anderson, Christopher M., and Louis Putterman. "Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism." Games and Economic Behavior 54.1 (2006): 1-24.

[3] d'Aspremont, Claude, and Louis-Andre Gérard-Varet. "Incentives and incomplete information." Journal of Public economics 11.1 (1979): 25-45.

[4] Bárány, Imre. "Fair distribution protocols or how the players replace fortune." Mathematics of Operations Research 17.2 (1992): 327-340.

[5] Barbieri, Stefano. "Communication and Early Contributions." Journal of Public Economic Theory 14.3 (2012): 391-421.

[6] Ben-Porath, Elchanan. "Correlation without mediation: Expanding the set of equilibrium outcomes by "cheap" pre-play procedures." Journal of Economic Theory 80.1 (1998): 108-122.

[7] Ben-Porath, Elchanan. "Cheap talk in games with incomplete information." Journal of Economic Theory 108.1 (2003): 45-71.

[8] Besley, Timothy, and Stephen Coate. "Centralized versus decentralized provision of local public goods: a political economy approach." Journal of public economics 87.12 (2003): 2611-2637.

[9] Buchanan, James M., and Gordon Tullock. "The calculus of consent." Vol. 3. Ann Arbor: University of Michigan Press (1962).

[10] Crawford, Vincent P., and Joel Sobel. "Strategic information transmission." Econometrica (1982): 1431-1451.

[11] Cho, Myeonghwan. "Externality and information asymmetry in the production of local public good." International Journal of Economic Theory 9 (2013): 177-201.

[12] Forges, Françoise. "Universal mechanisms." Econometrica (1990): 1341-1364.

[13] Gerardi, Dino. "Unmediated communication in games with complete and incomplete information." Journal of Economic Theory 114.1 (2004): 104-131.

[14] Groves, Theodore. "Incentives in teams." Econometrica (1973): 617-631.

[15] Laffont, Jean-Jacques, and Eric Maskin. "On the difficulty of attaining distributional goals with imperfect information about consumers." The Scandinavian Journal of Economics (1979): 227-237.

[16] Lockwood, Ben. "Distributive politics and the cost of centralization." The Review of Economic Studies 69.2 (2002): 313-337.

[17] Kessler, Anke S. "Communication in Federal Politics: Universalism, Policy Uniformity, and the Optimal Allocation of Fiscal Authority." Journal of Political Economy, 122.4 (2014): 766:805.

[18] Klibanoff, Peter, and Michel Poitevin. "A Theory of (De)centralization." (Unpublished paper, Northwestern University (2013).

[19] Matthews, Steven A., and Andrew Postlewaite. "Pre-play communication in two-person sealed-bid double auctions." Journal of Economic Theory 48.1 (1989): 238-263.

[20] Masclet, David, Noussair, Charles, Tucker, Steven, & Villeval, Marie-Claire. "Monetary and nonmonetary punishment in the voluntary contributions mechanism". American Economic Review 93.1 (2003): 366-380.

[21] Myerson, Roger B. "Incentive compatibility and the bargaining problem." Econometrica (1979): 61-73.

[22] Oates, Wallace E. "Fiscal federalism." (1972). Harcort, Brace, Jovanovich, New York.

[23] Samuelson, Paul A. "The pure theory of public expenditure." The review of economics and statistics (1954): 387-389.

[24] Schmidt, Klaus M. "The costs and benefits of privatization: an incomplete contracts approach." Journal of Law, Economics, and Organization 12.1 (1996): 1-24.