FAERE

French Association
of Environmental and Resource Economists

Working papers

Voluntary agreements as correlated equilibria of a subscription game

Anne-Sarah Chiambretto

WP 2017.23

Suggested citation:

A-S. Chiambretto *(2017).* Voluntary agreements as correlated equilibria of a subscription game. *FAERE Working Paper, 2017.23.*

ISSN number: **2274-5556**

www.faere.fr

Voluntary agreements as correlated equilibria of a subscription game[☆]

Anne-Sarah Chiambretto [1,*]

Aix-Marseille University (Aix-Marseille School of Economics), CNRS, & EHESS

Abstract

I develop a subscription game, modified so as to represent firms' incentives to participate to an environmental Voluntary Agreement (VA). Specifically, I assume the VA is preemptive, i.e. it occurs under the threat of a mandatory regulation. I suggest the use of a correlating device to strengthen firms participation, formalized by the concept of correlated equilibrium (CE). I characterize the multiple pure and mixed Nash equilibria (NE) of the game without the correlating device. I find that such a device not only solves the problem raised by multiplicity of NE, but also ensures that a higher expected aggregate payoff is reached for any given level of threat. I provide a full comparative efficiency analysis after the optimal CE is characterized, and study the impact of the threat stringency. Finally, I illustrate the general results in a specified example of pollution abatement model.

Keywords: collective voluntary agreements, pollution control, adoption costs, distributional effects, public goods, government policy.
JEL classification: C79, H23, H41, Q58.

[☆]The author thanks Francois Salanié and Bernard Sinclair-Desgagné for their helpful comments, as well as Hubert Stahn for the several productive discussions that helped in the starting of the present work. Of course the usual claim applies. The financial support of the Labex AMSE (ANR- 11-IDEX-0001-02) and the ANR GREEN-Econ (ANR-16-CE03-0005) are also gratefully acknowledged.

[*]Corresponding author

Email address: anne-sarah.chiambretto@inra.fr (Anne-Sarah Chiambretto)

[1]P-mail address: GREQAM, Centre de la Vieille Charité, 2, rue de la Charité, 13236 Marseille, France; Tel.: +33 (0) 4 91 14 07 21; Fax: +33 (0) 4 91 90 02 27

1. Introduction

Voluntary agreements (VAs hereafter) generically refer to the many and multiform schemes whereby a group of producers that generates some environmental externality engage in self-regulation. When these proactive behavior are motivated by a background regulatory threat, whether enacted (e.g. when used in a policy mix coupled with command and control, see Borkey et al. 4) or merely potential (Glachant 13), the literature mentions them as "preemptive". Since potential enactments most often turn out to be sectoral, preemptive VAs implicitly rely on a collective liability rule. Namely, if some global environmental cap cannot be voluntarily achieved by a group, the mandatory legislation applies regardless of individual voluntary efforts. Likewise, the preemption can possibly arise from any allocation rule of private reduction objectives as long as it fulfills the global target. Such an incentive structure is known by the literature on environmental VAs to generate two main issues. One is free-riding, widely studied in competitive game-theoretic frameworks without inter-firms communication (see Segerson and Wu 23, Dawson and Segerson 9, or Brau and Carraro 5). Another issue, hardly mentioned as a theoretical non-uniqueness property of the solution, is the multiplicity of burden sharing options, which may practically cause the preemptive VA to fail by lack of coordination.

. To our knowledge, only Glachant 12 mentions the need for devising procedural features to rule the inter-firms burden sharing process. After he stated the separability between the setting of the global objective and the means to collectively reach it, he briefly sketches such a process as a bargaining game between two asymmetric firms. However, a more complete understanding of multiplicity and free-riding in similar incentive contexts is provided by the general literature on voluntary provision of public goods (see Palfrey and Rosenthal 21, Dixit and Olson 10, Moulin 16). Thus, consistent with Glachant 12 separability principle, the present work formalizes preemptive VAs as a subscription game to a discrete public good, and focuses on firms' participation decisions. Specifically, the public good to be provided is the non rival and non excludable regulatory gains that the firms derive from the mandatory regulation preemption arising from a succeeding VA. These gains depend on the regulatory threat t, and the cost of providing the public good (i.e. the cost of satisfying the exogenous VA policy requirement). Full participation is socially optimal as it is assumed the total cost of providing the public good decreases in the number of participating firms. Such a framework is general enough to fit the numerous institutional background[2] within which VAs may take place.

Then, I apply Aumann 2's correlated equilibrium concept, as reflecting the potential for firms to coordinate. Since correlated equilibria are self-enforcing by nature, the solutions characterized are purely non-cooperative and voluntary. Beyond giving positive

[2] A widely adopted classification distinguishes between three main categories[3], set on the basis of the regulatory agency involvement level (OECD 19) : (i) *public voluntary agreements* (the agency elaborates engagements, to which the firms may voluntarily subscript), (ii) *negotiated agreements* (the engagements are collaboratively elaborated by the voluntary firms and the agency), and (iii) *unilateral commitments* (the engagements are elaborated by the voluntary firms solely).

insights about the burden sharing process between firms, a normative result is stated. I show that, if instead of participation being offered to all the firms from the sector to be regulated, some third party[4] makes private recommendation to each of them, it can implement a better participation equilibrium than the one that could have been possibly achieved in a case without private recommandation. This mechanism is called a correlating device, and such an equilibrium is characterized in the general n-player case.

. The results of this paper relate to the seminal Palfrey and Rosenthal 21 that studies, inter alia, the mixed Nash equilibria of a subscription game. My findings differ from their own in two ways. While Palfrey and Rosenthal focus on the relationship between the set of Nash equilibria of provision games with (i.e. subscription games) and without a refund, I elaborate on the concept of correlated equilibrium and focus on a slightly different game design. In the present work, the 'greed' motivation for free-riding can indeed be manipulated by the regulator through the tax threat, while assumptions on provision costs involve that the participation of a subgroup of players is socially inefficient. This specification reflects the participation incentives typically arising from preemptive VAs with an implicit collective liability. Dixit and Olson 10 extends Palfrey and Rosenthal 21 by assuming the fixed total provision cost is equally shared within participating players in a second stage. While the sharing is said to be the result of a cosean approach, the first stage is, once again, solved as the Nash equilibrium of a participation game. Same remarks as for Palfrey and Rosenthal 21 apply.

In contrast, Cavaliere 6 studies the provision of discrete public good by correlated equilibria. Nevertheless, he restricts his study to the same general game *without* a refund as studied in Palfrey and Rosenthal 21. He finds that any convex combinations of pure strategies Nash equilibria are efficient correlated equilibria. He then introduces payoff externalities by assuming that both the consumption benefits and the contribution amount increases with the number of contributor, which still substantially differs from the game analyzed in the present work. In addition, his results rely on a restrictive hypothesis of strategies symmetry. Last but not least, Arce and Sandler 1 make a point similar as the present work by relating CE to International Environmental Agreements (IEAs). While undoubtedly fruitful, their contribution consists of highlighting conceptual correspondances and illustrating it by 3-players examples under different aggregation technologies. Moreover, while central to the present paper, the specific issue of the threat is not investigated since irrelevant in the application context of IEAs.

. This work is organized as follows. In section 2, I present the base game and the correlating device. The multiple Nash equilibria are characterized in section 3, as benchmarks. In section 4, I study the set of correlated equilibria and identify the optimum. Section 5 compares the efficiency of the optimal correlated equilibrium and the nash equilibria. Then, a static analysis with respect to the threat and the number of firms is provided. Finally, I numerically illustrate, in section 6, the results in a specified example of pollution

[4] Of course, such a third party may be real or notional.

abatement.

2. The model

2.1. The basic game

Consider a regulator willing to achieve a global regulation cap that involves the production activities[5] of n identical firms. Each firm, indexed by $i \in \{1, ..., n\}$, can voluntarily choose to participate ($s_i = 1$) or not to participate ($s_i = 0$) to the achievement of the target. The modalities of participation to the VA depend on $\sum_i s_i \equiv m$, the number of participating firms. Indeed, for any fixed cap level, implementation costs are equally shared among the participating firms. Specifically, let $c(m)$ be the individual cost involved in participating, and assume besides the total cost of achieving the target, $mc(m)$, is decreasing in m:

$$(m\ c\,(m))' = m\ c'\,(m) + c\,(m) < 0. \tag{1}$$

This hypothesis may be seen as representing synergistic effects in the target implementation and/or convexity of individual costs. As for the second interpretation, think of m firms that have to share equally an aggregate abatement level. If their individual abatement costs are increasing in individual abatement, then clearly individual participation costs $c(m)$ are decreasing. If, in addition, individual abatement costs are convex, it is easily shown that total abatement costs $mc(m)$ are decreasing[6].

Moreover, the regulator may consider the achievement of a given target by too few agents is cost-inefficient or not feasible. It is thus assumed that if some participation threshold $w > 1$ is not reached, the VA fails and a mandatory regulation is enforced instead. As a collective threat, the mandatory regulation typically applies to each of the n firms, whatever their individual willingness to undertake voluntary action may be. Conversely, if the threshold is reached, the VA succeeds and only the $m \geq w$ participating firms implement the global target, each of them bearing the cost $c(m)$. It follows firm i's payoffs are defined by:

$$u(s_i, m) = \begin{cases} -c(m) & \text{if } s_i = 1,\ m \geq w \\ 0 & \text{if } s_i = 0,\ m \geq w \\ -t & \text{if } m < w, \end{cases} \tag{2}$$

where $-t$ is the cost to comply to the mandatory regulation. I also assume that a VA is cost-effective, meaning that if at least w firms participate, they prefer, as a whole, the VA to the mandatory regulation: $t > c(w)$. This condition is one of profitability in the sense

[5]Such a cap may be for instance to conform some given share of a sector output with an efficiency standard (e.g. the ACEA agreement), or to reduce targeted agents' toxically releases to some limit value (e.g. the EPA's 33/50 Program).

[6]If the aggregate abatement level is A, and individual abatement cost function is $d(.)$, then total costs are $md\,(A/m)$, and this function is decreasing when $d(.)$ is convex.

of self-enforcing equilibrium,[7] and it has for direct consequence that the Nash solution is non trivial.

Aside from profitability, the participation threshold and the threat parameter are exogenous in this model, so as to keep the analysis as general as possible. However, their setting may be constrained if considered within some given institutional context[8], as illustrated in section 6.

I now turn to the presentation of the correlating device.

2.2. The correlating device

. The device by which participation decisions correlate builds on the basic game presented above. Specifically, suppose each firm is now privately recommended to play either $s_i = 0$ or $s_i = 1$ before she actually makes her decision. Beyond being private, such a recommendation is nominal, i.e. it depends on i. I furthermore assume it results from a random selection process on $S = \{1, 0\}^n$, the set of participation profiles. For ease of presentation, the random selection process can be decoupled as follows. First, it selects $m = k$, the number of firms that participate, according to some discrete distribution denoted by $\{p_k\}_{k \in \{0,\dots,n\}}$, with $p_k \geq 0$ and $\sum_{k=0}^{n} p_k = 1$. Let S_k denotes the set of participation profiles with exactly k participants. Second, contingent upon the realization of $m = k$, each firm i has a probability $p(i, k)$ to participate, with $\sum_{i=1}^{n} p(i, k) = k$. It follows, by construction, that the marginal distribution of firm i's participation is given by:

$$\sum_{k=0}^{n} p(i, k) p_k := Q_i, \tag{3}$$

while $(1 - Q_i)$ represents i's chances not to participate. Please remark that the given of $\{p(i, k), p_k\}_{k=0}^{n} := p$ for all i, is equivalent, in terms of information, to the resulting joint distribution on S. Most importantly, assume such a distribution is public knowledge at the time firms deliberate.

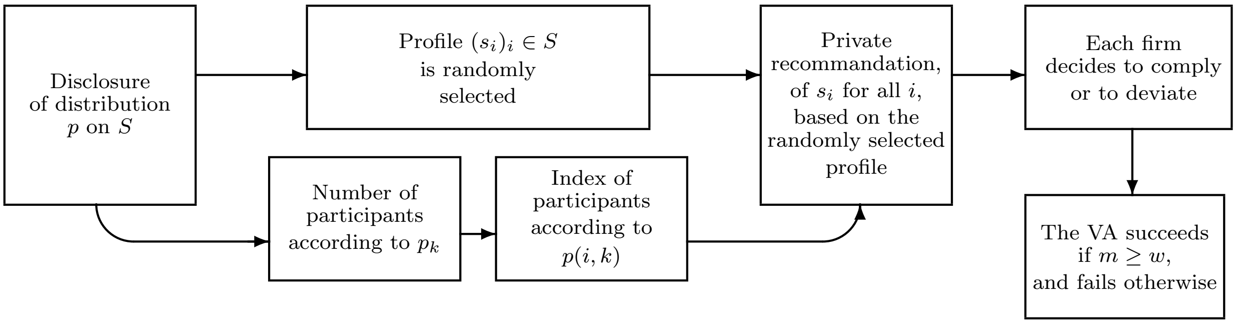
This device is a straightforward application of correlated strategies with private[9] signaling developed by Aumann 2. Combined with the VA, it thus proceeds according to the following timing. After p was disclosed to the n firms, and some participation profile was randomly selected, each firm is only told his individual strategy s_i, from the outcome of the selection. Again, the firm can freely follow or reject the private recommendation, and the VA succeeds as long as w is reached (see figure 1).

[7] A coalition is said to be self-enforcing (d'Aspremont et al. 8, Dawson and Segerson 9) iif participation is (i) profitable, (ii) internally and externally stable. While (ii) is conceptually equivalent to a Nash equilibrium condition, profitability tests for simultaneous (as opposed to unilateral) participating agents' deviations. In public good games, the use of a breaking rule conjointly with a assumption on payoffs such that it is better to participate than not to participate at the threshold (here $t > c(w)$), allow to mimic coalitional incentives within a non-cooperative framework.

[8] Let the mandatory regulation threat be, for instance, a pigouvian tax. In this case, t amounts to $c(n)$ net of the tax payment, which determines, in turn, a minimum profitable value for w (see Dawson and Segerson 9). Such a refinement remains a subcase of the general game.

[9] As will later be shown, some of the forthcoming results may also be reachable through public signaling.

Figure 1: The timing of the correlated VA



Of course, firms' decisions to participate now rely on their *expected* payoffs, inferred from the information provided by the device : the announced p and the recommendation.

2.3. The conditional payoffs

At the point some distribution p is publicly announced, eq. (2) and eq. (3) imply firm i's expected payoffs write:

$$V(i,p) = -t\sum_{k=0}^{w-1} p_k - \sum_{k=w}^{n} p_k \left(p(i,k)c(k)\right), \qquad \forall i. \tag{4}$$

Then, based on its private prescription, each firm may be able to revise the distribution on S according to its new piece of information. Let $V(i,p,s_i)$ denote i's expected payoffs contingent upon prescription s_i, for a given announced p. Using Bayes' theorem, I obtain:

$$V(i,p,0) = -t\sum_{k=0}^{w-1} p_k \left(\frac{1-p(i,k)}{\sum_{l=0}^{n} p_l(1-p(i,l))}\right) \qquad \forall i \tag{5a}$$

$$V(i,p,1) = -\sum_{k=0}^{w-1} p_k \frac{p(i,k)}{\sum_{l=0}^{n} p_l p(i,l)} t - \sum_{k=w}^{n} p_k \frac{p(i,k)}{\sum_{l=0}^{n} p_l p(i,l)} c(k) \qquad \forall i, \tag{5b}$$

where the ratio $\frac{p(i,k)}{\sum_l p_l p(i,l)}$ reads as the probability k firms participate, knowing firm i is one of them. Please remark for later uses, that expected payoffs (4), (5a) and (5b) all continuously depend on the threat level, t.

. I now turn to the equilibria analysis.

3. The outcome of the uncorrelated VA

The present section investigates, as a benchmark case, the voluntary effort when there is no burden sharing process. Such a scenario is captured by the basic game of VA, in which the n firms play a Nash equilibrium. Indeed, without preplay prescription or devised lottery on the set of participation profiles, each firm may independently mix over $\{0,1\}$ according to its own strategy, $\{Q_i,(1-Q_i)\}_{i=1}^n$. As the game is finite and symmetric, it is known there alway exists[10] a symmetric mixed NE, and that it is the mixed strategy profile such that the individual participation probability, Q (where index i is dropped for symmetry), satisfies:

$$\binom{w-1}{n-1} Q^{w-1}(1-Q)^{n-w} t = \sum_{k=w-1}^{n-1} \binom{k}{n-1} Q^k (1-Q)^{n-1-k} c(1+k), \tag{6}$$

which is the algebraic form of the condition that to contribute and not to contribute must yield the same expected gains for each agent. A general characterization[11], i.e. which also includes partial supports, is provided in Appendix A. From condition (6), I establish the following proposition:

Proposition 1. *The uncorrelated VA has a unique symmetric mixed NE, given by* $\{\hat{Q}(n,t),(1-\hat{Q}(n,t))\}$*, with* $\frac{\partial \hat{Q}(n,t)}{\partial t} > 0$*,* $\hat{Q}(n,0)=0$*, and* $\lim_{t\to+\infty}\hat{Q}(n,t)=1$*, where* $\hat{Q}(n,t)$ *is defined as the inverse function of:*

$$t(Q) = \sum_{k=w-1}^{n-1} \frac{\binom{k}{n-1}}{\binom{w-1}{n-1}} c(1+k)\left[\frac{Q}{(1-Q)}\right]^{k-w+1}. \tag{7}$$

Proof 1. Algebraic manipulations lead us to rewrite (A.4) *with* $j=m=0$ *as* $t(Q)$*, which is strictly increasing in* q*:*

$$\frac{\partial}{\partial q} t(Q) = \sum_{k=w-1}^{n-1} \frac{\binom{k}{n-1}}{\binom{w-1}{n-1}} c(1+k)\left[\frac{1}{(1-Q)^2}\right]^{k-w+1} > 0. \tag{8}$$

Since $t(0)=0$ *and* $\lim_{q\to 1} t(Q)=+\infty$*, it follows* $t(Q)$ *is invertible on our interval of interest,* $Q\in[0,1]$*.*

As regards the pure NE of the game, it is easily shown such equilibria are multiple by using an argument similar as in Palfrey and Rosenthal 21. First, observe any strategy profile such that $m<w-1$ is a pure NE, since a unilateral deviation will not affect the VA status nor the corresponding payoffs (agents are not pivotal). Conversely, for any profile such that $m=w-1$, each non-participating firm has a unilateral incentive to deviate and pay $c(w)$ instead of t. Likewise, for any profile such that $m>w$, each participating firm has a unilateral incentive not to participate since it will not affect the VA status but will avoid her the participation cost. Finally, all the profiles such that $m=w$ are pure NE since any deviation of a non-participating firm will trigger tax enforcement, while non-participating firms have no interest in participating knowing that the VA is provided anyway. These results can be summarized as follows.

[10] See **?** for a complete proof of the existence of symmetric NE in finite and symmetric games.

[11] These results are an extension of Palfrey and Rosenthal 21 to subscription games that feature our more general payoffs structure.

Remark 1. *If $n \geq 2$ and $w \geq 2$, the $|\hat{S}| = \binom{w}{n} + \sum_{k=0}^{w-2} \binom{k}{n}$ pure NE of the VA without the correlating device are given by $\hat{S} = \{S_w, (S_k)_0^{w-2}\}$, none of which correspond to the socially optimal allocation, $(1)_{i=1}^{n}$. Otherwise, $|\hat{S}| = \binom{w}{n}$ and $\hat{S} = \{S_w\}$.*

Remark also that, by the assumption on costs, participation profiles can be Pareto-ranked simply based on m. In particular, the full participation profile maximizes the deterministic aggregate payoff. Two questions then arise from these preliminary results.

Even though pure NE such that the VA succeeds have been identified, a notable issue is, of course, one of efficiency. But primarily, multiplicity even raises the question of the VA feasibility : how will firms coordinate amongst the several subsets of Pareto equivalent pure NE ?

. The rest of the paper precisely addresses these issues.

4. Characterizing the optimal correlated equilibrium

In this section, I study the set of CE under the correlating device, and characterize the distribution that ensures the most efficient expected welfare.

4.1. The set of CE

Still using Myerson 17 interim definition of correlated equilibria, the strategic incentive constraints can be written as follows. When agents are privately told to play $s_i = 0$, the distribution that was prealably announced must be such that they have no incentive to unilateraly deviate, ie.

$$-\sum_{k=0}^{k\leq w-1} p_k \frac{(1-p(i,k))}{\sum_{l=0}^{n}(1-p(i,l))\,p_l} t - \sum_{k=w}^{n-1} p_k \frac{(1-p(i,k))}{\sum_{l=0}^{n}(1-p(i,l))\,p_l} 0 \geq -\sum_{k=0}^{k\leq w-2} p_k \frac{(1-p(i,k))}{\sum_{l=0}^{n}(1-p(i,l))\,p_l} t - \sum_{k=w-1}^{n-1} p_k \frac{(1-p(i,k))}{\sum_{l=0}^{n}(1-p(i,l))\,p_l} c(k+1)$$

must holds for all i, which simplifies to:

$$p_{w-1}\,(1-p(i,w-1))\,(t-c(w)) - \sum_{k=w}^{n-1} p_k\,(1-p(i,k))\,c(k+1) \leq 0. \tag{9}$$

Such a condition ensures agent i's expected payoff is higher if she complies than if she decides to participate, provided her prior (i.e. the announced distribution), has been revised taking into account the fact that the profile randomly selected necessarily features $s_i = 0$. Likewise, when the prescription is $s_i = 1$,

$$-\sum_{k=1}^{k\leq w-1} p_k \frac{p(i,k)}{\sum_{l=0}^{n} p(i,l)\,p_l} t - \sum_{k=w}^{n} p_k \frac{p(i,k)}{\sum_{l=0}^{n} p(i,l)\,p_l} c(k) \geq -\sum_{k=1}^{k\leq w} p_k \frac{p(i,k)}{\sum_{l=0}^{n} p(i,l)\,p_l} t - \sum_{k=w+1}^{n} p_k \frac{p(i,k)}{\sum_{l=0}^{n} p(i,l)\,p_l} 0$$

must hold for all i, which simplifies to:

$$p_w p\,(i,w)\,(t - c(w)) - \sum_{k=w+1}^{n-1} p_k p\,(i,k)\, c(k) - p_n c(n) \geq 0. \tag{10}$$

Again, condition (10) guarantees agent i's expected payoff is higher if she complies than if she decides not to participate, provided her prior has been revised taking into account that the profile the regulator selected necessarily features $s_i = 1$. Finally, the probability constraints are given by:

$$\begin{cases} p(i,k) \geq 0 \quad \forall i, \qquad k \in \{0,...,n\}, \\ \sum_{k=0}^{n} p_k = 1, \\ \text{ and } \sum_{i=1}^{n} p(i,k) = k, \qquad k \in \{0,...,n\}. \end{cases} \tag{11}$$

Conjointly with the $2n$ inequalities defined by (9) and (10), for all i, these probability constraints fully characterize the set of CE of the correlated VA. Remark that the set of CE is by definition self-enforcing or, in other words, voluntary.

4.2. The optimal CE

In a normative perspective, one might be interested in implementing p such that the social welfare is maximized. Remark that the aggregate payoffs is actually given by:

$$\begin{aligned} \sum_{i=1}^{n} V(i,p) &= -nt \sum_{k=0}^{w-1} p_k - \sum_{k=w}^{n} p_k \sum_{i=1}^{n} p(i,k) c(k) \\ &= -nt \sum_{k=0}^{w-1} p_k - w p_w c(w) - \sum_{k=w+1}^{n} k p_k c(k) \end{aligned} \tag{12}$$

Such a distribution should be incentive compatible to be workable, i.e. it has to verify, for all i, the conditions (9) and (10) characterized above. The maximization problem therefore writes :

$$\max_{\{p_k, p(i,k)\}_0^n} - nt \sum_{k=0}^{w-1} p_k - w p_w c(w) - \sum_{k=w+1}^{n} k p_k c(k) \tag{13a}$$

$$\text{st.} \qquad (9), (10) \quad \text{and } (11). \tag{13b}$$

This non-linear program becomes solvable as soon as one remark that the $p(i,k)$ are not in the objective. Indeed, as firms are symmetric, the repartition of the burden within firms for a given k, does not impact the social welfare. In other words, the program (13) can be decomposed into two separable subproblems. In a first step, summing conditions (9) and (10) over i, I obtain two necessary conditions, which defines a reduced program that is solvable by linear programming. In a second step, I substitute the solutions obtained for $\{p_k\}_{k=0}^n$ into (13), and solve it in $p(i,k)$.

Step 1. Thus, summing (9) and (10) over i, and using that $\sum_{i=1}^{n} p(i,k) = k$, I obtain respectively:

$$
\begin{cases}
p_{w-1}\,(n-w+1)\,(c(w)-t) + \displaystyle\sum_{k=w}^{n-1}(n-k)p_k\ c\,(k+1) & \geq 0 \qquad \text{(14a)} \\
wp_w\,(t-c(w)) - \displaystyle\sum_{k=w+1}^{n} kp_k c\,(k) & \geq 0 \qquad \text{(14b)}
\end{cases}
$$

One is now able to recognize a linear programming problem, which can be converted into its augmented form by introducing two slack variables, denoted x_1 (into constraint (14a)), and x_2 (into constraint (14b)). Note that the generated standard program is only composed of $j = 3$ equality constraints and $l = n + 2$ variables, since the symmetry of firms implies the objective is actually optimized in $\{p_k\}_0^n$. Such a program can be solved by applying the two-phase simplex algorithm (Padberg 20). However, this method may imply numerous (and cumbersome) iterations if started without proceeding to a preliminary heuristic analysis. The next proposition states the result of the program, while the proof provides both an intuitive and a formal (Appendix B) argument.

Proposition 2. *The optimal CE is such that $\{p_k^*\}_{k=0}^n$ recommends participation either to everyone, $p_n^* = \frac{w(t-c(w))}{w(t-c(w))+nc(n)}$, either to w firms, $p_w^* = \frac{nc(n)}{w(t-c(w))+nc(n)}$.*

Proof 2. First, observe that (14a) and (14b) do not depend on $\{p_k\}_0^{w-2}$. Since $t >$ ¸(w), it is straightforward the optimal distribution must assign a probability 0 to the corresponding participation profiles. Then, remark that p_{w-1} appears in (14a) solely, and that it does not need to be strictly positive. However, as p_w must be strictly positive for (14b) to hold, we know (14a) cannot be equal to 0 at the optimum. Finally, $\frac{\partial mc(m)}{\partial m} < 0$ implies (14b) will be binding, since it is now obvious the highest probability should be put on profile $(1)_i$. We therefore know $\{x_1,\ p_w,\ p_n\}$ is the basis which has to be tested in order to confirm our heuristic solution is optimal. This is done in Appendix B.

Step2. Then, substituting $\{p_k^*\}_{k=0}^n$ into the general program, the latter rewrites :

$$
\max_{\{p(i,w)\}_{i=1}^n} -p_w^* c(w) \sum_{i=1}^{n} p(i,w) - n(1-p_w^*)c(n) \tag{15a}
$$

$$
\text{st.} \qquad p(i,w) \geq \frac{(1-p_w^*)c(n)}{p_w^*(t-c(w))} = \frac{w}{n} \quad \forall i, \qquad \text{and} \quad \sum_{i=1}^{n} p(i,w) = w. \tag{15b}
$$

where I used the fact that $p_w^* + p_n^* = 1$. It is now straightforward the solution is $p^*(i,w) = \frac{w}{n}$ for all i, meaning that the optimal marginal distributions are symmetric.

Proposition 3. *Under the optimal CE, the probability to participate and not to participate are given, for all i, by $1 + p_w^* \left(\frac{w}{n} - 1\right)$ and $p_w^* \left(1 - \frac{w}{n}\right)$, respectively.*

Proof 3. $\sum_k p^*(i,k)p^*_k = p^*(i,w)p^*_w + p^*_n = 1 + p^*_w(p^*(i,w) - 1)$ *where* $p^*(i,w) = \frac{w}{n}$.

The best self-enforcing distribution, $\{p^*(i,k), p^*_k\}_{k=0}^n := p^*$, that is implementable by the correlating device is such that each firm's fear of being pivotal (and loosing $(t - c(w))$), just compensates for its fear of being superfluous on the full participation profile (and loosing $c(n)$). Of course, such a trade off depends on the threat, t, and the number of firms.

. Next section performs a comparative efficiency analysis and derives practical results about the way participation could be enhanced through the use of the correlating device.

5. A social welfare analysis

I show that not only the correlating device solves the coordination issue, but it allows to reach a higher social welfare than the uncorrelated VA.

5.1. Coordination and efficiency

A general result of Aumann (Aumann 3, 2 and Nisan et al. 18) states that any convex combination of Nash equilibria is a correlated equilibrium. It has two main normative implications in this framework, which the next proposition details.

Proposition 4. *The correlating device allows : (i) firms to coordinate on any NE of the uncorrelated VA. In particular, the participation profiles that are pure NE can be implemented through public signaling. (ii) to achieve, for all level of threat t, a higher social welfare than the uncorrelated VA, given by:*

$$nV(p^*) = -n\frac{tc(n)}{(t - c(w)) + \frac{n}{w}c(n)}, \tag{16}$$

Proof 4. See Appendix C.

Any joint distribution derived from a CE can be implemented by the correlating device since it fulfills both (9) and (10) as well as the probability constraints. In practice, some third party can perform the drawing and the private prescriptions. Depending on the institutional background, it may be an industry association (negotiated VAs) or the regulatory body itself (public VAs).

In particular, the distribution that assigns $1/\binom{w}{n}$ for all $(s_i)_i \in S_w$ and 0 otherwise, is a CE that can be implemented by public signaling. Indeed, one can easily check that the incentive constraints rewrite $(n-w)p_w c(w+1) \geq 0$ and $wp_w(t - c(w)) \geq 0$, with $p_w = 1$. Moreover, as each profile that may be drawn is a pure NE, the equilibrium is robust to prescriptions' disclosure. This holds for any distribution over a subset of strong Nash equilibria (see Moreno and Wooders 11 for a general demonstration).

Public signaling may for instance take the form of published lists, but it can also plausibly be argued to be exogenous, and thus to represent inter-firms communication arising from self-regulation initiatives in the absence of a proper third-party. The present framework is therefore fully relevant to analyse preemptive VAs in all its variations, including the case of self-regulation.

Another practical result of interest regards preemption:

Corollary 1. *When the optimal CE is implemented, the VA always succeeds. In contrast, it fails with probability* $\sigma(t,n) = \sum_{k=0}^{w-1} \binom{k}{n} \hat{Q}^k (1-\hat{Q})^{n-k}$ *under the mixed NE, in which case the mandatory regulation is implemented.*

This last statement directly follows from proposition 2 and the definition of $\hat{Q}(n,t)$.

5.2. Comparative statics

In the rest of this section, I investigate the effects of the number of firms and the tax threat, both on free-riding and social welfare.

Free-riding. In participation games, the expected participation rate provides a direct measure for free-riding, defined as the ratio of participating firms. Remark that under the optimal CE, participation profiles with the same number of participating firms are equiprobable (see propositions 2 and 3). Likewise, by definition of mixed NE, $\hat{Q}(n,t)$ induces a probability $\binom{n}{k}\hat{Q}^n(1-\hat{Q})^{n-k} = \frac{\hat{p}_k}{k}$ that k firms participate to the VA. Some elementary combinatorics then show that $\hat{p}(i,k) = \binom{k-1}{n-1}/\binom{k}{n} = \frac{k}{n}$ for all i.

It follows that the mixed NE induces the distribution $\{\frac{k}{n}, \hat{p}_k\}_{k=0}^n := \hat{p}$, and that the following can be stated:

Remark 2. *The expected participation rates under the VA with the mixed NE or the optimal CE, verify* $\frac{1}{n}\left(\sum_k kp_k\right) = Q$ *in both cases.*

Proof 5. Under p^* *and* $\hat{p}$*, it holds for all firm* i *that :* $p(i,k) = \frac{k}{n}$, $k \in \{0,...,n\}$. *It follows* $\frac{1}{n}\left(\sum_k kp_k\right) = \sum_{k=0}^n p(i,k)p_k$*, which is the marginal distribution of participation.*

This statement[12] implies an analysis of free-riding can be provided simply by studying

$$Q^*(n,t) = \frac{w}{n}p_w + p_n^* \quad \text{and} \quad \hat{Q}(n,t) = \sum_{k=0}^n \frac{k}{n}\hat{p}_k.$$

Note, for comparison purpose, that the pure NE's participation rate of a succeeding VA is given by $\frac{w}{n}$. It follows the tax threat has no impact on the participation rate of the pure NE (as long as $t > c(w)$), but it decreases the free-riding both under the optimal CE:

$$\frac{\partial Q^*(n,t)}{\partial t} = \frac{\partial}{\partial t}\left(1 + p_w^*\left(\frac{w}{n} - 1\right)\right) = \frac{tnc(n)\,(n-w)}{n\,(w\,(t - c\,(w)) + nc(n))^2} > 0.$$

[12] Notice that this results holds for any distribution on S such that profiles with the same number of participating firms are equiprobable, and that such a restriction is not as strong (namely sufficient yet not necessary) as assuming firms' individual participation probabilities are symmetric.

and under the symmetric mixed NE (see that eq. (8) in proposition 1 is positive).

Then, the effects of a rise of n on the incentives to rely on voluntary efforts of other firms can be described as follows. Under the pure NE, free-riding unambiguously increases in n : $\frac{\partial}{\partial n}\frac{w}{n} = -\frac{w}{n^2} < 0$. However, under the optimal correlated VA, while a rise of n directly enhances the incentive to free-ride (firms are less afraid of being pivotal, $\frac{\partial}{\partial n}p^*(i,w) < 0$), this effect is mitigated by the decrease of $nc(n)$ (the losses arising from being superfluous are lower):

$$\frac{\partial}{\partial n}\left(\frac{w}{n}p_w^* + (1-p_w^*)\right) = \frac{\overbrace{w\left(w-n\right)c'(n)\left(t-c(w)\right)}^{>0} + \overbrace{wc(n)\left[\left(c(w)-t\right)-c(n)\right]}^{<0}}{\left(w\left(t-c(w)\right)+nc(n)\right)^2}$$

As regards $\hat{p}$, the reader can report to Hong and Lim 15 for a similar analysis.[13] Some results will also be detailed in the specified example.

Social welfare. Obviously, the aggregate payoff does not depend on the threat under the pure NE. In contrast:

Corollary 2. *As the level of threat intensifies, the expected aggregate payoff under p^* increases and tends toward the level of aggregate welfare generated by the full participation profile, $-nc(n)$.*

Proof 6.

$$\frac{\partial}{\partial t}nV(p^*) = \frac{-wnc\left(n\right)+w^2tnc\left(n\right)}{w\left(t-c\left(w\right)+nc\left(n\right)\right)^2} > 0 \qquad \textit{and} \qquad \lim_{t\to+\infty}\frac{w\left(t-c(w)\right)}{w\left(t-c(w)\right)+nc(n)} = 1.$$

Furthermore, even though the effects of a change in t are non monotonic under the mixed symmetric NE, the expected aggregate payoffs can be shown to converge towards $nc(n)$. Hence, the following result can be stated:

Corollary 3. *The difference between the expected aggregate payoffs generated by the optimal CE and the symmetric NE decreases in the level of threat from $t \geq t_1$:*

$$\lim_{t\to+\infty} n\left(V(p^*)-V(\hat{p})\right) = 0, \qquad t \geq t_1,$$

where t_1 is defined as the minimum threat level beyond which an increase of t leads to an increase of $V(\hat{p})$.

Proof 7. Observe that: $\partial\hat{p}_k/\partial\hat{Q} = k\hat{Q}^{k-1}(1-\hat{Q})^{n-k} - (n-k)(1-\hat{Q})^{n-k-1}$, hence the probability on profiles in S_k increases in $\hat{Q}(t,n)$ if and only if $k > n\hat{Q}(t,n)$, the expected participation, and $n\frac{\partial}{\partial t}\hat{Q}(t,n) > 0$. Thus, an increase of the threat induces two effects on

[13] Hong and Lim 15 provides an analysis for free-riding in the case of a lumpy good provision with constant total costs.

the cost of failure, $t\sigma(n,t)$. On one hand, $\sigma(n,t)$ decreases for the benefit of more efficient participation profiles, which increases $V(\hat{p})$. On the other hand, the tax applied in case of failure is higher, which lowers $V(\hat{p})$. Appendix D shows that from some $t_1 > c(w)$, the first effect dominates the second one. Moreover, proposition 1 implies $\lim_{t\to+\infty}\hat{Q}(n,t)^n = 1$, hence:

$$\lim_{t\to+\infty} -\sum_{w}^{n} k\hat{Q}(n,t)^k(1-\hat{Q}(n,t))^{n-k}c(k) = \lim_{t\to+\infty} nV(p^*) = -nc(n).$$

I close this comparative static analysis by a quick study of the effect of n on average expected aggregate payoffs :

$$\frac{\partial}{\partial n}V(p^*) = -tw\frac{c'(n)w\,(t-c\,(w)) - c(n)^2}{(w\,(t-c\,(w)) + nc\,(n))^2} > 0 \qquad \text{and} \qquad -\frac{\partial}{\partial n}\frac{c(w)}{n} = \frac{c(w)}{n^2} > 0.$$

In other words, even if there is more incentives for free-riding, which may decrease firms individual participation probabilities: (i) the probability on full participation under the optimal CE does increase, $\frac{\partial p_n^*}{\partial n} > 0$, while (ii) under the pure NE, the VA still succeeds at the threshold, which implies decreasing average cost.

6. A numerical example

I apply the model of VA to a specified pollution abatement model. I numerically illustrate the general results stated in the previous sections.

6.1. The pollution abatement model

Consider n symmetric firms producing a good and pollutant emissions which are engaged in Cournot competition. Let $\pi\,(e_i)$ be the indirect profit function, with $\frac{\partial^2\pi(e_i)}{\partial^2 e_i} \leq 0$. Specifically, assume that:

$$\pi\,(e_i) = ae_i - be_i^2 + k, \tag{16}$$

with $a, b > 0$ and a constant $k \geq 0$. The optimal level of emission at laisser-faire is $e^{LF} = {}^{a}/_{2b}$. Let $n\varepsilon$ be the emission target, where $\varepsilon \in [0, e^{LF}]$, and which amounts to an aggregate abatement $n(e^{LF} - \varepsilon)$. Assume the regulator chooses w such that it is feasible, i.e. formally $w = \lceil n(1 - {}^{\varepsilon}/_{e^{LF}})\rceil$. As in the general participation game, I denote the threat by t. It follows the payoffs of firm i in the basic game are given by :

$$u(s_i, m) = \begin{cases} -c(m) = -\frac{n^2(a-2b\varepsilon)^2}{4bm^2} & s_i = 1 \text{ and } m \geq w \\ 0 & s_i = 0 \text{ and } m \geq w \\ -t & m < w \end{cases} \tag{17}$$

where $c(m)$ is the participation cost which, by the concavity of indirect profits, decreases in m. Following the general game, $c(m)$ is obtained as the difference between the profit at laisser faire and the profit when m firms participate :

$$c(m) = \pi\left(e^{LF}\right) - \pi\left(\frac{n\varepsilon - (n-m)\,e^{LF}}{m}\right) = \frac{n^2(a-2b\varepsilon)^2}{4bm^2}.$$

Likewise, t is the difference between the profit at laisser faire and the profit under the mandatory regulation scheme. This illustration also requires I specify what would be the mandatory regulation in this context. I assume it is a two-part scheme, primarily composed of a pigouvian tax, i.e. a tax on emissions set at a level t^{pig} such that each of the n firms produces $e_i = \left(e_i^{LF} - \varepsilon\right)$ and the target $n\varepsilon$ is reached:

$$t^{pig} = \left.\frac{\partial \pi\left(e_i\right)}{\partial e_i}\right|_{e_i=\varepsilon} = a - 2b\varepsilon.$$

The second part of the scheme is a lump-sum transfer, which may be interpreted as transaction costs, denoted τ. It follows the expression for t:

$$t = \pi(e^{LF}) - \pi\left(\varepsilon\right) + t^{pig}\varepsilon + \tau = \frac{a^2}{4b} - b\varepsilon^2 + \tau. \tag{18}$$

I assume furthermore $\tau \in [\tau_{min}, \tau_{max}]$, where τ_{max} corresponds to a zero profit condition under the pigouvian tax :

$$\tau_{max} = \pi\left(\varepsilon\right) - t^{pig}\varepsilon = \frac{4b^2\varepsilon^2}{4b} + k. \tag{19}$$

The lower bound τ_{min} stands for the minimum tax level such that the profitability condition is satisfied at w, defined as the feasible threshold:

$$\pi\left(e^{LF}\right) - \pi\left(e^{LF} - \frac{n\left(e^{LF} - \varepsilon\right)}{\lceil n(1 - \varepsilon/e^{LF})\rceil}\right) > t \qquad \Rightarrow \qquad \tau_{min} = b\varepsilon^2 \tag{20}$$

if $\lceil n(1-\varepsilon/e^{LF})\rceil = n(1-\varepsilon/e^{LF})$. I can now calculate the optimal CE, the unique symmetric mixed NE and the social welfare associated with the optimal CE, as functions of n, w and the target $n\varepsilon$.

6.2. The symmetric mixed NE and the optimal CE:

The optimal CE of the game is defined by:

$$p_n^* = \frac{\left(\frac{a^2}{4b} - b\varepsilon^2 + \tau\right)w - \frac{1}{b}\frac{n^2(a-2b\varepsilon)^2}{4w}}{\left(\frac{a^2}{4b} - b\varepsilon^2 + \tau\right)w - \frac{1}{b}\frac{n^2(a-2b\varepsilon)^2}{4w} + \frac{(a-2b\varepsilon)^2}{4b}} \qquad \text{and} \qquad p_w^* = 1 - p_n^*.$$

The induced expected social welfare is given by:

$$nV(p^*, \tau) = n\frac{w(a-2b\varepsilon)^2(\frac{a^2}{4b} + \tau - b\varepsilon^2)}{(1-\frac{n}{w})n(a-2b\varepsilon)^2 + 4bw\left(\frac{a^2}{4b} - b\varepsilon^2 + \tau\right)}.$$

Finally, I characterize the symmetric mixed NE:

$$t(Q) = \sum_{w-1}^{n-1}\left(\frac{(w-1)!\,(n-w)!}{k!\,(n-1-k)!}\right)\frac{n^2(a-2b\varepsilon)^2}{4b(k+1)^2}\left(\frac{Q}{1-Q}\right)^{k-w+1}. \tag{21}$$

and $t = \tau - b\varepsilon^2 + \frac{a^2}{4b}$, which yield the inverse function $\hat{Q}(n, t(\tau))$.

6.3. The numerical results

The results are based on an example with $a = 40$ and $b = 2$. The optimal emission level at laisser faire is $e^{LF} = 10$ which yields $\pi\left(e^{LF}\right) = 200 + k$. Before proceeding to the numerical analysis, please remark that eq. (21) rewrites as an hypergeometric function, $t(Q) = \frac{n^2(a-2b\varepsilon)^2}{w^2 4b}\,{}_3H_2[\{w-n,1,w\}\{1+w,1+w\},\frac{Q}{Q-1}]$. Using the Euler integral transform rule, the latter can be decomposed as a function of ${}_2H_1[\{w-n,1\}\{1+w\},\frac{Q}{Q-1}]$ as follows:

$${}_3H_2\left[\{w-n,1\}\{1+w\},\frac{Q}{Q-1}\right] = w^2\int_0^1 x^{w-1}\underbrace{\int_0^1 (1-x)^{w-1}(1+x)^{w-1}\left(\frac{Q}{1-Q}\right)^{n-w}}_{=\frac{\Gamma(w+1)}{\Gamma(1)\Gamma(w)}{}_2H_1[\{w-n,1\}\{1+w\},\frac{Q}{Q-1}]} dxdx$$

The probability to participate of the mixed NE is now analytically solvable as a function of the threat τ, thus I am able to fully study the static comparative with respect to n.

The analysis is first performed for an emission target $\varepsilon = 4$ and $n = \{5, 10\}$. The cost of the mandatory regulation is $t(\tau) = 168 + \tau$, where the pigouvian tax per unit of emission is $t^{pig} = 24$, while $\pi(\varepsilon) = 128 + k$. Since participation is feasible for $w = \frac{3n}{5}$, the profitability condition requires a minimum level of $\tau > \tau_{min} = 32$. Likewise, the threat can not exceed the profits under the mandatory regulation, i.e. $\tau_{max} = 32 + k$.

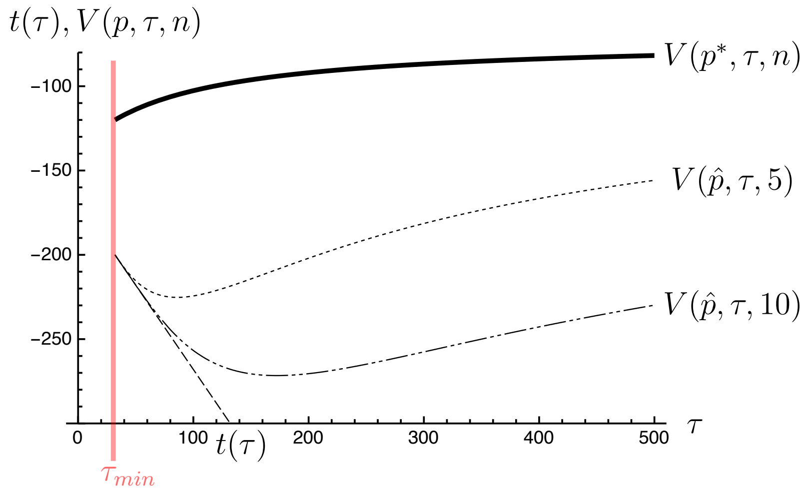
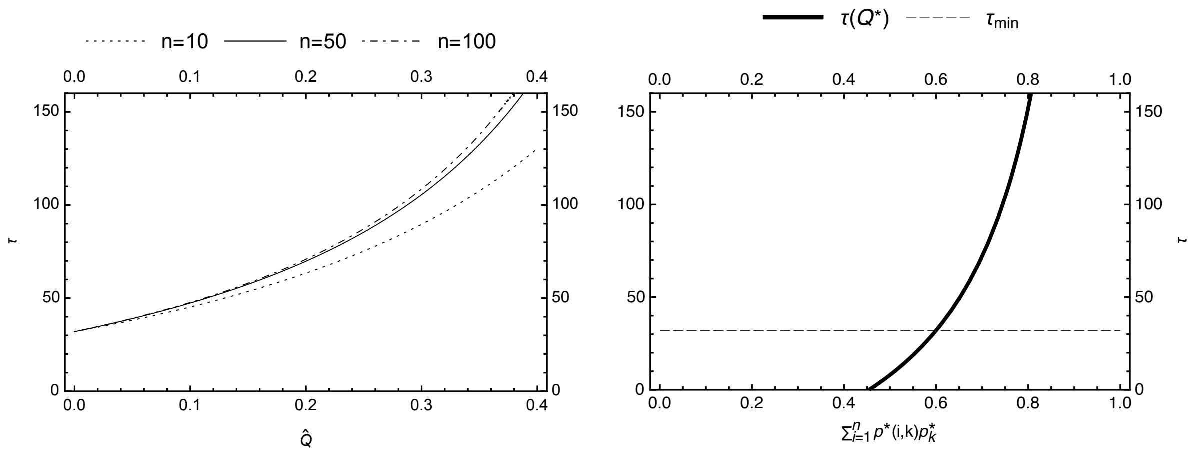


Figure 2: The probability to participate under the mixed NE and the optimal CE as the number of firms rises from 10 to 100.

. Figure 2 shows the participation rates and the expected payoffs under the optimal CE, the mixed NE and the mandatory regulation, expressed as functions of the punishment parameter $\tau \in [\tau_{min}, \tau_{max}]$. A first observation is that both the optimal CE (thick line) and the mixed NE (dotted lines) generate higher payoffs than the mandatory regulation (dashed lines) within $[\tau_{min}, \tau_{max}]$, since by assumption: $t > c(w)$. A second observation is that the expected payoff per player under the optimal CE does not depend on the number of player. Two explanations are: (i) since the target is expressed in terms of individual and equally shared burden, an increase of n does not imply a decrease of the individual abatement (hence nor $c(n) = \frac{(a-2b\varepsilon)^2}{4b} = 72$, neither $t(\tau)$ (see eq. 18), depend on n) (ii) by definition of feasibility, the participation threshold w is linear in n (hence $p^*(i,w) = \lceil (1 - \varepsilon/e^{LF}) \rceil$ is a constant). Then, observe that the individual expected payoffs

Figure 3: The probability to participate under the mixed NE and the optimal CE as the number of firms rises from 10 to 100.



under the mixed NE is lower when $n = 10$, which highlights the efficiency losses due to free-riding. Specifically:

$$\int_0^1 x^{n\left(1-\frac{\varepsilon}{e^{LF}}\right)-1} [\int_0^1 \left(n\left(1 - \frac{\varepsilon}{e^{LF}}\right) - 1\right)(1-x)^{n\left(1-\frac{\varepsilon}{e^{LF}}\right)-2}\left(1 + x\frac{Q}{1-Q}\right)^{n\frac{\varepsilon}{e^{LF}}} dx$$
$$+ \int_0^1 \left(n(1 - \frac{\varepsilon}{e^{LF}}) - 1\right) x^{n\left(1-\frac{\varepsilon}{e^{LF}}\right)-2} dx \int_0^1 (1-x)^{n\left(1-\frac{\varepsilon}{e^{LF}}\right)-1}\left(1 + x\frac{Q}{1-Q}\right)^{n\frac{\varepsilon}{e^{LF}}} dx$$
$$+ \int_0^1 n\frac{\varepsilon}{e^{LF}}(1-x)^{n\left(1-\frac{\varepsilon}{e^{LF}}\right)-1}\left(1 + x\frac{Q}{1-Q}\right)^{n\frac{\varepsilon}{e^{LF}}-1} dx]\, dx = \frac{\partial}{\partial n}\frac{{}_3H_2}{w^2}. \geq 0$$

As a result, it can be concluded that:

$$\frac{\partial t(Q,n)}{\partial n} = \frac{\partial}{\partial n}\left(\frac{n^2(a-2b\varepsilon)^2}{w^2 4b} \quad {}_3H_2\right) \geq 0.$$

This means, as illustrated by figure 3, that $\hat{Q}(n, t(\tau))$ decreases in the number of firms for all τ, i.e. induces free-riding.

On the other hand, notice that a more stringent threat induces a higher expected participation $nQ(n, t(\tau))$, i.e. deters free-riding. However, as shown in the general case for $n > 30$ (see corollary 3), an increase of participation under the mixed NE does not necessarily involve higher expected aggregate payoffs though, since the rise of τ also increases agents' costs in case of failure. Table 1 gives probability that the uncorrelated VA fails, $\sigma(t(\tau), n)$, for several values of the tax threat.

Threat parameter, τ	32	50	100	150	200	500	$+\infty$
n=5 firms	1	0.974	0.721	0.507	0.370	0.107	0
n=10 firms	1	0.999	0.926	0.770	0.622	0.226	0

Table 1: Probability of failure of the VA under the mixed NE, when $\varepsilon = 4$.

Thus, depending on the number of players and $\tau_{max}(k)$, the use of deterrence may be inefficient if the VA is not augmented with private prescriptions. In particular, the background threat induces a gain of expected payoffs from $\tau = 84$ (respectively $\tau = 171$) when $n = 5$ (respectively, $n = 10$). The use of the correlating device is therefore especially recommended when profits strongly rely on the level of emissions to be regulated. Namely, here, if k is such that $\tau_{max}(k) \leq t_1$.

7. Summary and concluding remarks

This study developed a subscription game with a payoff structure representing the participation incentives of firms that face a preemptive VA. Two key features are the exogenous sectoral target (modeled by a minimum participation threshold, w) and the use of a tax threat with a collective liability rule ($t > c(w)$ applies to each firm if the VA fails). Most importantly, this basic VA game is augmented with a correlation device, cast as mimicking the burden sharing process. All the results are demonstrated in the general n-player case.

First, the unique symmetric mixed NE and the pure NE of the basic VA game are characterized, and it is pointed out they can be implemented by the correlating device. Such a device may be embodied by an industry association, the regulator itself or, in the case of the pure NE, some exogenous source of public signal.

Then, the paper shows the correlating device does not only solve the problem raised by multiplicity, but it also ensures efficiency gains. Specifically, the best CE is such that it randomly implements either the full or the minimum participation level so as to optimally take advantage of the fear for each firm to be pivotal. A related finding is that the optimal CE yields, for all $t > c(w)$, a higher aggregate payoff than the symmetric mixed NE. In addition, the free-riding is shown to decrease in the threat level under both distributions. One last result is that, from some tax level $t_1 > c(w)$, the relative efficiency gain of the optimal CE decreases in t. Nevertheless, this claim must be qualified by the fact that a credibility requirement would actually limit the regulator in his choice of the threat stringency (section 6).

Remark that, conversely to standard cooperative (Foley 11, Rosenthal 22) or mechanism design approaches (e.g. Groves and Ledyard 14, Clarke 7) to the burden sharing process, the mechanism developed in this paper does not require a pre-stage of adherence.[14]

Two research directions are left for further work. First, remark there is some potential for considering asymmetric firms in this framework, where it may be assumed that the third party only know the agents' type distribution.

It would also be worth studying experimentally correlated strategies of participation to a public good game, as an extension of Hong and Lim 15's protocol, where a direct application could be license plate-based driving restrictions.

Appendix

Appendix A. Mixed strategies Nash equilibria of the basic game

As in Palfrey and Rosenthal 21, we restrict our analysis to the cases such that there are j agents with a strategy support $\{0\}$ and m agents with a strategy support $\{1\}$. The rest of agents is mixing in the support $\{0,1\}$ according to a symmetrical distribution :

$$p_i(s_i) = \begin{cases} Q & \text{if } s_i = 1 \\ 1-Q & \text{if } s_i = 0 \end{cases}. \tag{A.1}$$

with S strictly positive, and (m, j, w, n) an admissible vector of parameters such as defined in Palfrey 1984 (ref.), namely

$$(m, j, w, n) \in \{(m, j, w, n) \mid 0 \le j \le n - w \text{ and } 0 \le m \le w - 1\}.$$

Admissibility both guarantees that the parameters define a partition of the set of players (ie. $n-m-j \geq 0$), and that there exists a unique and strictly positive best response probability q for mixing agents. Indeed, if $j > n - w$, mixing agents are not pivotal and any $Q \in [0, 1]$ is a best response. Likewise, if $m > w - 1$, the unique best response of potentially mixing agents is $Q = 0$. For the sake of convenience, let us introduce the notation

$$A = n - m - j \tag{A.2}$$

$$B = w - m. \tag{A.3}$$

before seeking for admissible combinations of parameters values (w, m, j, Q) and n, such that no player has an incentive to unilaterally change his strategy in the model. It is the case when the following incentive constraints are simultaneously satisfied

$$\begin{aligned} -\binom{m}{m+j} \sum_{k=B-1}^{A-1} \binom{k}{A-1} Q^k (1-Q)^{A-1-k} c\,(1+m+k) \\ -\binom{m}{m+j} \sum_{k=0}^{B-2} \binom{k}{A-1} Q^k (1-Q)^{A-1-k}\, t = \\ -t \binom{m}{m+j} \sum_{k=0}^{B-1} \binom{k}{A-1} Q^k (1-Q)^{A-1-k} \end{aligned}$$

[14]Dixit and Olson 10 pointed out that the literature on devices to achieve efficiency in the public good provision game implicitly assume that all players agree to participate to such devices.

$$\Leftrightarrow \binom{B-1}{A-1} Q^{B-1}(1-Q)^{A-B}\, t = \sum_{k=B-1}^{A-1} \binom{k}{A-1} Q^k (1-Q)^{A-1-k}\, c\,(1+m+k)\,, \tag{A.4}$$

which is the algebraic form of the condition that to contribute and not to contribute must yield the same expected gains for the mixing agents, and

$$\begin{aligned} -\binom{m-1}{m+j-1} \sum_{k=B}^{A} \binom{k}{A} Q^k (1-Q)^{A-k}\, c\,(m+k) \\ -\binom{m-1}{m+j-1} \sum_{k=0}^{B-1} \binom{k}{A} Q^k (1-Q)^{A-k}\, t \geq \\ \binom{m-1}{m+j-1} \sum_{k=0}^{B} \binom{k}{A} Q^k (1-Q)^{A-k}\, t \end{aligned}$$

$$\Leftrightarrow \binom{B}{A} Q^{B}(1-Q)^{A-B}\, t \geq \sum_{k=B}^{A} \binom{k}{A} Q^k (1-Q)^{A-k}\, c\,(m+k)\,, \tag{A.5}$$

$$\begin{aligned} -\binom{m}{m+j-1} \sum_{k=0}^{B-1} \binom{k}{A} Q^k (1-Q)^{A-k}\, t \geq \\ -\binom{m}{m+j-1} \sum_{k=B-1}^{A} \binom{k}{A} Q^k (1-Q)^{A-k}\, c\,(m+k) \\ -\binom{m}{m+j-1} \sum_{k=0}^{B-2} \binom{k}{A} Q^k (1-Q)^{A-k}\, t \end{aligned}$$

$$\Leftrightarrow \binom{B-1}{A} Q^{B-1}(1-Q)^{A-B+1} t \leq \sum_{k=B-1}^{A} \binom{k}{A} Q^k (1-Q)^{A-k}\, c\,(m+k)\,, \tag{A.6}$$

which are the algebraic forms for the conditions that (i) contributing is at least as good than not contributing for participating agents (A.5) and (ii) not contributing is at least as good than contributing for non participating agents (A.6). These results are an extension of Palfrey and Rosenthal 21 to subscription games with our more general payoffs structure.

Note that if (i) $m = 0$, only conditions (A.4) and (A.6) need to be satisfied (ii) $j = 0$, only conditions (A.5) and (A.6) need to be satisfied (iii) $j = m = 0$, only condition (A.6) applies and (iv) for $m+j = n-1$, the admissibility constraints hold at equality and the conditions rewrite as the pure Nash equilibria conditions with $Q \in \{0, 1\}$.

Appendix B. Proof of Proposition 2

Let us write the standard form constraints, and rearrange them so as to isolate the basis variables x_1, p_w and p_n:

$$\begin{aligned}
x_1 = p_{w-1}&\left[\left(\frac{(n-w)\,nc\,(n)}{w\,(c(w)-t)-nc\,(n)}c\,(w+1)\right)+(n-(w-1))\,(c(w)-t)\right]\\
+\sum_{k=w+1}^{n-1}p_k&\left((n-k)\,c\,(k+1)+\left(\frac{nc\,(n)-kc\,(k)}{w(c(w)-t)-nc\,(n)}(n-w)\,c\,(w+1)\right)\right)\\
&-\frac{(n-w)\,nc\,(n)}{w(c(w)-t)-nc\,(n)}c\,(w+1)-\frac{x_2\,(n-w)}{w(c(w)-t)-nc\,(n)}(c\,(w+1))\\
&+\sum_{k=0}^{w-2}p_k\left(\frac{(n-w)\,nc\,(n)}{w(c(w)-t)-nc\,(n)}c\,(w+1)\right)
\end{aligned}$$

$$\begin{aligned}
p_w = &\sum_{k=0}^{w-1}p_k\left(\frac{nc\,(n)}{w(c(w)-t)-nc\,(n)}\right)-\frac{nc\,(n)}{w(c(w)-t)-nc\,(n)}\\
&+\sum_{k=w+1}^{n-1}p_k\left(\frac{nc\,(n)-kc\,(k)}{w(c(w)-t)-nc\,(n)}\right)-\frac{x_2}{w(c(w)-t)-nc\,(n)}
\end{aligned}\tag{B.1}$$

$$\begin{aligned}
p_n = 1-&\sum_{k=0}^{w-1}p_k\left(\frac{w(c(w)-t)}{w(c(w)-t)-nc\,(n)}\right)+\frac{nc\,(n)}{w(c(w)-t)-nc\,(n)}\\
&-\sum_{k=w+1}^{n-1}p_k\left(\frac{w(c(w)-t)-kc\,(k)}{w(c(w)-t)-nc\,(n)}\right)+\frac{x_2}{w(c(w)-t)-nc\,(n)}
\end{aligned}\tag{B.2}$$

We substitute (B.1) and (B.2) into the objective (12). It follows the objective depends only on the n non-basic variables :

$$\begin{aligned}
&\left[\sum_{k=0}^{w-1}p_k\left(\left(\frac{nc\,(n)\,(w(c(w)-t))}{w(c(w)-t)-nc\,(n)}\right)-\left(\frac{wc\,(w)\,nc\,(n)}{w(c(w)-t)-nc\,(n)}\right)-nt\right)\right]\\
+&\left[\sum_{k=w+1}^{n-1}p_k\left(nc\,(n)\left(\frac{w(c(w)-t)-kc\,(k)}{w(c(w)-t)-nc\,(n)}\right)-wc\,(w)\left(\frac{nc\,(n)-kc\,(k)}{w(c(w)-t)-nc\,(n)}\right)-kc\,(k)\right)\right]\\
&+nc\,(n)\left(\frac{tw}{w(c(w)-t)-nc\,(n)}\right)+\left(\frac{wc\,(w)-nc\,(n)}{w(c(w)-t)-nc\,(n)}\right)x_2
\end{aligned}\tag{B.3}$$

Rearranging (B.3) we obtain :

$$\sum_{k=0}^{w-1}p_k\left(\frac{-nc\,(n)\,tw}{w(c(w)-t)-nc\,(n)}-nt\right)+nc\,(n)\left(\frac{tw}{w(c(w)-t)-nc\,(n)}\right)\tag{B.4}$$

$$+\sum_{k=w+1}^{n-1}p_k\left(\frac{tw\,(kc\,(k)-nc\,(n))}{w(c(w)-t)-nc\,(n)}\right)+\left(\frac{wc\,(w)-nc\,(n)}{w(c(w)-t)-nc\,(n)}\right)x_2\tag{B.5}$$

Then, let us remark that:

$$\underbrace{\left(\frac{nc\,(n)}{nc\,(n)+w(t-c(w))}\right)}_{<1} w < n. \tag{B.6}$$

Coefficients (dual variables) associated to $\{p_k\}_0^{w-1}$ are therefore negative. Assumptions on cost and profitability imply the remaining components of (B.5) are negative as well.

Since non-basic variables are all set to 0 while constrained to non-negativity, it can be concluded the program is solved for:

$$\begin{cases} (n-w)\,p_w c\,(w+1) & = x_1 \\ w\ p_w\,(t-c(w)) - np_n c\,(n) & = 0 \\ p_w + p_n & = 1 \end{cases}$$

$$\Rightarrow \begin{cases} x_1^* & = \frac{(n-w)nc(n)c(w+1)}{w(t-c(w))+nc(n)} \\ p_w^* & = \frac{nc(n)}{w(t-c(w))+nc(n)} \\ p_n^* & = \frac{w(t-c(w))}{w(t-c(w)+nc(n)}. \end{cases} \tag{B.7}$$

Appendix C. The symmetric mixed NE is a CE

Remark that symmetry, added to the independence of individual participation decisions, imply for all i:

$$\frac{p_k p\,(i,k)}{\sum_{k=0}^{n} p(i,k)p_k} = \left(\binom{n}{k} Q^k(1-Q)^{n-k}\right)\frac{k}{nQ}$$

$$\frac{p_k(1-p(i,k))}{1-\sum_{k=0}^{n} p(i,k)p_k} = \left(\binom{n}{k} Q^k(1-Q)^{n-k}\right)\frac{n-k}{n}\frac{1}{1-Q},$$

Now, using the two previous equalities, we can rewrite condition (9) of CE as follows:

$$\binom{n}{w-1}Q^{w-1}(1-Q)^{n-w+1}\frac{n-w+1}{n(1-Q)}\left(t-\frac{c}{w}\right) - \sum_{k=w}^{n-1}\binom{n}{k}Q^k(1-Q)^{n-k}\frac{n-k}{n(1-Q)}c(k+1) \le 0$$

$$\binom{n}{w-1}Q^{w-1}(1-Q)^{n-w+1}\frac{n-w+1}{n(1-Q)}\left(t-\frac{c}{w}\right) - \frac{1}{(1-Q)\,n}\sum_{k=w}^{n-1}\binom{n}{k+1}Q^k(1-Q)^{n-k}(k+1)c(k+1) \le 0$$

$$(n-w+1)\binom{n}{w-1}Q^{w-1}(1-Q)^{n-w+1}\left(t-\frac{c}{w}\right) - \sum_{k=w+1}^{n}\binom{n}{k}Q^{k-1}(1-Q)^{n-k+1}kc(k) \le 0 \tag{C.1}$$

Likewise, condition (10) of CE rewrites:

$$\left(\binom{n}{w} Q^w(1-Q)^{n-w}\right)\frac{w}{nQ}\left(t-\frac{c}{w}\right) + \sum_{k=w+1}^{n}\left(\binom{n}{k} Q^k(1-Q)^{n-k}\right)\frac{k}{nQ}\,(-c(k)) \ge 0$$

$$\left(\binom{n}{w} Q^w(1-Q)^{n-w}\right)\frac{w}{nQ}\left(t-\frac{c}{w}\right) - \frac{1}{(1-q)\,n}\sum_{k=w+1}^{n}\left(\binom{n}{k} Q^{k-1}(1-Q)^{n-k+1}\right)kc(k) \ge 0$$

$$(n-w+1)\binom{n}{w-1}Q^{w-1}(1-Q)^{n-w+1}\left(t-\frac{c}{w}\right) - \sum_{k=w+1}^{n}\binom{n}{k}Q^{k-1}(1-Q)^{n-k+1}kc(k) \ge 0 \tag{C.2}$$

Now, notice that (C.1) and (C.2) imply:

$$(n-w+1)\left(\binom{n}{w-1}Q^{w-1}(1-Q)^{n-w+1}\right)\left(t-\frac{c}{w}\right)=\sum_{k=w+1}^{n}\left(\binom{n}{k}Q^{k-1}(1-Q)^{n-k+1}\right)kc(k).$$

$$\Leftrightarrow (n-w+1)\binom{n}{w-1}Q^{w-1}(1-Q)^{n-w+1}\left(t-\frac{c}{w}\right)=n\sum_{k=w}^{n-1}\binom{n-1}{k}Q^{k}(1-Q)^{n-k}c(k+1)$$

$$\Leftrightarrow \binom{n-1}{w-1}Q^{w-1}(1-Q)^{n-w+1}\left(t-\frac{c}{w}\right)=\sum_{k=w}^{n-1}\binom{n-1}{k}Q^{k}(1-Q)^{n-k}c(k+1),$$

which is the symmetrical mixed NE condition when m=j=0. It follows the unique mixed NE is a symmetric CE. As p^* is the optimum, it generate a higher aggregate expected payoffs than the symmetric mixed NE.

Appendix D. Proof of Corollary 3

First, remark that the probability constraint involves the derivative of $V(\hat{p})$ with respect to t can be rearranged as follows:

$$-\frac{d}{dt}\sum_{w}^{\lceil nQ(t)\rceil-1}p_k\left(\frac{kc(k)}{n}-t\right)-\frac{d}{dt}\sum_{\lceil nQ(t)\rceil}^{n}p_k\left(\frac{kc(k)}{n}+t\right)-\sum_{0}^{w}p_k, \tag{D.1}$$

since $\frac{d}{dt}\sum_0^{w-1}p_k+\frac{d}{dt}\sum_w^{\lceil nQ(t)\rceil-1}p_k=\frac{d}{dt}\sum_{\lceil nQ(t)\rceil}^{n}p_k$. As m obviously follows the binomial $\mathcal{B}(n,Q(t))$, it can be approximated by the normal distribution \(0, 1), with change of variable: $M=\frac{m-nQ(t)}{\sqrt{nQ(t)(1-Q(t))}}$. The derivative (D.1) therefore rewrites :

$$\int_{\frac{w-nQ(t)}{\sqrt{nQ(t)(1-Q(t))}}}^{0}\frac{1}{2\pi}e^{-\frac{x^2}{2}}dx-\underbrace{\int_0^{\frac{n-nQ(t)}{\sqrt{nQ(t)(1-Q(t))}}}\frac{1}{2\pi}e^{-\frac{x^2}{2}}dx-\int_{\frac{-nQ(t)}{\sqrt{nQ(t)(1-Q(t))}}}^{\frac{w-nQ(t)}{\sqrt{nQ(t)(1-Q(t))}}}\frac{1}{2\pi}e^{-\frac{x^2}{2}}dx}_{:=A} \tag{D.2a}$$

$$+\frac{1}{2\pi}e^{-\frac{\left(\frac{w-nQ(t)}{\sqrt{nQ(t)(1-Q(t))}}\right)^2}{2}}\left(t-\frac{wc(w)}{n}\right)\left(\frac{n\left(w+Q(t)\left(n-2w\right)\right)}{2\left(nQ(t)\left(1-Q(t)\right)^{3/2}\right)}\right) \tag{D.2b}$$

$$+\frac{1}{2\pi}e^{-\frac{\left(\frac{n-nQ(t)}{\sqrt{nQ(t)(1-Q(t))}}\right)^2}{2}}\left(\frac{nc(n)}{n}+t\right)\left(\frac{n}{2Q(t)\sqrt{nQ(t)\left(1-Q(t)\right)}}\right) \tag{D.2c}$$

Assuming $t>c(w)$ and $(mc(m))'<0$, we know the terms (D.2b) and (D.2c) are positive. Finally, as $Q(t)$ increases in t, and $w-nQ(t)<0<n-nQ(t)$ for all $Q(t)>0$, we know the first term (respectively, the second term) of (D.2a) increases (respectively, decreases) in t. It follows there exists t_0 such that $\int_{\frac{w-nQ(t_0)}{\sqrt{nQ(t_0)(1-Q(t_0))}}}^{0}\frac{1}{2\pi}e^{-\frac{x^2}{2}}dx=-A$, as well as $t_1<t_0$ from which the whole expression becomes positive.

Appendix E. The two-player example

In order to illustrate the general study, let us consider the game when $n=2$ and $w=1$. The strategic form is the following :

	$s_2 = 1$	$s_2 = 0$
$s_1 = 1$	$-c(2), -c(2)$	$-c, 0$
$s_1 = 0$	$0, -c$	$-t, -t$

where $wc(w) = c(1) \equiv c$. From the study above, we know this game has two pure Nash equilibria corresponding to the strategy profiles such that $m = w$, namely $(1,0)$ and $(0,1)$, and that the set of admissible vectors is

$$(m, j, w, n) = \{(0,0,1,2), (0,1,1,2)\}.$$

Substituting the parameters values in the relevant general conditions (A.6) and (A.4), we get

$$\begin{cases} (1-Q)\,t = (1-Q)\,c + Qc\,(2) & \text{if } j = 0 \\ t = c \text{ and } 0 \leq Qc & \text{if } j = 1 \end{cases}, \tag{E.1}$$

conjointly characterizing the symmetrical mixed Nash equilibria of the game. Note then that the inequalities induced by $j = 1$ imply that mixed Nash equilibria with asymmetrical supports exist if and only if $t = c$, in which case one player i does not contribute while any distribution on S_{-i} is also a best response for its opponent. Finally, we can derive a unique mixed strategy equilibrium from the first case equality, with the symmetrical distribution

$$\hat{Q} = \frac{2\,(c-t)}{2\,(c-t) - 2c\,(2)} \quad \text{and} \quad 1 - \hat{Q} = \frac{2c\,(2)}{2c\,(2) + 2\,(t-c)}. \tag{E.2}$$

More general conditions to characterize Nash equilibria such that both agents do mix strategies are easy to derive in this simple game, and show that the distribution exhibited in the symmetrical case actually exhausts the mixed strategy equilibria. The corresponding aggregate payoff is

$$\begin{aligned} \sum_{s \in S} \left(\prod_i p_i(s_i)\,(u_1(s) + u_2(s)) \right) &= (2\,(c-t) - 2c(2))\,Q^2 + (2Q-1)\,2t - 2Qc \\ &= -\frac{2t}{2t + 2c\,(2) - 2c} 2c\,(2)\,. \end{aligned} \tag{E.3}$$

From our assumption on costs, we know that $2c(2) < 2c$, which implies that $\left(\hat{Q}, 1-\hat{Q}\right)$ yields a smaller payoff than the socially optimal pure allocation. For a minimum tax $t = c$, the payoff under $\left(\hat{Q}, 1-\hat{Q}\right)$ is $-2c < -c$, and then strictly increases in the threat stringency with

$$\lim_{t \to +\infty} -\frac{2t}{t + c\,(2) - c} c\,(2) = -2c(2) \tag{E.4}$$

Specifically, mixed strategies are Pareto improving compared to the pure Nash equilibria allocations for a tax level

$$t > \frac{c(2c(2) - 2c}{2(2c(2) - c)} \tag{E.5}$$

Thus, extending the set of pure strategies to mixed strategies allows to reach higher expected aggregate payoffs for a high enough tax threat, provided agents find a way to coordinate on equilibria multiplicity. Let us see what would be the set of reachable payoffs in the voluntary agreement with mediated communication such as described in our coordination device. We already know that any mixed strategies Nash equilibria of a game is also a correlated equilibria of this game (Myerson 17), meaning that the VA with the coordination device certainly allows to implement $(\hat{Q}, 1-\hat{Q})$. But we want to check if even higher payoffs could be implemented as correlated equilibria for a *given* threat level, since we know that a credibility requirement would actually limit the regulator in his choice of t. Accordingly to the general case, we denote p_{kl} the probability assigned by the regulator to the pure strategy profile $(s_1 = k, s_2 = l) \in S$, with $\sum_{s\in S} p_s = 1$. Then, following Myerson 17's interim definition of correlated equilibrium, let us write the strategic incentive constraints in our two-player game

$$\frac{p_{11}}{p_{11}+p_{10}}c(2)+\frac{p_{10}}{p_{11}+p_{10}}c \leq \frac{p_{10}}{p_{11}+p_{10}}t$$
$$\frac{p_{00}}{p_{01}+p_{00}}t \leq \frac{p_{01}}{p_{01}+p_{00}}c(2)+\frac{p_{00}}{p_{01}+p_{00}}c$$
$$\frac{p_{11}}{p_{11}+p_{01}}c(2)+\frac{p_{01}}{p_{11}+p_{01}}c \leq \frac{p_{01}}{p_{11}+p_{01}}t$$
$$\frac{p_{00}}{p_{10}+p_{00}}t \leq \frac{p_{10}}{p_{10}+p_{00}}c(2)+\frac{p_{00}}{p_{10}+p_{00}}c$$

with the probability constraint

$$\begin{cases} p_{11}+p_{10}+p_{01}+p_{00}=1 \\ p_{11}\geq 0,\ p_{10}\geq 0,\ p_{01}\geq 0 \text{ and } p_{00}\geq 0 \end{cases}, \tag{E.7}$$

which is the algebraic form for the condition that for any individual suggestion from the regulator to an agent, and provided a given probability distribution on S that was prealably announced, the agent has no incentive not to follow the suggestion. The incentive constraints rewrite

$$p_{10}\left(\frac{t}{c(2)}-\frac{c}{c(2)}\right)-p_{11}\geq 0 \tag{E.8a}$$

$$p_{00}\left(\frac{c}{c(2)}-\frac{t}{c(2)}\right)+p_{01}\geq 0 \tag{E.8b}$$

$$p_{01}\left(\frac{t}{c(2)}-\frac{c}{c(2)}\right)-p_{11}\geq 0 \tag{E.8c}$$

$$p_{00}\left(\frac{c}{c(2)}-\frac{t}{c(2)}\right)+p_{10}\geq 0 \tag{E.8d}$$

and a maximization program for the regulator can be formulated as follows

$$\underset{p_{11},p_{10},p_{01},p_{00}}{max} \quad -c(p_{10}+p_{01})-2c(2)p_{11}-2p_{00}t \tag{E.9}$$
$$s.t. \qquad \text{(E.8a), (E.8b), (E.8c), (E.8d) and (E.7).}$$

Using the assumption $t>c$, and denoting correlated strategies as vectors

$$(p_{11}\ p_{10}\ p_{01}\ p_{00})^T,$$

we first notice that the set of vectors solving program E.7 must be a subset of $(p_{11}\ p_{10}\ p_{01}\ 0)^T \in \mathbb{R}^4$. Specificaly, by substituting $p_{00}=0$ into (E.8b), (E.8d) and (E.7), we get the set of candidates $(p_{11}\ p_{10}\ p_{01}\ 0)^T \in \mathbb{R}^4$ such that

$$\begin{cases} p_{10}\geq p_{11}\frac{C(2)}{2(t-c)} \\ p_{01}\geq p_{11}\frac{C(2)}{2(t-c)} \\ p_{11}\geq 0 \\ p_{11}+p_{10}+p_{01}=1 \end{cases} \tag{E.10}$$

Geometricaly, it is the area bounded by the inequalities

$$p_{10}\geq \frac{2c(2)}{2(t-c)+2c(2)}-\left(\frac{2c(2)}{2(t-c)+2c(2)}\right)p_{01} \tag{E.11}$$

$$p_{10}\geq 1-\left(1+\frac{2(t-c)}{2c(2)}\right)p_{01} \tag{E.12}$$

$$p_{10}\leq 1-p_{01} \tag{E.13}$$

on the affine hyperplane defined by $p_{11} = (1 - p_{10} - p_{01})$, with the two first inequalities intersecting in

$$p_{01} = p_{10} = \frac{c(2)}{2c(2) + t - c}\ , \tag{E.14}$$

as illustrated in figure E.4. Provided our assumption on costs, it is now obvious the regulator maximizes

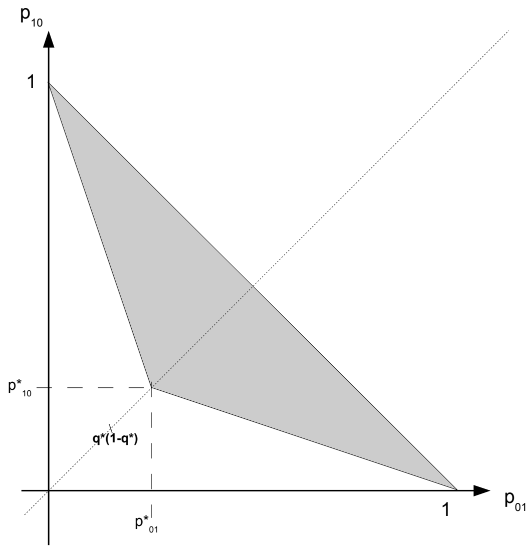


Figure E.4: The set of correlated equilibria in the two-player example.

the objective by assigning a maximal probability to the full-participation profile. Consequently, the solution of the program is

$$\begin{pmatrix} p^*_{11} \\ p^*_{10} \\ p^*_{01} \\ p^*_{00} \end{pmatrix} = \begin{pmatrix} \frac{t-c}{2c(2)+t-c} \\ \frac{1}{2}\frac{2c(2)}{2c(2)+t-c} \\ \frac{1}{2}\frac{2c(2)}{2c(2)+t-c} \\ 0 \end{pmatrix}, \tag{E.15}$$

and the value of the objective (or expected aggregate gain) is

$$\begin{aligned} V(p^*, t) &= \sum_{s \in S} p^*_s \left(u_1(s) + u_2(s)\right) \\ &= -c\left(p_{10} + p_{01}\right) - 2c(2)\, p_{11} \\ &= -\frac{2c(2)\, t}{2c(2) + t - c}. \end{aligned} \tag{E.16}$$

Note that both $p^*_{01} = p^*_{10}$ are decreasing in t, while $p^*_{11} = 1 - (p^*_{01} + p^*_{10})$ is increasing in t. In other words, a higher threat rises the probability on the socially optimal allocation

$$\frac{\partial}{\partial t}\left(\frac{t - c}{2c(2) + t - c}\right) = \frac{2c(2)}{\left(2c(2) + t - c\right)^2} > 0. \tag{E.17}$$

Or, geometrically, the grey area extends along the bisectrice on figure E.4. The aggregate payoff tends then to the payoff corresponding to the socially optimal allocation

$$\lim_{t \to +\infty} V(p^*, t) = -2c(2). \tag{E.18}$$

Finally, let us remark as previously mentioned, that the pure and mixed Nash equilibria of the game do satisfy the correlated equilibria conditions (easily verified by substituting $p_{11} = (q^*)^2$, $p_{10} = p_{01} = Q^*(1 - Q^*)$ and $p_{00} = (1 - q^*)^2$ into (E.8a)-(E.8d)). But we know now that they yield a smaller aggregate payoff than the optimal correlated equilibrium for all t. Specifically, we have the following ranking

$$V(p^*, t) > -2\frac{c(2)\,t}{c(2) + (t - c)} > -c, \tag{E.19}$$

implying that the coordination device not only solve the problem raised by multiplicity, but also ensures that a higher expected aggregate payoff is reached for a given credible level of threat.

References

[1] Arce, D. G. and Sandler, T. (2001). Transnational public goods: strategies and institutions. *European Journal of Political Economy*, 17:493–516.

[2] Aumann, R. J. (1974). Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1(1):67–96.

[3] Aumann, R. J. (1987). Correlated equilibrium as an expression of bayesian rationality. *Econometrica*, 55(1):1–18.

[4] Borkey, P., Glachant, M., and Leveque, F. (1998). Voluntary approaches for environmental policy in oedc countries: An assessment. Technical report, Paris: CERNA, Centre d'économie Industrielle.

[5] Brau, R. and Carraro, C. (2011). The design of voluntary agreements in oligopolistic markets. *Journal of Regulatory Economics*, 39(2):111 – 142.

[6] Cavaliere, A. (2001). Coordination and the provision of discrete public goods by correlated equilibria. *Journal of Public Economic Theory*.

[7] Clarke, E. (1971). Multipart pricing of public goods. *Public Choice*, 11(1):19–33.

[8] d'Aspremont, C., Jacquemin, A., Gabszewicz, J. J., and Weymark, J. A. (1983). On the stability of collusive price leadership. *Canadian Journal of Economics*, 16(1):17–25.

[9] Dawson, N. L. and Segerson, K. (2008). Voluntary agreements with industries: Participation incentives with industry-wide targets. *Land Economics*, 84(1):97–114.

[10] Dixit, A. and Olson, M. (2000). Does voluntary participation undermine the coase theorem. *Journal of Public Econommics*, 76:309–335.

[11] Foley, D. K. (1970). Lindahl's solution and the core of an economy with public goods. *Econometrica*, 38(1):66–72.

[12] Glachant, M. (1994). The setting of voluntary agreements between industry and government: Bargaining and efficiency. *Business Strategy and the Environment*, 3(2):43–49.

[13] Glachant, M. (2007). Non-binding voluntary agreements. *Journal of Environmental Economics and Management*, 54(1):32–48.

[14] Groves, T. and Ledyard, J. (1977). Optimal allocation of public goods: A solution to the free rider problem. *Econometrica*, 45(3):783–809.

[15] Hong, F. and Lim, W. (2016). Voluntary participation in public goods provision with cosean bargaining. *Journal of Economic Behavior and Organization*, (126):102–119.

[16] Moulin, H. (1986). Characterizations of the pivotal mechanism. *Journal of Public Economics*, 31:53–78.

[17] Myerson, R. B. (1997). *Game Theory Analysis of Conflict*, chapter 6. Harvard University Press.

[18] Nisan, N., Roughgarden, T., Tardos, E., and Vijay, V. (2007). *Algorithmic Game Theory*. Cambridge University Press.

[19] OECD (1999). Voluntary approaches for environmental policy : An assessment. Technical report, OECD, Paris.

[20] Padberg, M. (1999). *LINEAR OPTIMIZATION AND EXTENSIONS*, volume 12 of *0937-5511*. Springer-Verlag.

[21] Palfrey, T. R. and Rosenthal, H. (1984). Participation and the provision of discrete public goods: a strategic analysis. *Journal of Public Economics*, 24(2):171–193.

[22] Rosenthal, R. (1976). Lindahl's solution and values for a public-goods example. *Journal of Mathematical Economics*, 3:37–41.

[23] Segerson, K. and Wu, J. (2006). Nonpoint pollution control: Inducing first-best outcomes through the use of threats. *Journal of Environmental Economics and Management*, 51(2):165–184.