J Regul Econ (2011) 39:111–142
DOI 10.1007/s11149-010-9134-z

ORIGINAL ARTICLE

The design of voluntary agreements in oligopolistic markets

Rinaldo Brau · Carlo Carraro

Published online: 11 November 2010
© Springer Science+Business Media, LLC 2010

Abstract This paper analyses the conditions under which a group of firms have the incentive to sign a voluntary agreement (VA) to control polluting emissions even in the presence of free-riding by other firms in the industry. We consider a policy framework in which firms in a given industry decide whether or not to sign a VA proposed by an environmental regulator. We identify the features that a VA should possess in order to provide firms with an incentive to participate in the VA and to enhance its economic and environmental effectiveness. Under very general conditions on the shape of the demand schedule, we obtain the following results. First, a VA does not belong to the equilibrium of the coalition game when benefits from voluntary emission abatement are a pure public good, unless an industry emission target is set by the regulator. Second, in the presence of partial spillovers—i.e. when signatories obtain more benefits from the VA than non-signatories—a VA can belong to the equilibrium only if a minimum participation rule is guaranteed. Third, a VA with a minimum participation rule and a minimum mandatory emission abatement may improve welfare (and even industry profits) compared to a VA in which firms are free to set their own profit maximizing abatement level.

R. Brau (✉)
University of Cagliari and CRENoS,
Viale S. Ignazio 17, 09123 Cagliari, Italy
e-mail: brau@unica.it

C. Carraro
University of Venice, Venice, Italy
e-mail: ccarraro@unive.it

C. Carraro
Fondazione Eni Enrico Mattei, Milan, Italy

C. Carraro
CMCC, Venice, Italy

Keywords Voluntary agreement · Voluntary approaches · New policy instruments · Environmental regulation · Coalition structures · Emission standards

JEL Classification K32 · D21 · Q58

1 Introduction

Since the early 1990s, traditional environmental policy based on command-and-control regulation or on price incentives (i.e. environmental taxes and subsidies) has been challenged by the increasing popularity of various forms of voluntary commitments by firms to improve their environmental performance beyond what is required by the law. The terms "voluntary approaches" and "voluntary agreements" (henceforth VAs) are the most common labels attached to this kind of environmental policy instruments.[1] Given the documented increased use of VAs in addressing domestic and international environmental problems (e.g. OECD 2003), several authors have analyzed their main pros and cons.[2] On the one hand, VAs' flexibility and potential to stimulate innovation and the diffusion of cleaner technologies has been emphasized. On the other, concerns have been expressed about VAs' real effectiveness and possible misuse as a tool for regulatory capture.

When studying the economic functioning of VAs, the first question that arises is "why do profit maximizing firms decide to adhere to a VA?" By and large, the answer falls into two main categories (Brau and Carraro 2006): (i) VAs can *increase market demand*, and therefore profits, mainly by enhancing a firm's green reputation. (ii) VAs can be used to *achieve cost advantages*, which may take the form of avoiding the costs of public regulation designed to address the environmental problem, or technical and financial incentives granted by regulators, or advantages from information and innovation sharing among signatory firms. Specific explanations and institutional arrangements for all forms of VAs have been widely studied in the recent environmental economics literature (e.g. for public voluntary programs, see Wu and Babcock 1999; Lyon and Maxwell 2003; Glachant 2007; for negotiated agreements, Segerson and Miceli 1998; Manzini and Mariotti 2003; for unilateral commitments, Bagnoli and Watts 2003; Lutz et al. 2000).

Most of these articles restrict their analysis to the bargaining process between an industry representative (or monopolist) and a public regulatory body, or to the strategic interactions among firms which offer environmentally differentiated products. A less researched area has focused on the interactions between firms that negotiate a VA

[1] According to a well established classification (e.g. OECD 1999; Carraro and Lévêque 1999), the term *voluntary approaches* refers to *unilateral commitments*, which consist of environmental improvement programs established by polluters themselves and communicated to their stakeholders; *public voluntary programs*, in which participating firms agree to frameworks and standards developed by public authorities; *negotiated agreements*, which are contracts between public authorities and industry. The term *voluntary agreements* should mainly be used for the latter two categories, but is often used to encompass all forms of voluntary commitments by firms (e.g. Glachant 2007).

[2] See the surveys by Alberini and Segerson (2002) and Lyon and Maxwell (2002), and the books edited by Carraro and Lévêque (1999), Morgenstern and Pizer (2007) and Baranzini and Thalmann (2004).

with the environmental regulator. In this setting, it is important to take into account the fact that the benefits firms gain by adhering to a VA often exhibit some degree of non-rivalry and non-excludability, i.e. once signed and implemented, the VA yields positive spillovers to non-signatories (e.g. see Xepapadeas and Passa 2004; Dawson and Segerson 2008, for theoretical models dealing with this issue). The empirical relevance of this situation can be related not only to those cases where a VA by a group of firms pre-empts mandatory regulation for the whole industry, but also to those where VAs can essentially be seen as devices suited for disseminating cost-effective emission abatement practices (e.g. see Lyon and Maxwell 2007 who point out that this is the very nature of a large share of the US Environmental Protection Agency (EPA) public voluntary programs; see also Delmas and Montes-Sancho 2010 for an analysis of participation in the 'Climate Challenge Program' between the Department of Energy (DOE) and representatives of the national electric utilities). In the latter case, a dichotomy has arisen between firms which obtained substantial results and firms which undertook only 'symbolic cooperation', a situation which could arise given that "*most VAs lacked explicit measures to sanction firms that pursue only symbolic cooperative strategies or sometimes even to identify such firms*" (Delmas and Montes-Sancho 2010, p. 2).

These examples suggest that there are two particularly important issues for industry-wide VAs: (a) strategic interaction among firms, given that a problem of free-riding is likely to arise; (b) appropriate design of VA incentives in order to maximize environmental effectiveness and overall social benefits. This paper focuses on VAs between several firms and an environmental regulator and analyses whether firms in a given industry have an incentive to sign a VA when the benefits derived from the agreement may be enjoyed not only by the participants, but also spill over to others. This situation is widely discussed in the public finance and industrial organization literature, where it is often argued that the incentives to free-ride are so strong that no public good is actually provided (or no R&D is carried out, or no information is collected). In our setting, the incentive to free ride may result in no VA being implemented. However, the literature on international environmental agreements (IEAs) on global commons has shown that even in those cases where a good is purely public, a number of agents—in this literature sovereign countries—may decide to pay the cost of providing the public good (technically, they form a coalition within which the burden of providing the public good is shared).[3] If the public good is imperfect, then the coalition size usually increases (Yi 1997, 2003). Relying explicitly on this literature, we assess whether these results also apply to VAs, and above all under what conditions a 'coalitional VA' emerges, i.e. a group of firms decides to sign the VA even if other firms choose to free-ride.

This issue has already been studied by Dawson and Segerson (2008, henceforth D-S). They assume that the benefit provided by the VA can only be obtained if an industry-wide environmental target—set by the regulator—is met. Each firm knows that the VA will be beneficial only if a group of firms is committed to achieving the industry's abatement target defined by the regulator. If this is not the case, all firms

[3] See the seminal papers by Carraro and Siniscalco (1993), Barrett (1994); and the surveys by Finus (2003), Barrett (2002) and Carraro (2003).

will be penalized. The structure of the VA is such that the regulatory body cannot make any differentiation contingent on firms' participation in the VA, and its signing can therefore be seen as a pure public good. At equilibrium, some firms sign the agreement in order to meet the target, even if the remaining firms free-ride, and are indifferent between being in the coalition or behaving as singletons.

The approach adopted in this paper is somewhat related to the one described above, but enables us to analyze a more general situation where a broader set of participation incentives can be designed so as to maximize VAs environmental effectiveness throughout the industry. We consider a policy scenario where a binding government policy (e.g. a carbon tax) is *likely* to be enforced *unless* a concerted effort by a polluting industry, in terms of costly emission reductions, is made within a VA scheme. The latter *may* foresee that a share of the advantages (e.g. in terms of dissemination of information and technical assistance, or simply of tax reduction) for individual and joint over-compliance are *signatory-firms specific*, but also imply the industry-wide advantage of abolishing (or reducing the likelihood of) the binding government policy. In this setting, the regulator can try to stimulate firms' cooperation by means of two different types of VA. One in which firms are free to set their own optimal additional abatement level by maximizing profits and VAs' related cost advantages, and one where the regulator requires a minimum threshold on the number of firms entering into the agreement and a minimum 'voluntary' abatement level by each firm. The features of this kind of VAs are investigated in a model without specific functional form assumption on market demand.

We focus in particular on the conditions under which at least some firms adopt the VA and we devote special attention to the policy design implications of the analysis. VAs are often criticized—at least in those cases in which they do not foresee an explicit environmental target—for being cost-effective but environmentally ineffective, because firms within a VA tend to pre-empt regulations or access subsidies but achieve poor emission reductions. This paper examines this criticism of VAs and discusses their proper design in terms of minimum mandatory requirement (a minimum abatement target for each firm or other comparable policy measures at the industry level), and minimum number of participants.

From a positive viewpoint, we find that VAs do not belong to the equilibrium of the coalition game in the case of a pure public good, i.e. when spillovers are such that all firms benefit equally from the abatement achieved by the signatories, unless a situation of 'coalition unanimity' occurs when the profitability of the VA ceases (e.g. due to the impossibility of reaching the collective target set by the regulator), in this case allowing for the formation of a VA with partial participation. Second, in the case of an 'impure public good', namely partial spillovers, the VA can be an equilibrium of the coalition game provided that a minimum participation rule in terms of number of participants is satisfied. In this case, if the minimum participation constraint is met, then strategic interactions induce all firms to sign the VA. We also find that VAs with partial participation are still possible only in the presence of relatively low industry-wide targets. Therefore, with partial spillovers, under fairly general conditions either no firms sign the VA or all firms sign it, which is an important difference compared to the literature on IEAs—where usually a subset of agents (countries) decides to sign

the agreement (a partial coalition is formed)—essentially driven by the presence of a superior public body which can ensure the partial appropriability of VA advantages.

As for policy design issues, we examine how a regulator should design the VA in order to increase both the number of participants and the VA's economic and environmental effectiveness. Besides the opportunity to discriminate, with specific devices, between signatory and non-signatory firms, we show that setting a minimum threshold on the number of participants for the VA's entry into force is an effective instrument for neutralizing firms' incentives to free ride. This is all the more important, the lower the ability to distinguish between signatories and non-signatories. Moreover, we find that a VA with a minimum firm-specific mandatory regulation (or alternatively with a more stringent industry target) may represent a welfare improving policy *vis à vis* one where firms are free to set their own optimal abatement level. More surprisingly, firms' profits may also increase (within a specific parameter interval), since a regulation in an oligopolistic setting may induce firms to adopt cartel-like production levels (this result is similar to those contained in Dung 1993; Katsoulacos and Xepapadeas 1996 for taxation; and Farzin 2003; Farzin and Akao 2006 for quality standards; but has not yet been shown for VAs). Finally, "closed membership rules" should be avoided in the VA given the presence of incentives for signatory firms to limit the maximum number of participants in the agreement.

The remainder of the paper is structured as follows. Section 2 describes the main features of the VA model. Section 3 studies the equilibrium of the game and identifies the main properties of the equilibrium 'coalitional' VA. Section 4 discusses the policy implications of our analysis and further develops the model by analyzing the effects of a minimum mandatory abatement target. Finally, the main results of the paper are summarized in Sect. 5, where some directions for further research are also outlined.

2 Regulatory-cost minimizing VAs with intra-industry spillovers

As previously stated, there are several benefits that firms could reap by signing an environmental VA. In this paper, we focus on benefits arising from the expected reduction of firms' production costs. This includes absolute cost reductions (due for example to co-operation on innovation),[4] as well as relative cost reductions (e.g. due to preemption of a regulatory threat. Cf. Segerson and Miceli 1998).

These benefits may also spill over to non-signatory firms. For example, Lyon and Maxwell (2007) demonstrate that one of the most common elements of the US-EPA's voluntary programs is the dissemination of shared information about abatement techniques throughout firms, but this information may be received by participants, as well as by nonparticipants. More generally, in a political-economic framework, one can consider a situation where a group of firms can prevent implementation of a cost-ineffective regulation by undertaking a VA: all firms in the industry will benefit therefrom, not only those that sign the VA. A share of these benefits *can* be exclusive (depending

[4] This group also comprises VAs with 'information sharing activities' (Glachant 1999; Cavaliere and Frontoso Silvestri 2002; Lyon and Maxwell 2007); innovation oriented' VAs with R&D activities by firms (Aggeri and Hatchuel 1999); VAs with 'effects on input prices' (e.g. Konar and Cohen 2001); and 'tailored regulation schemes' (Blackman and Boyd 2002).

on the type of marketing strategy or regulation); but at least some portion of these benefits will be enjoyed by firms that do not adhere to the VA. Hence, a firm's decision on whether or not to adopt a VA (better, to really engage in emission reductions)[5] is strongly affected by incentives to free-ride.

Allowing for asymmetric benefits (i.e. partial spillovers) is important since regulatory authorities do often discriminate between firms adhering and those not adhering to voluntary initiatives. This occurs, first, when the main characteristics of VAs is a regulator's subsidy to participants, which may take the form of technical assistance or financial incentives (e.g. Wu and Babcock 1999), and secondly, when VAs are mainly aimed at pre-empting a mandatory threat. In Europe this is particularly true of some policy mixes, such as the Danish CO_2 agreement scheme (see Krarup and Millock 2007) and the British Climate Change Agreements (Cf. ten Brink et al. 2003; De Muizon and Glachant 2004). In both cases, public authorities clearly discriminate between participants and non-participants, by granting the former high reductions on CO_2 taxations.[6]

In order to consider the possibility that only a subgroup of firms belonging to the industry decides to sign the VA, while all or part of the benefits gained by signatories may spill over to the remaining firms in the industry, in this paper we set up a model having the following main characteristics. First, free-riding incentives are modeled as a continuous function of the voluntary abatement effort undertaken by firms, thus allowing for more general cases of self-regulation. Let the share of total benefits that spill over to non-signatory firms be denoted by β, with $\beta \in [0,1]$. Full spillovers (i.e. the pure public good case) occur when $\beta = 1$. Otherwise, partial spillovers occur.

Secondly, the benefits accruing to firms from adopting a VA take the form of expected cost reductions *vis à vis* those they would incur were mandatory policy measures applied: i. benefits grow with abatement levels, i.e. the capacity of a VA to offset the implementation of mandatory regulation (or to gain direct incentives) increases with the level of emission abatement effected by each signatory firm; ii. benefits are increased by the cooperative behavior of firms that sign the VA and the number of participants. An example of a framework compatible with these hypotheses is the case studied by Aggeri and Hatchuel (1999), where firms adopting a VA agree to common technological innovation efforts in order to attain a cleaner production goal. A second example is represented by the 'information-dissemination nature' of the above mentioned US-EPA VAs.

2.1 The model

Consider an oligopolistic industry with n identical firms that share the same technology and compete *à la Cournot*. In this framework, three externalities can be identified:

[5] If a firm formally joins a VA but does not improve its environmental performance, it is actually behaving as a free-rider. Henceforth we will call "signatory firms" only those really engaging in pollution abatement.

[6] As is well known, energy-intensive firms in Denmark could considerably reduce their carbon tax payments by signing a VA with the Danish Energy Agency in which they commit to undertake investments for energy efficiency (for an evaluation of these policy interventions see Bjørner and Jensen 2002). In the UK, signatories of the Climate Change agreements have obtained an 80% reduction of the climate change levy.

(i) a *Cournot* (or 'combined-profits') *externality* that leads firms to excess production at the Nash equilibrium compared to the case where they behave cooperatively and choose abatement levels without maximizing overall industry profits[7]; (ii) a *VA participation externality,* through which the benefit achieved by a firm increases with the abatement effected by the other signatories[8]; (iii) a *public good externality,* which enables a non-signatory firm to receive a share β of the benefits produced by implementation of the VAs and the related behavior adopted by the firms which sign the VA.[9]

Firms supply a homogeneous good and pollute the environment while a regulatory agency proposes a VA to reduce polluting emissions. A firm's decision variables are the production level, the amount of emission abatement, and whether or not to sign the VA proposed by the regulator. Each firm sets its decision variables by maximizing profits.

The regulator decides the type of VA to propose by taking into account the consequent participation incentives and emission levels. His objective is the maximization of social welfare. The regulator we are considering is fairly similar to that of Segerson and Miceli (1998), Glachant (2007) or Lyon and Maxwell (2003), i.e. a public authority which has incomplete control of a mandatory threat of taxation or command and control intervention, but which, by promoting the use of a VA, is able to eliminate or reduce the severity of such mandatory threat and provide firms with technical or financial incentives. A structured description in line with our framework is offered by Lyon and Maxwell (2003), who model the behavior of a regulator that prefers a publicly subsidized and environmentally weaker VA to the promotion of a (uncertain) legislative intervention in cases of strong political opposition.[10] Note that the preference for the VA may be even more likely where real efficiency gains such as those recalled above are assumed to take place.[11]

[7] From a firm's perspective, this kind of externality exists in all market structures different from cartels, i.e. perfect competition, monopolistic competition and Cournot competition. This effect is referred to as 'combined-profits' externality in the R&D coalition literature (e.g. Kamien et al. 1992).

[8] In the case of the policy mixes cited above, as well as in other examples in climate change policies (Cf. ten Brink et al. 2003), the very structured nature of the schemes apparently seems to eliminate the collective nature of these VAs, i.e. each firm is faced with a take-or-leave-it offer—irrespective of whether or not it decides to accept the scheme—and the nature of the equilibrium seems not to be affected by the decisions of other firms. We consider this view as too simplistic. Indeed, apart from some public voluntary programs unilaterally promoted by regulators, the final written provisions of many VAs are influenced by the number of polluters involved in the negotiation and the extent of the promised environmental commitments.

[9] The latter two externalities actually originate from the same economic effect, but they differ in size (save for the case of perfect non-excludability). We keep them separate in order to better emphasize the role of their different intensity and the form of free-riding incentive.

[10] This does not necessarily conflict with the pursuit of social welfare. Be it a lesser environmental objective than the mandatory policy, or the granting of subsidies, the regulator could simply be evaluating the certainty of the agreed voluntary environmental objective at least as the lesser environmental objectives or the cost of public funds.

[11] We agree with Lyon and Maxwell (2003) that also allowing for the genuine cost advantages related to VAs can make it too easy to reach a conclusion about the superiority of VAs. However, the focus is different here. We are not saying that VAs are always better. Nevertheless, once the VA option is on the agenda, we believe that there is still room for policy design improvement to make even a suboptimal (but feasible) policy better.

The above choices are cast in a three-stage game. In the first stage, we consider a policy scenario where the regulator chooses between two main types of VAs: (i) a VA in which signatory firms are free to choose their profit maximizing additional abatement levels; and (ii) a VA in which each firm is required to attain at least a minimum abatement level in order to avoid direct regulation or an emission tax. Many unilateral commitments undertaken by industries can be included in the first type of VA.[12] The second type of VA, which has more intensive direct regulation and is actually a policy mix between voluntary and binding regulation, can easily be reconciled with the Danish and British VAs schemes on CO_2 emission reductions mentioned above, where obligations for every firm are clearly specified.

In the second stage, the *coalition game* takes place. Given the type of VA proposed by the regulator, firms decide non-cooperatively whether or not to sign it. Each firm chooses its profit maximizing strategy by taking into account its interaction with the other firms, and the implications of the second stage choice for the production and abatement decisions that will be taken in the third stage of the game. The equilibrium in the second stage is a non-cooperative Nash equilibrium and each firm's strategy space is [*to sign the VA, not sign the VA*]. Decisions are simultaneous and the membership rule is open (see Carraro and Marchiori 2003; Yi 2003 for a detailed description of these assumptions and their implications). In the third stage—the *Cournot game* – the n firms choose their output and optimal abatement levels simultaneously and non-cooperatively (with or without the constraint on minimum abatement levels imposed by the regulator).

The formal model. Let y_i be firm i's production level. Total output is therefore $Y = \sum_{i=1}^{n} y_i$. The following conditions are needed on the demand side in order to ensure stability (Seade 1980):

$$P(y_i)' - C(y_i)'' < 0 \tag{1a}$$

$$P(y_i)' + y_i P(y_i)'' < 0, \tag{1b}$$

where $C(y_i)$ are the overall production costs and $P(y_i)$ is the inverse demand function, which is only assumed to be weakly convex.

The production activity is polluting and firms may plan to carry out some abatement effort e_i. Emission abatement is costly. Let us assume that firms deciding to sign a VA have an expected convex abatement cost equal to $\phi_i(e_i)$, where $\phi_i'(0) = 0$, $\phi_i'(e_i) \geq 0$ for $e_i > 0$, and $\phi_i''(e_i) > 0$. Also, assume that abatement costs are independent of the emission reduction being voluntarily undertaken or in observance of a mandatory

[12] One of the most celebrated examples is the 'Declaration' by 5 federations representing the entire German industry in 1996 announcing a 20% reduction in energy consumption by 2005. Also for some negotiated agreements it may appear that the public counterpart largely accepts self-designed commitments. This seems to be the case for agreements signed in 1996 and 1997 in France between the aluminum and the glass packaging industries and the Ministry of Environment, OECD 2003), and for the Climate Challenge program established by the U.S. Department of Energy and electric utility representatives, where no greenhouse gas emissions limits were established for participant firms, nor penalties for firms that did not meet their commitments (Delmas and Montes-Sancho 2010). Also in some EPA public voluntary programs, there are no penalties for failure to participate or failure to achieve the stated goals (the "33/50" program), or for not observing reporting commitments (e.g. the "Waste-Wise" program: see Delmas and Keller 2005).

requirement. The traditional interpretation is that of an end-of pipe-technology which absorbs pollutants.[13] A complementary interpretation, particularly suited to climate change policy agreements, is to consider $\phi_i(e_i)$ as an emission reduction cost function representing the cost from accessing a backstop emission trading market or other 'Kyoto mechanisms'. Finally, let 'A' indicate an 'Agreeing', or VA adopting firm, whilst 'NA' denotes a 'Non-Agreeing' firm.

The crucial element of the model is the total cost function TC (including both production costs $C(y_i)$ and abatement costs $\phi_i(e_i)$), which must be a reduced form of at least three variables: (a) the levels of emission abatement by participants, (b) the number of participants in the VA, and (c) the spillovers obtained by both 'agreeing' and 'non-agreeing' firms. In order to capture these different effects, we borrow from a kind of functional form specifications which can often be found in RJV or R&D coalition literature (e.g. Katz 1986; D'Apremont and Jacquemin 1988; Kamien et al. 1992; Suzumura 1992; Beath et al. 1998) and we assume that the adoption of a VA affects both production and abatement costs (the latter in an indirect way, as will appear from equilibrium conditions).

(i) Let c denote the marginal production cost in the business-as-usual situation (i.e. when no VA is introduced) and let $e_{iA} - e_{i0}$ be the increase in the abatement level agreed by a signatory firm *vis-à-vis* the emission abatement level e_{i0} achieved when no VA is available to firms. Henceforth, we let e_{i0} be zero for the sake of simplicity.[14]
(ii) Let us also assume that the cost advantages (whether 'regulatory gains' or real efficiency gains) obtained by a VA signatory are proportional to the incremental abatement effort e_{iA}.
(iii) To allow for cases in which VAs imply significant participation costs, let us reduce the previous cost advantage by a share proportional to the individual abatement effort. Let d parameterize this possible marginal cost increase following the adoption of a VA.
(iv) To model the *VA participation externality*, let us relate a firm's cost advantage to the number of signatories by means of a positive parameter γ.
(v) Finally, let $K \in [1, n]$ be the number of signatory firms.

As a consequence, the VA participation externality in a symmetric equilibrium is $\gamma(K-1)e_{iA}$, and when all the above elements are added up, the expected total cost function of a signatory firm is:

$$TC_{iA} = (c + z_{iA})\, y_{iA} + \phi(e_{iA}), \tag{2}$$

[13] For an analysis of the relationship between VAs and the market for pollution abatement technologies see David and Sinclair-Desgagné (2005).

[14] It can be shown that $e_{i0} = 0$ is actually the equilibrium solution for our model in the absence of the VA option, given the simplified abatement cost function hypothesized. The solution could have been different had we allowed for the presence of other more complicated forms of marginal cost effects arising from voluntary abatement (e.g. end-of-pipe cost modifications).

where

$$z_{iA} = de_{iA} - \gamma \sum_{k=1}^{K} e_{kA}, \quad d \in [0, \gamma) \tag{3}$$

represents the net marginal cost advantages accruing to firm i only once it has signed the VA.[15] In order to set a limit to the extent of cost advantages, we allow for an upper bound e^{UB}, which can conveniently be expressed in emission abatement units. This implies that $z_{iA} \geq -e^{UB}$. In order to ensure positive production levels, we also assume that $z_{iA} \geq -c$. Without loss of generality and to keep the algebra as simple as possible, the value of the parameter γ can be normalized to 1 by introducing a parameter $\delta = d/\gamma$.[16] Hence, the expected profit function for firms which sign the VA is:[17]

$$\pi_{iA} = \left\{ P(Y) - \left[c + \delta e_{i0} - \sum_{k=1}^{K} e_{kA} \right] \right\} y_{iA} - \phi(e_{iA}), \quad \delta \in [0,\ 1] \tag{4}$$

Recalling that a share $\beta \in [0, 1]$ of the benefits produced by adopting the VA spills over onto the non-signatory firms, the latter therefore have the following cost function:

$$TC_{iNA} = \left[c - \beta \sum_{k=1}^{K} (e_{kA} - e_{k0}) \right] y_{iNA}, \quad \beta \in [0, 1] \tag{5}$$

where the potential business-as-usual abatement cost related to the potential abatement effort which both signatories and non-signatories may effect even in the absence of a VA has been considered equal to zero, having set $e_{i0} = 0$.

While the abatement cost function is the same regardless of whether or not the VA is adopted, non-signatories do not pay the costs of implementing the VA, which is δe_A. Thus, their expected profit function is:

$$\pi_{iNA} = \left\{ P(Y) - \left[c - \beta \sum_{k \neq 1}^{K} e_{kA} \right] \right\} y_{iNA} \tag{6}$$

Finally, let us look at the regulator's behavior. As we have said, the regulator may set a minimum abatement level ($\bar{e}$) in order to rule out the adoption of 'cosmetic

[15] We could add some structure to z_{iA} so as to formally distinguish between regulatory gains and genuine efficiency gains. While remarking again that our framework may comprise both effects, we prefer not to do so, in order to keep the analytical structure as simple as possible.

[16] Note that model's results are unaffected by d being equal to zero. Also note that we are imposing $\gamma \leq d$, and of course this must be true for a Cournot-Nash competition framework. Had we allowed for the case in which $d > \gamma$, no VA could be created unless we assume that firms collude in the choice of abatement level.

[17] We usually refer to expected costs and expected profits because the cost reduction achieved through the VA is relative to the expected costs under a different form of regulation, which becomes less likely as total voluntary abatement increases.

agreements'. The regulator makes this choice in the first stage of the game, by anticipating how the industry will behave in the subsequent stages. Hence, the regulator decides to bind the firm to an abatement level $e_{iA} \geq \bar{e}$ if its welfare function (defined as usual by the sum of industry profits, consumers' surplus and environmental gains), when a minimum binding abatement level is imposed, is larger than in those cases where signatory firms are free to adopt their profit-maximizing abatement strategy.[18]

3 Equilibrium coalitional VAs

The equilibrium of the game described in the previous section is computed by moving backwards from the last stage. First recall that when no VA is adopted, the equilibrium output coincides with the standard Cournot production level. When a VA is proposed to firms and the first K firms sign it ($K = 1, 2, \ldots, n$), the expected profit functions can then be re-written as:

$$\pi_{iA}(K) = P(Y)\, y_{iA} - \left(c - (1-\delta)\, e_{iA} - \sum_{k \neq i}^{K} e_{kA} \right) y_{iA} - \phi(e_{iA}) \quad \text{for } i=1, \ldots, K \tag{7a}$$

$$\pi_{iNA}(K) = P(Y)\, y_{iNA} - \left(c - \beta \sum_{k=1}^{K} e_{kA} \right) y_{iNA} \quad \text{for } i=K+1, \ldots, n \tag{7b}$$

If no firms sign the VA, equilibrium profits are those of the standard Cournot-Nash competition framework, i.e.

$$\pi_{iNA}(0) = [P(Y) - c]\, y_{iNA} \tag{8}$$

The differentiation of (7a) and (7b), respectively w.r.t. y_{iA} and e_{iA} and w.r.t. y_{iNA}, yields the following first order conditions[19] which, jointly with the condition $e_{iA} \leq e^{UB}$, define the Nash equilibrium values $(y^*_{iA}, y^*_{iNA}, e^*_{iA})$ of the third stage of the game contingent on K firms having signed the VA in the second stage:

$$P'(Y)\, y_A + P(Y) = c + (\delta - K)\, e_A \tag{9a}$$

$$(1-\delta)\, y_A = \phi'(e_A) \tag{9b}$$

$$P'(Y)\, y_{NA} + P(Y) = c - \beta K e_A \tag{9c}$$

where $\delta < 1$ and $Y = K y_A + (n-K)\, y_{NA}$. The index i has been dropped given the symmetry of the equilibrium. Note that the equilibrium production and abatement level $(y^*_{iA}, y^*_{iNA}, e^*_{iA})$ are a function of K, the number of firms that decide to sign

[18] Following a useful insight by an anonymous referee, in Sect. 4.1 we discuss how this situation is linked to one where the regulator can vary the level of an industry-wide target.

[19] The solutions to this problem are internal. Hence, we do not discuss the weak inequality of the Kuhn-Tucker conditions.

the VA. To clarify the notation, here and in the sequel we shall write this relationship explicitly.

The equilibrium outputs are obtained by first solving (9b) for y_A:

$$y_A = \frac{\phi'(e_A)}{1-\delta} \equiv \varphi^{-1}(e_A) \tag{10}$$

from which $e_A = \varphi(y_A)$. Then, considering the upper bound constraint, the equilibrium value y_A^* is:

$$y_A^* = \varphi^{-1}\left(e_A^*\right), \quad \text{where} \quad e_A^* \equiv \psi\left(y_A^*\right) = \min\left\{\begin{matrix}\varphi(y_A) \\ e^{UB}\end{matrix}\right. . \tag{11}$$

Given that abatement costs are assumed to be convex, the abatement level is a monotonically non-decreasing function of output.[20] Substituting the expression for Y^* and e_A^* in (9a) and (9c) and using symmetry again, we get:

$$P'\left(Ky_A^* + (n-K)y_{NA}^*\right)y_A^* + P\left(Ky_A^* + (n-K)y_{NA}^*\right) = c + (\delta - K)\,\psi\left(y_A^*\right) \tag{12a}$$

$$P'\left(Ky_A^* + (n-K)y_{NA}^*\right)y_{NA}^* + P\left(Ky_A^* + (n-K)y_{NA}^*\right) = c - \beta K\psi\left(y_A^*\right). \tag{12b}$$

Equations 12a and 12b implicitly define the equilibrium output levels which determine the equilibrium abatement levels. These equilibrium values (all dependent on K) can then be substituted in the profit functions (7a) (7b) to analyze the equilibrium of the second stage of the game where firms decide whether or not to sign the VA, i.e. the stage where the equilibrium number of signatories, K^*, is determined.

Second stage of the game

Let $\pi_{iA}^*(K) \geq 0, \pi_{iNA}^*(K) \geq 0$ be the values of the profit functions, for signatories and non-signatories respectively, when output and abatement levels take the equilibrium values determined by Eqs. 12a, 12b and 11. A necessary condition for the existence of an equilibrium coalition is that profits for VA signatories are greater than their profits when no VA is signed, i.e.

$$\Pi_{iA}\left(K^*\right) \equiv \pi_{iA}^*\left(K^*\right) - \pi_{iNA}^*(0) \geq 0 \quad \text{for all } i = 1, 2, \ldots K \tag{13}$$

This profitability (or rationality) condition must hold jointly with the so-called stability conditions (Cf. D'Apremont et al. 1983; Carraro and Siniscalco 1993):

$$\pi_{iA}^*\left(K^*\right) \geq \pi_{iNA}^*\left(K^*-1\right), \quad \text{for all } i = 1, 2, \ldots, n. \tag{14a}$$

$$\pi_{iA}\left(K^*+1\right) \leq \pi_{iNA}\left(K^*\right), \quad \text{for all } i = 1, 2, \ldots, n. \tag{14b}$$

[20] This result confirms the general applicability of the abatement cost specification that we have chosen.

As can be seen, by regarding competitors' decisions as given, in the stability conditions each firm evaluates its own advantages of leaving or joining a VA by considering unchanged participation or non-participation decisions by other firms. The profitability and stability conditions are necessary and sufficient to define the Nash equilibrium of the second stage of the game. Hence, when players are symmetric, the equilibrium K^* is the largest admissible integer smaller than the unique root of the following equation:

$$L\left(K^*\right) \equiv \pi_A\left(K^*\right) - \pi_{NA}\left(K^* - 1\right) = 0, \tag{15}$$

where $L(K^*)$ is such that $L'(K^*) < 0$, $\Pi_A(K^*) \geq 0$ for all $K \in [1, n]$, and the index i has been dropped because of symmetry.[21]

Note that, applying (14), if $L(K) < 0$ for all $K \in [1, n]$, then there is only a trivial equilibrium coalition, because the incentive to free-ride—$\pi_{NA}(K-1) - \pi_A(K)$—is positive for all $K \in [1, n]$.

Under the above fairly general functional forms, it is not possible to derive an analytical expression for K^*. Nonetheless, in the following we argue that general conclusions can be drawn. We will show that if $L(K) > 0$ for all K larger than a given value $K^\wedge$, where $1 \leq K^\wedge < n$, then the 'grand coalition' ($K^* = n$) forms at the equilibrium, provided that a coordination mechanism (e.g. a minimum participation constraint) guarantees that at least $K^\wedge$ firms join the coalition.[22] We summarize the positive results of the model in the following:

Proposition 1 *Under open membership, if benefits for signatories are partially excludable ($\beta < 1$), the stability function L(K) becomes positive for some $K \geq K^\wedge$ and its slope is positive. This implies $K^* = n$, i.e. all firms in the industry decide to sign the VA, provided that $K^\wedge \leq n$. If benefits for signatories are not excludable ($\beta = 1$), no firm adopts the VA.*

Proof The proof of the proposition can be divided in two main steps or lemmas:

(i) **Lemma 1** *If $\beta < 1$, there exists a value $K^\circ = \delta/(1-\beta)$ such that, for all $K \geq K^\circ (K \leq n)$, the output of non-signatories is smaller than signatories, provided that $K^\circ \leq n$. For $\beta = 1$, the output of non-signatories will be always greater than signatories.*

(ii) **Lemma 2** *There exists a value $K^\wedge \geq K^\circ$ for which the stability condition is satisfied, and above which it is always positive, so that in equilibrium $K^* = n$ firms adopt the VA, provided that $K^\wedge \leq n$.* □

[21] $L(K)$ is usually referred to as the "stability function" (e.g. Carraro and Marchiori 2003; Hoel and Schneider 1997).

[22] The presence of minimum participation constraint is very common in international environmental treaties. According to Rutz (2001), only 2 out of the 122 multilateral environmental agreements provided by the Center for International Earth Science Information Network do not contain any minimum participation rule. In 81 cases, the participation rule asks for a minimum number of signatories. In 22 cases, unanimity is required for the treaty to come into force, namely all negotiating countries must sign and ratify the agreement for it to be effective. In the remaining 17 cases, the minimum participation rule is coupled with other requirements, i.e. for these agreements it is not sufficient for a certain number of countries to ratify the treaty, but these countries also have to satisfy other, additional criteria. Interestingly, we are not aware of equivalent rules in the case of firm-level VAs.

Proof of Lemma 1 Subtracting Eq. 12b from 12a yields:

$$P'\left(Y^*\right)\left(y_A^* - y_{NA}^*\right) = \left[\delta - (1-\beta)\,K\right]\psi\left(y_A^*\right), \tag{16}$$

which implies that firms signing the VA produce a higher output if and only if they have lower marginal production costs.[23] This is the case when $\delta < (1-\beta)\,K$, or equivalently:

$$K > \frac{\delta}{1-\beta} = K^{\circ}. \tag{17}$$

Hence, only when the group of signatories is sufficiently large, can a signatory produce more than a non-signatory. In this case, a signatory also has a larger market share. Note that the positive relationship between profits and equilibrium output implies that (17) is a necessary condition for the (internal) stability condition (14a) to be met.[24]

Also note that (17) is more easily satisfied when the spillover rate β is low and the VA-related cost parameter δ is also low. For high values of β and δ, (17) is never likely to be satisfied (i.e. $K^{\circ} > n$). This is always true for $\beta = 1$, since from (17) we have $K^{\circ} = \infty$. The fact that non-signatory firms always produce at least as much as signatory firms implies that $\pi_A^*(K) - \pi_{NA}^*(K) < 0$ for all $K \in [1, n]$. In turn, the negativity of the stability function in the whole support of K implies that no firm is willing to sign the VA. Hence, as the cost effects of the VA approach those of a pure public good, the incentive to free-ride is always greater than the benefits deriving from adhering thereto.

Proof of Lemma 2 When a value $K^{\circ} \leq n$ exists, first note that a lower marginal cost is only a necessary condition for signatories to have larger profits than non-participants, because signatories also pay for the emission abatement costs $\varphi(e)$. This implies that profits of signatories are larger than non-signatories only for a value of K above K°. Let us denote this value with $K^{\wedge}$, where $K^{\wedge} \geq K^{\circ}$.

This value generally exists independently of whether profits increase monotonically with K for all $K = 1, 2, \ldots, n$ or not, and the characteristics of the profits functions are such that once the critical participation threshold $K^{\wedge}$ is reached, in equilibrium $K^* = n$ firms join the VA. Figures 1 and 2 below exemplify these two cases, showing the possible shapes of the functions making up the internal stability function, i.e. $\pi_A(K)$ *and* $\pi_{NA}(K-1)$, whilst the formal proof is left to the Appendix.

In both cases, for $K > K^{\wedge}$ the profits of signatories are the highest. Hence, if $K^{\wedge} < n$, $L(K) > 0$ for $K^{\wedge} < K \leq n$, then there is an incentive to join the coalition for all $K \in [K^{\wedge}, n]$. The equilibrium coalition is thus the grand coalition ($K^* = n$), whenever a co-ordination mechanism is introduced which induces at least $K^{\wedge}$ firms to

[23] Note that y_A^* and y_{NA}^* are actually $y_A^*(K)$ and $y_{NA}^*(K)$. While we take this into account in the following analysis, we omit reporting the K argument so as to simplify notation.

[24] This conclusion is true independently of ϕ because Eq. 11 implies that equilibrium abatement is a monotonically non-decreasing function of output. Hence, the greater the function $\phi(e)$, the lower y^*; therefore the lower e^* and the value of $\phi(e^*)$.

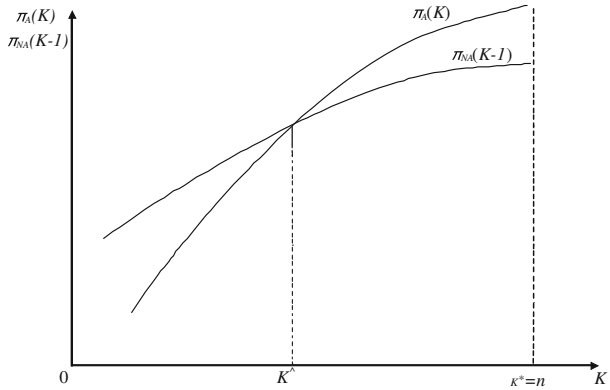

Fig. 1 Shape of functions $\pi_A(K)$ and $\pi_{NA}(K-1)$ for $\beta < 1$ when $\pi^*_{NA}(K)$ increases in $K^\wedge$

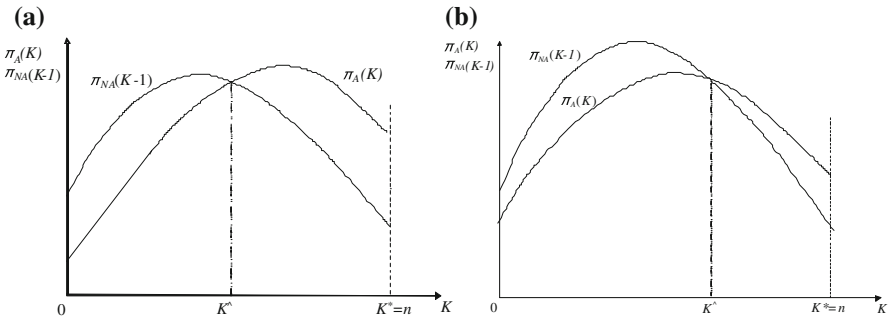

Fig. 2 Shape of functions $\pi^*_A(K)$ and $\pi^*_{NA}(K-1)$ for $\beta < 1$ when $\pi^*_{NA}(K-1)$ decreases in $K^\wedge$

sign the VA. The coordination mechanism could be a minimum participation constraint imposed by the regulator. (Cf. Carraro et al. 2009; Finus 2003).

3.1 Allowing for industry-wide targets

How can we compare our results with those obtained by D-S (2008), who actually state that a stable group of firms will adopt a VA in equilibrium even when non-signatories equally benefit from it? We claim that the difference between the two results is only apparent and determined by the absence or presence of an additional assumption on the coalition membership rule. In their paper, D-S actually argue that, due to the presence of an industry-wide target, "a level of participation that is not profitable will lead to a dissolution of the coalition". This would imply that each firm realizes that, for a K^P such that the profitability condition (13) is barely satisfied,[25] its defection would break the VA completely. Profits for non-signatories would therefore

[25] In the coalition theory jargon, K^P can be referred to as the *minimum profitable coalition size*, which is the smallest coalition size for which a member obtains a higher payoff than if all firms are independent (e.g. Block 2003).

be $\pi^*_{NA}\left(K^P - 1\right) = \pi^*_{NA}(0)$, and realizing this risk, firms belonging to a VA of size K^P would never defect.

When the whole coalition collapses when one member defects, "coalition unanimity" is assumed. Usually, the game-theoretic literature applies this rule to the whole support of K, and it is shown that with positive spillovers any $K \geq K^P$ would be an equilibrium of the game (e.g. Block 2003 for a survey). Only by making it possible that the threat of the whole coalition collapsing is real just for $K = K^P$,[26] can one obtain the minimum participation threshold K^P which satisfies the profitability condition as the equilibrium of the game.[27] The case of an industry target set by the regulator which firms know must be met in order to access VAs cost advantages nicely implements coalitional unanimity.[28]

In fact, if we represent D-S's analysis in our framework considering 'local' coalition unanimity in the neighborhood of K^P in the form of a minimum industry-wide target,[29] we get the following partial participation equilibrium results.

Proposition 2 *When firms' strategies are such that "coalition unanimity rule" for VA membership is adopted at the minimum profitable participation level K^P (for example because of an industry target which is required in order for VAs benefits to come into place), the following results hold:*

(i) *For $\beta = 1$, a VA with K^P participant firms is an equilibrium of the game.*
(ii) *For $\beta < 1$, $K^* = K^P$ is an equilibrium if K^P is lower than $K^\wedge$, the minimum participation threshold for which signatories' profits are larger than non-signatories'; $K^* = n$ otherwise.*

Proof – For $\beta = 1$ The internal stability condition is satisfied since the VA becomes null and void if any firm pulls out of the agreement, i.e. $\pi^*_A\left(K^P\right) \geq \pi^*_{NA}\left(K^P - 1\right) = \pi^*_{NA}(0)$. The external stability condition $\pi_{iA}\left(K^P + 1\right) \leq \pi_{iNA}\left(K^P\right)$ is satisfied since from the proof of Lemmas 1 and 2 we have $\pi^*_A(K) - \pi^*_{NA}(K) < 0$ and $\frac{d\pi^*_A}{dK} < \frac{d\pi^*_{NA}}{dK}$ for all $K \in [1, n]$ (Cf. inequality A.5.b).

– For $\beta < 1$. If $K^P < K^\wedge$, this part of the proposition is satisfied using the above reasoning. If $K^P \geq K^\wedge$, non-signatories have the incentive to join the VA, i.e. $\pi_{iA}\left(K^P + 1\right) \geq \pi_{iNA}\left(K^P\right)$ for all $K \in [K^\wedge, n]$. As a consequence, the grand coalition $K^* = n$ is achieved in equilibrium.

We can also re-interpret the last part of the previous result with the following Corollary 1:

[26] Since for K^P an industry target can be barely collectively satisfied, each firm knows that, if it defects, the whole VA collapses since the remaining firms would be unable to meet the target. Formally, this implies that $\pi^*_{NA}\left(K^P - 1\right) = \pi^*_{NA}(0)$.

[27] Dissatisfaction with simultaneous coalition games has led the coalition formation literature to propose structures of 'sequential games' of coalition formation. Considering the case of positive spillovers, the unique equilibrium in these games is actually K^P (Block 2003).

[28] We thank an anonymous referee for having drawn our attention to this point.

[29] An additional assumption they make is the orthogonality of the reactions function of non-signatories, which in our model is equivalent to setting $d\pi^*_{NA}/dK = 0$ for $K \epsilon [1, n]$, $K \neq K^P$.

Corollary 1 *If benefits for signatories are partially excludable ($\beta < 1$), an industry-wide target can yield the same full participation result summarized in Proposition* 1 *provided that the target is set at a sufficiently high level.*

Within our framework, it is therefore possible to attain partial participation equilibriums also when spillovers are partially appropriable. Note that, in addition to equity considerations, in this case an inefficiency clearly arises. As similarly remarked by D-S in the absence of appropriable advantages, the fact that emission reduction costs are assumed to be weakly convex implies that a VA with K^P participants produces higher total abatement costs than where all the n firms would have signed it. However, in the present framework, this unsatisfactory equilibrium can be replaced by an efficient full participation situation either where at least $K^\wedge$ firms sign the VA, or a 'tough enough' industry-wide target is chosen, (i.e. such that $K^P \geq K^\wedge$). This leads us to a more general series of policy recommendations which are summarized in the next Section.

4 The design of VAs

The characterization of participation incentives described in the previous section has direct and important policy implications on the design of VAs in order to possibly achieve full participation. Three recommendations follow from our results. First, the regulator should design the VA in such a way that most of the benefits are reaped by signatories, thus minimizing spillovers to non-signatories. This can be done, for example, through a policy-mix which penalizes (or excludes from some benefits) firms which do not join the VA.[30] Note that by minimizing β, the regulator not only encourages firms to join the VA, but also reduces the minimum participation constraint required to ensure that all firms adhere.

Second, the regulator should introduce a minimum participation constraint. Namely the agreement only enters into force if at least a minimum number of firms belonging to the industry decides to adhere. Given that this threshold value is inversely related to the share of VA benefits accruing only to signatories, the more difficult it is to discriminate VA's advantages between participants and non-participants, the greater the importance of this minimum participation constraint becomes.

Note that from the results summarized in Corollary 1 full participation can also be obtained by means of a sufficiently stringent industry-wide target. This possibility proves to be particularly useful also when, for institutional reasons, the spillovers in favor of non-participants cannot be modified and are so high that the implied minimum participation threshold $K^\wedge$ is larger than n. Thus, the only available option would be to establish a VA with inefficient partial participation K^P by means of a credible threat associated with an appropriate industry-wide target.

[30] In a similar vein, Lyon and Maxwell (2007) have recommended that in order to improve the effectiveness of EPA's PVPs, greater benefits in terms of public recognition should be granted only to those firms that join these programs during the initial phase of the program, when information collection and sharing with government takes place.

A third important policy implication is drawn from the likely hump-shape of the profit functions, as shown in Fig. 2. This shape is the result of the asymmetry produced by introducing the VA. Firms which sign the VA have a competitive advantage over other firms, because they enjoy lower marginal production costs (if $\beta < 1$ and when a stable coalition forms). This advantage implies a higher market share, and therefore higher profits, only if some firms are excluded from the VA. When all firms sign, they all enjoy the same expected reduction in costs and therefore have the same market share. Therefore, each firm desires to obtain the benefits provided by the VA, but not to share these benefits with all firms in the industry. Thus, firms have the incentive to form a coalition that is smaller than the one in which all firms participate.[31] This is not possible under an open membership rule whereby firms are free to join or leave the VA, but it may be possible if an exclusive membership rule applies, where access is conditional on the consensus of the other signatories (Cf. Yi 1997; Carraro and Marchiori 2003). Consequently, if the regulator's objective is to achieve the greatest environmental benefit, he/she should design the VA in such a way that firms are unable to adopt rules or behaviors that enable them to exclude other firms therefrom.[32]

All the above mentioned policy design directions can be summarized in the following proposition, which directly follows from the analysis carried out in the previous section

Proposition 3 *If the regulator wants to ensure an efficient VA participation equilibrium in which all firms adopt the VA, then: a) benefits for signatory firms must be partially excludable, b) an adequately defined minimum participation constraint* $K \geq K^{\wedge}$*(or alternatively an industry-wide target such that* $K^P \geq K^{\wedge}$*) must be imposed, c) an open membership rule must be enforced for the VA.*

4.1 Welfare enhancing mandatory abatement constraints

So far, we have only stressed that if VAs are properly designed, they will be signed by all firms in the industry even if some free-riding may occur. However, their effectiveness in terms of environmental control remains uncertain, given that the profit maximizing abatement level set by signatory firms may be too low compared to the level of the environmental externality. This raises the question of whether a policy mix that incorporates a minimum abatement constraint associated with the possibility of firms having access to a VA scheme, cannot only enhance the VA environmental effectiveness, but also improve welfare.

From a pessimistic viewpoint, one may argue that since a minimum abatement constraint increases the abatement effort required by firms, it will also reduce their output and profits and could raise equilibrium prices, thus resulting in an overall reduction in

[31] Examples exist in which the EC Directorate-General Competition has made inapplicable the provision of a VA which restricted some advantages only to original members. e.g. see Gremminger et al. (2001) on the *Eco-Emballages* waste recovering system.

[32] The case is more ambiguous if the regulator aims to maximise total welfare, since the greater environmental benefit achieved when all firms participate in the VA and the higher consumer surplus resulting from price reductions, must be measured against the lower total profit in the industry.

welfare (recall that welfare is defined as the sum of industry profits, consumer surplus and environmental benefits). Hence, a minimum abatement level may not be welfare improving.

From an optimistic viewpoint, (i) either environmental benefits may outweigh profits and consumer surplus losses from price increase (in the literature on voluntary overcompliance, this is the case for models with vertically differentiated firms; e.g. Arora and Gangopadhyay 1995; and Lutz et al. 2000); (ii) or the introduction of an additional bias may force firms to "pursue their own interests" by diverting them from adopting a Cournot-Nash behavior in the choice of voluntary abatement levels.

The existence and the size of the latter 'second best argument' can be analyzed by solving the regulator's decision problem in the first stage of our policy framework. Having ascertained that in order to promote full participation, the regulator should adopt the recommendations summarized in Proposition 3, we now come back to the question whether he or she should choose a VA in which firms are free to set their profit maximizing abatement level or a VA in which a minimum abatement level is imposed. The regulator decides to impose a minimum abatement requirement only if: (i) the profit maximizing abatement level is below the minimum abatement standard; and (ii) total welfare increases in the presence of a minimum abatement level.

In order to see whether the second condition can be met, let us analyze a firm's choice in the third stage of the game when the VA requires the achievement of a minimum abatement level $\bar{e}_A$. If $\bar{e}_A < e_A^*$, then the equilibrium is the one previously derived. *Vice versa*, if $\bar{e}_A \geq e_A^*$, then a firm only chooses output in the third stage of the game, and the equilibrium is defined by the standard Cournot output choices. Indeed, if emissions by the VA adopting firms are equal to $\bar{e}_A$, the profit functions become:

$$\bar{\pi}_A = P\left(\bar{Y}\right) y_A - \left(c + \delta\bar{e}_A - K\bar{e}_A\right) y_A - \phi\left(\bar{e}_A\right), \tag{18a}$$

$$\bar{\pi}_{NA} = P\left(\bar{Y}\right) y_{NA} - \left(c - \beta K\bar{e}_A\right) y_{NA}. \tag{18b}$$

The equilibrium first order conditions are therefore:

$$P'\left(\bar{Y}^*\right)\bar{y}_A^* + P\left(\bar{Y}^*\right) = c + \left(\delta - K\right)\bar{e}_A \tag{19a}$$

$$P'\left(\bar{Y}^*\right)\bar{y}_{NA}^* + P\left(\bar{Y}^*\right) = c - \beta K\bar{e}_A, \tag{19b}$$

where the upper bar denotes the equilibrium values when a minimum abatement constraint exists.

Similarly to the unconstrained case, the previous equilibrium conditions are a prerequisite for solving the second stage of the game, in which firms decide whether or not to sign the VA proposed by the regulator. In order to determine K^*—the Nash equilibrium number of firms which join the VA—we must compare profits of signatory and non-signatory firms. The results for the equilibrium coalition are essentially identical to those summarized by Proposition 1, given that the equilibrium conditions (19a) and (19b) are analogous to conditions (9a) and (9c) when no abatement constraints exist. Hence, the presence of a minimum abatement level does not modify the shape of the profit functions of signatories and free-riders. Therefore, in equilibrium, $K^* = n$.

This finally allows us to move on to the first stage of the game, by comparing outputs and profits in the presence and absence of an abatement constraint and by checking whether an interval exists in which welfare is an increasing function of the level of mandatory abatement. With the minimum abatement constraint, output is clearly higher since it is a decreasing function of marginal costs, which are lower under a binding minimum constraint.

As for profits, note that they are identical in the two cases when $\bar{e}_A = e_A^*$. Moreover, profits for signatory firms in the unconstrained regime are maximized when $e_A = e_A^*$; thus to compare profits, we just need to evaluate the derivative $\frac{d\bar{\pi}_A}{d\bar{e}_A}$ at $\bar{e}_A = e_A^*$. By differentiating (18a) at the Nash equilibrium with respect to $\bar{e}_A$ and using the envelope theorem (so that $\frac{\partial \bar{\pi}_{iA}}{\partial \bar{y}_{iA}^*} \frac{d\bar{y}_{iA}^*}{d\bar{e}_A} = 0$) we have:

$$d\bar{\pi}_{iA} = P'\bar{y}_A^* \left[(K-1)\, d\bar{y}_A^* + (n-K)\, d\bar{y}_{NA}^*\right] + \left[(K-\delta)\, \bar{y}_A^* - \phi'(\bar{e}_A)\right] d\bar{e}_A.$$

Using the generalized reaction function $\bar{R}_{iA}$ defined in the Appendix, this yields

$$\frac{d\bar{\pi}_{iA}}{d\bar{e}_A} = P'\bar{y}_A^* \bar{R}_{iA} \frac{dy_A^*}{d\bar{e}_A} + (K-\delta)\, \bar{y}_A^* - \phi'(\bar{e}_A)\,. \tag{20}$$

At the equilibrium of the second stage of the game, i.e. $K^* = n$, we have:

$$d\bar{\pi}_{iA} = P'\bar{y}_A^* (n-1)\, d\bar{y}_A^* + \left[(n-\delta)\, \bar{y}_A^* - \phi'(\bar{e}_A)\right] d\bar{e}_A,$$

from which:

$$\frac{d\bar{\pi}_{iA}}{d\bar{e}_A} = P'\bar{y}_A^* \bar{R}_A^n \frac{dy_A^*}{d\bar{e}_A} + (n-\delta)\, \bar{y}_A^* - \phi'(\bar{e}_A)\,, \tag{21}$$

where $\bar{R}_A^n$ is the 'standard' negative generalized reaction function of a symmetric Cournot oligopoly, and $dy_A^*/d\bar{e}_A$ is clearly positive as can be seen from (19a).

Hence, if the minimum standard is imposed when the grand coalition is the equilibrium coalition, profits of signatories may increase provided that:

$$\phi'(\bar{e}_A) < (n-\delta)\, \bar{y}_A^* + P'\bar{y}_A^* \bar{R}_A^n \frac{dy_A^*}{d\bar{e}_A}. \tag{22}$$

It can easily be verified that introducing a minimum abatement requirement when $K < n$ does not change the quality of the result.[33] Hence, Eq. 22 identifies an interval in which profits may be an increasing function of the minimum mandatory abatement level. The size of this interval depends on the coalition size, equilibrium output (which is affected by the set of demand and supply parameters considered) and the magnitude of marginal abatement costs, which in turn are determined by the extent of the minimum abatement constraint due to convexity.

[33] Simply look at Eq. 20 to see that a positive interval also exists when $K < n$.

Condition (22) also holds for the case of an industry-wide target, thus providing further indications on how to set these targets. Namely, in this case each firm knows that by signing a VA it will incur abatement costs equal to $\phi\,(E/K)$, where E is the target to be collectively met. At the final equilibrium $K^* = n$, this obviously yields a per-firm burden $\bar{\bar{e}}_A = E/n$. Hence the effects on profits of a variation $\frac{d\bar{\bar{\pi}}_A}{dE}$ can be indifferently studied by totally differentiating (18a) at the Nash equilibrium with respect to $\bar{\bar{e}}_A$, which implies that the same condition (22) which applies to $\bar{e}_A$ is recovered. Hence, for sufficiently low marginal abatement costs, the regulator can increase social welfare by tightening the industry-wide target ahead of what would be required for merely achieving full participation.

The implications of Eqs. 21 and 22 are summarized in the following proposition.

Proposition 4 *If a coalitional VA forms by means of an adequately defined minimum participation constraint $K \geq K^{\wedge}$ (alternatively, an industry-wide target such that $K^P \geq K^{\wedge}$) social welfare might be increased by introducing a minimum binding abatement constraint (strengthening the industry target beyond that required for full firm participation only). In particular, for sufficiently low marginal abatement costs, there is an interval of abatement levels across which all components of the welfare function—including profits—increase when the VA is characterized by an appropriate minimum firms' specific abatement constraint (sufficiently high industry-wide target).*

Proof . If there is an appropriate minimum binding abatement constraint, i.e. (22) is met, then:

(a) environmental benefits are greater, because total abatement is greater than it would be were firms free to set their own abatement levels.
(b) consumer surplus is higher, given that the greater reduction in costs when an emission abatement constraint is in force entails a higher level of production.
(c) firms' profits are higher, since the regulator will choose a minimum abatement level which satisfies condition (22).

The reason behind the fact that a mandatory abatement is required in order to increase production (and possibly profits) is to be found in the positive participation externality that the VA brings to the industry, whether in the form of genuine advantages like information diffusion about pollution prevention or simply avoiding alternative compulsory policies. This participation externality is related to the number of participating firms and to the total extent of their effort. When this external effect is present[34] and when a firm decides its abatement effort, it also takes into account the effect that its choice will have on the marginal costs of its competitors. In particular, it realizes that its decision to abate emissions will also benefit its competitors. This obviously diminishes the firm's propensity to abate emissions.

The Cournot-Nash competition scheme will consequently lead to an under-provision of abatement efforts—as is usually the case when a positive externality is present—because firms take the abatement choice of other members of the VA as given,

[34] Note that this effect is analogous to the 'competitive-advantage externality' in the RJV literature (Kamien et al. 1992).

instead of coordinating their abatement strategies. A minimum abatement constraint is simply a way of persuading firms to produce this external effect, moving them closer to co-operative abatement levels. In this case, although abatement costs increase, there may be an interval of emission abatement levels in which the cost increase is dominated by the higher revenue generated by higher production, which yields higher profits. Not surprisingly, the models which have unveiled this sort of effects for emission taxation (e.g. Dung 1993; Carraro and Soubeyran 1996) and quality standards (e.g. Farzin 2003) are developed in an oligopolistic framework and in the presence of free-riding incentives.

5 Conclusions

The main objective of this paper was to identify the conditions under which a group of firms in a given industry consider signing a VA the best option even when there are incentives to free-ride. These incentives arise from the non-excludability of the benefits achieved by firms adhering to the VA. Non-signatories may indeed gain all or part of these benefits as well. What conditions encourage a firm to participate in a VA is a crucial consideration to make for the proper design of VAs in an oligopolistic framework.

We have seen that the equilibrium of the game, which describes the interaction between regulators and firms and among all firms in the industry, is characterized by several strategic incentives and external effects. The first externality is that produced by the (partial or total) public good nature of the VA. This is the origin of the aforementioned free-riding incentives. A second strategic incentive arises from the fact that the cost saving effectiveness of the VA may increase with the number of firms adhering thereto and the extent of their abatement. Finally, a third strategic incentive is created by the oligopolistic nature of the market, which leads firms to overproduce compared to the case where they would jointly maximize industry profits and also not perfectly co-ordinate their voluntary initiatives on emission abatement.

The first effects may suggest that a VA is unlikely to be signed or, if it is signed, only a small number of firms will agree to adhere (Cf. Carraro and Marchiori 2003; Barrett 1994). However, the second effect may offset the first, because firms know that the greater the number of signatories, the greater the benefits to be gained. We have shown that, under certain conditions concerning the extent of the spillover effect, the second effect can prevail, in particular when a minimum participation constraint is met. The third effect explains why a cut in production and/or increase in abatement levels produced by the VA may increase welfare and even industry profits.

We have analyzed the interplay of the abovementioned three strategic incentives in a coalitional game in which firms decide whether or not to participate in the VA. We have shown that, in the absence of mandatory requirements by the regulator, the following results apply. If the benefits that firms achieve from adhering to the VA are non-excludable, i.e. non-signatories get the same benefits as signatory firms, then no firms decide to sign the VA at the equilibrium, except for the 'industry-wide target case' studied by Dawson and Segerson (2008). If the benefits are partially excludable, i.e. signatories get more benefits that non-signatories, then an industry-wide VA is signed only if an appropriate minimum participation constraint is satisfied. Hence,

this constraint could be opportunely imposed by the regulator. However, if benefits are partially excludable, firms signing the VA may find it optimal to exclude other firms from participating in the VA, because of the strategic advantage that a VA provides to signatories in an oligopolistic market (as seen in Sect. 3, a VA produces both absolute and relative cost advantages). Finally, the optimal voluntary abatement effort produced by the firms in the VA may be smaller than the socially optimal one, i.e. signatory firms still tend to over-emit if they are not constrained by a minimum abatement requirement.

In the light of the above results, we can make some normative policy recommendations or provide some suggestions on how to optimally design a VA in an oligopolistic industry. First, the regulator should design the VA such that mainly the SigNaTory firms reap the related benefits (in terms of regulatory pre-emption, technological co-operation, financial and technical incentives, etc.). Free-riders should be at least partly excluded from these benefits (e.g. they are taxed whereas signatories are not; they do not receive the financial incentives, codified information, additional flexibility or public recognition granted to participants).[35] Second, the VA should contain a minimum participation clause, namely it comes into force only if a minimum number of firms accepts to adhere thereto. This minimum participation constraint reduces free-riding incentives (Cf. Carraro et al. 2009) and enables firms to achieve a level of benefits that creates a band-wagon effect, i.e. when the minimum number of signatories is achieved, all remaining firms find it optimal to sign the VA as well. Note that we have seen that full participation can similarly be obtained by setting a sufficiently strong mandatory target at the industry level. Also note that the greater the share of benefits reaped by the signatories, the lower the minimum participation constraint will be. Third, if the first two recommendations hold, then the VA may oblige signatory firms to achieve at least a minimum individual mandatory abatement level. This can prove useful not only for avoiding negligible emission reductions, but also for partly offsetting the negative effects on production and profits of non-cooperative (oligopoly-specific) behavior regarding abatement levels. Finally, the regulator may decide to guarantee open access to the VA in order to exclude an anti-competitive, discriminatory use of the VA by some firms in the industry.

In view of the previous considerations, one can see that an optimal VA design is actually a policy mix design, that is an environmental policy where the co-existence of voluntary and 'traditional' (whether 'command and control' or 'economic') elements should be foreseen. From this viewpoint, some of the negative reputation accumulated by voluntary agreements can probably be attributed to past implementations characterized by a naïve trust in a merely self-enforcing kind of mechanism, rather than to the intrinsic unsuitability of this policy tool for implementing environmental regulation policies.

Further research could explore the issue of the optimal design of VAs in different contexts still characterized by intra-industry spillovers. First, the consideration of a

[35] Given that one of the goals of many VAs is to spread information on cost-effective abatement practices across the whole industry, a trade-off between information sharing and participation incentives is likely to arise. One possible solution suggested by Lyon and Maxwell (2007) could be to foresee asymmetric advantages for the VA participants at least in its early phases.

framework where the regulator has an imperfect control (or heterogeneous monitoring costs) could provide further explanations for the case where formally only some firms sign the VAs, or where there is *de facto* partial participation since only some firms actually undertake substantial emission abatement. Moreover, it would be interesting to analyze the case of 'green reputation enhancing' VAs, i.e. where the VA is adopted by firms because of its positive effects on market demand. Finally, our analysis is based on the assumption of symmetric firms. The adoption of VAs in an industry with asymmetric firms (e.g. with different efficiency levels) could be a topic of interest.

Acknowledgments The Authors are grateful to the Editor and two anonymous referees for many helpful comments and suggestions which greatly improved the paper. All remaining errors are obviously ours.

Appendix

Proof of Lemma 2

Let us study the behaviour of the profit functions at the Nash equilibrium. In general, we have:

$$\frac{d\pi_i^*}{dK} \equiv \frac{d\pi\left(y_i^*, y_{-i}^*, e_A^*, K\right)}{dK} = \left(\sum_{j\neq i} \frac{\partial \pi_i^*}{\partial y_j^*} \frac{dy_j^*}{dK}|_{y_A=y_A^*, y_{NA}=y_{NA}^*, e_A=\psi\left(y_A^*\right)}\right) + \frac{\partial \pi_i^*}{\partial K}|_{y_A=y_A^*, y_{NA}=y_{NA}^*, e_A=\psi\left(y_A^*\right)}, \tag{A.1}$$

Hence, by totally differentiating (7a) at the Nash equilibrium and using symmetry, we get:

$$d\pi_A^* = P'\left(Y^*\right) y_A^* \left\{(K-1)dy_A^* + (n-K)\, dy_{NA}^* + \left[\left(y_A^* - y_{NA}^*\right) + \frac{\psi\left(y_A^*\right)}{P'\left(Y^*\right)}\right] dK\right\}. \tag{A.2a}$$

where the index i has been omitted because of symmetry. Let $R_A = \frac{d}{dy_A^*}\left(\sum_{j\neq i}^n dy_j\right)$ denote the aggregate industry reaction function and recall that R_A is negative for the so-called generalized Hahn condition for a Cournot oligopoly (Dixit 1986). Using (16) to substitute for $\left(y_A^* - y_{NA}^*\right)$, the previous equation becomes:

$$d\pi_A^* = P'\left(Y^*\right) y_A^* \left\{R_A dy_A^* + \frac{[1+\delta-(1-\beta)\,K]\,\psi\left(y_A^*\right)}{P'\left(Y^*\right)} dK\right\}. \tag{A.2b}$$

Similarly, by totally differentiating (7b), we obtain:

$$d\pi_{NA}^{*} = P'\left(Y^{*}\right) y_{NA}^{*} \left\{ \left[K + \frac{\beta K \psi'}{P'\left(Y^{*}\right)} \right] dy_{A}^{*} + (n - K - 1)\, dy_{NA}^{*} + \left[\left(y_{A}^{*} - y_{NA}^{*}\right) + \frac{\beta \psi\left(y_{A}^{*}\right)}{P'\left(Y^{*}\right)} \right] dK \right\}. \tag{A.3a}$$

By using (16) and the definition of R_A, this becomes:

$$d\pi_{NA}^{*} = P'\left(Y^{*}\right) y_{NA}^{*} \left\{ R_A dy_{A}^{*} + \frac{\beta K \psi'}{P'\left(Y^{*}\right)} dy_{A}^{*} + \left(dy_{A}^{*} - dy_{NA}^{*}\right) + \frac{\left[\beta + \delta - (1 - \beta) K\right] \psi\left(y_{A}^{*}\right)}{P'\left(Y^{*}\right)} dK \right\}. \tag{A.3b}$$

Let us examine Eqs. A.2b and A.3b. First note that the terms within square brackets at the end of both expressions, respectively $[1 + \delta - (1 - \beta) K]$ and $[\beta + \delta - (1 - \beta) K]$, become negative for all $K > K^1 = (1 + \delta)/(1 - \beta)$ and $K > (\beta + \delta)/(1 - \beta) = K^1 - 1$ respectively, with $K^1 - 1 > K^{\circ}$, for $\beta < 1$.[36]

Then observe that the relationship between profits and signatories in the two equations depends on $\frac{dy_A^*}{dK}$, the derivative of a signatory's equilibrium production with respect to K, and on the difference $dy_{iA}^* - dy_{iNA}^*$. By assuming weakly convex demand functions and defining with K^A and K^{NA} the coalition size for which the equilibrium production of signatories and non-signatories is maximised, we now show that the following conditions hold:

$$\frac{dy_{iA}^{*}}{dK} \geq 0 \text{ for } K \in \left[K^{\circ}, K^{A}\right]; \frac{dy_{iA}^{*}}{dK} < 0 \text{ for } K > K^{A}, \text{ with } K^{A} \geq \frac{n}{2} + \frac{1 + \delta}{2(1 - \beta)} = \frac{n}{2} + \frac{K^{1}}{2} > K^{1}; \tag{A.4.a}$$

$$K^{NA} < K^{A} \tag{A.4.b}$$

$$dy_{iA}^{*} - dy_{iNA}^{*} > 0, \quad \text{for } K \in \left[K^{\circ}, n\right] \text{ and } \beta < 1 \tag{A.5.a}$$

$$dy_{iA}^{*} - dy_{iNA}^{*} < 0, \quad \text{for } K \in [1, n] \text{ and } \beta = 1; \tag{A.5.b}$$

[36] Hence, when considering $\pi_A^*(K)$ and $\pi_{NA}^*(K - 1)$, the last term within curly brackets in (A.2b) and (A.3b) has the same root $K^1 = (1 + \delta)/(1 - \beta)$.

To prove *inequality* (A.4.a), let us introduce the following notation in order to identify the well-known negative expressions which define the second order conditions of the profit maximisation problem in a Cournot oligopoly:

$$\begin{aligned} a_{iA} &\equiv P' + P'' y^*_{iA} + (K - \delta)\,\psi', \\ b_{iA} &\equiv P' + P'' y^*_{iA}, \\ a_{iNA} &\equiv P' + P'' y^*_{iNA}, \\ b_{iNA} &\equiv P' + P'' y^*_{iNA} + \beta K \psi' \end{aligned}$$

By focussing on weakly convex demand functions, we also have $P''(Y) \geq 0$.

Total differentiation of (12a) and (12b) yields:

$$\begin{aligned} \left(a_{iA} + P'\right) dy^*_{iA} + (K-1)\, a_{iA} dy^*_{-iA} + (n-K)\, b_{iA} dy^*_{NA} &= -\mu_K^{iA} dK, \\ \left(a_{iNA} + P'\right) dy^*_{NA} + (n-K-1)\, a_{iNA} dy^*_{-iNA} + K b_{iNA} dy^*_{A} &= -\mu_K^{iNA} dK; \end{aligned}$$

where symmetry has been exploited to add terms which are equal at the equilibrium. These equations can be re-written as:

$$\left(K a_{iA} + P'\right) dy^*_A + (n-K)\, b_{iA} dy^*_{NA} = -\mu_K^{iA} dK \tag{A.6}$$

$$\left((n-K)\, a_{iNA} + P'\right) dy^*_{NA} + K b_{iNA} dy^*_A = -\mu_K^{iNA} dK, \tag{A.7}$$

where μ_K^{iA} and μ_K^{iNA} are the direct effects of a variation of Kon the first order conditions (12a) and (12b):

$$\begin{aligned} \mu_K^{iA} &\equiv \frac{\partial^2 \pi_{iA}}{\partial y_{iA} \partial K} = \left(P' + P'' y^*_{iA}\right)\left(y^*_A - y^*_{NA}\right) + \psi \\ &= b_{iA}\left(y^*_A - y^*_{NA}\right) + \psi, \\ \mu_K^{iNA} &\equiv \frac{\partial^2 \pi_{iNA}}{\partial y_{iNA} \partial K} = \left(P' + P'' y^*_{iNA}\right)\left(y^*_A - y^*_{NA}\right) + \beta\psi \\ &= a_{iNA}\left(y^*_A - y^*_{NA}\right) + \beta\psi \end{aligned}$$

By applying Cramer's rule to (A.6) and (A.7) we have:

$$\frac{dy^*_A}{dK} = \frac{\begin{vmatrix} -\mu_K^{iA} & (n-K)\, b_{iA} \\ -\mu_K^{iNA} & (n-K)\, a_{iNA} + P' \end{vmatrix}}{\begin{vmatrix} K a_{iA} + P' & (n-K)\, b_{iA} \\ K b_{iNA} & (n-K)\, a_{iNA} + P' \end{vmatrix}}, \tag{A.8}$$

$$\frac{dy_{NA}^*}{dK} = \frac{\begin{vmatrix} Ka_{iA} + P' & -\mu_K^{iA} \\ Kb_{iNA} & -\mu_K^{iNA} \end{vmatrix}}{\begin{vmatrix} Ka_{iA} + P' & (n-K)\, b_{iA} \\ Kb_{iNA} & (n-K)\, a_{iNA} + P' \end{vmatrix}}. \tag{A.9}$$

Since the determinant of the denominator (denoted by *DET*) is positive (this is implied by the second order conditions for profit maximisation), the sign of (A.8) and (A.9) is determined by their numerators. In particular, we have:

$$\frac{dy_{iA}^*}{dK} = -\frac{1}{DET}\left[(n-K)\left(\mu_K^{iA} a_{iNA} - \mu_K^{iNA} b_{iA}\right) + \mu_K^{iA} P'\right]. \tag{A.10a}$$

Using (16) to substitute for $y_A^* - y_{NA}^*$, we get:

$$\frac{dy_{iA}^*}{dK} = -\frac{\psi}{DET}\left[P' + (n-K)\left(a_{iNA} - \beta b_{iA}\right) + b_{iA}\left[\delta - (1-\beta)\, K\right]\right], \tag{A.10b}$$

which is clearly positive in a generic interval $[K^\circ, K^A]$. For a weakly convex demand function, the lowest K^A value is obtained for a linear demand function. In this case, we get:

$$\frac{dy_A^*}{dK} = -\frac{\psi P'}{DET}\left[n\,(1-\beta) + 1 + \delta - 2\,(1-\beta)\, K\right], \tag{A.10c}$$

which is positive for $K < \frac{n}{2} + \frac{1+\delta}{2(1-\beta)} = \frac{n}{2} + \frac{K^1}{2}$.

Hence $\frac{dy_{iA}^*}{dK} \geq 0$ for $K \in [K^\circ, K^A]$, where $K^A > K^\circ$ in general and $K^A \geq \frac{n}{2} + \frac{1+\delta}{2(1-\beta)} = \frac{n}{2} + \frac{K^1}{2} > K^1$ for weakly convex demand functions.

In order to prove *inequalities* (A.4.b, A.5.a, A.5.b), note that, from (A.9), $\frac{dy_{iNA}^*}{dK}$ is equal to:

$$\frac{dy_{iNA}^*}{dK} = -\frac{1}{DET}\left[K\left(\mu_K^{iNA} a_{iA} - \mu_K^{iA} b_{iNA}\right) + \mu_K^{iNA} P'\right]; \tag{A.11a}$$

that is, recalling that $b_{iNA} = a_{iNA} + \beta K \psi'$,

$$\frac{dy_{iNA}^*}{dK} = -\frac{1}{DET}\left[K\left(\mu_K^{iNA} a_{iA} - \mu_K^{iA} a_{iNA} - \mu_K^{iA} \beta K \psi'\right) + \mu_K^{iNA} P'\right]. \tag{A.11b}$$

Subtracting (A.11b) from (A.10a) it follows that:

$$\frac{dy_{iA}^*}{dK} - \frac{dy_{iNA}^*}{dK} = -\frac{1}{DET}\left[K\left(\mu_K^{iA}\beta K\psi' - \mu_K^{iNA}\left(a_{iA} - b_{iA}\right)\right) + \left(\mu_K^{iA} - \mu_K^{iNA}\right)P' + n\left(\mu_K^{iA}a_{iNA} - \mu_K^{iNA}b_{iA}\right)\right] \tag{A.12a}$$

To find the sign of this difference, re-write n as $n + (n - K)$ and use again $b_{iNA} = a_{iNA} + \beta K\psi'$. Then:

$$\frac{dy_{iA}^*}{dK} - \frac{dy_{iNA}^*}{dK} = -\frac{1}{DET}\left[K\left(\mu_K^{iA}b_{iNA} - \mu_K^{iNA}a_{iA}\right) + \left(\mu_K^{iA} - \mu_K^{iNA}\right)P' + \psi\left(n - K\right)\left(a_{iNA} - \beta b_{iA}\right)\right] \tag{A.12b}$$

The three terms within square brackets are all negative for $K > K^\circ$ and $\beta < 1$. This can be seen more easily by expanding (A.12b) as follows:

$$\frac{dy_{iA}^*}{dK} - \frac{dy_{iNA}^*}{dK} = -\frac{K\left[\left(y_A^* - y_{NA}^*\right)\left(b_{iNA}b_{iA} - a_{iNA}a_{iA}\right) + \psi\left(b_{iNA} - \beta a_{iA}\right)\right]}{DET} + \frac{P'\left(\mu_K^{iA} - \mu_K^{iNA}\right) + \psi\left(n - K\right)\left(a_{iNA} - \beta b_{iA}\right)}{DET} \tag{A.12c}$$

Hence:

(a) $\frac{dy_{iA}^*}{dK} - \frac{dy_{iNA}^*}{dK} > 0$ for $K \in [K^\circ, n]$;

(b) As a corollary, given that $\frac{dy_{iA}^*}{dK} \geq 0$ for $K \in [K^\circ, K^A]$, the equilibrium output for non-signatories starts decreasing for a value of K such that $K^{NA} < K < K^A$ in the interval $[K^\circ, n]$.

When $\beta = 1$, Eq. A.12.c becomes:

$$\frac{dy_{iA}^*}{dK} - \frac{dy_{iNA}^*}{dK} = -\frac{K\left[\left(y_A^* - y_{NA}^*\right)\left(b_{iNA}b_{iA} - a_{iNA}a_{iA}\right) + \psi\left(b_{iNA} - a_{iA}\right)\right]}{DET} + \frac{\left(a_{iNA} - b_{iA}\right)\left[\psi\left(n - K\right) - P'\left(y_A^* - y_{NA}^*\right)\right]}{DET} \tag{A.13a}$$

That is, by using (16):

$$\frac{dy_{iA}^*}{dK} - \frac{dy_{iNA}^*}{dK} = -\frac{\psi}{DET}\left\{K\left[\frac{\delta}{P'}\left(b_{iNA}b_{iA} - a_{iNA}a_{iA}\right) + \left(b_{iNA} - a_{iA}\right)\right] + \left(a_{iNA} - b_{iA}\right)\left(n - K - \delta\right)\right\}. \tag{A.13b}$$

It can easily be verified that the two terms within curly brackets are positive, so that, for $\beta = 1$, $\frac{dy_{iA}^*}{dK} - \frac{dy_{iNA}^*}{dK} < 0$, or $dy_{iA}^* - dy_{iNA}^* < 0$, for $K \in [1, n]$.

The first implication of the above conditions on dy_A^*/dK (A.4) and on the difference $dy_{iA}^* - dy_{iNA}^*$ (A.5) is the following:

- *both π_A^* and π_{NA}^* may first increase and then decrease with K.*

Indeed, using (A.4) and (A.5) for studying (A.2b) and (A.3b) we see that profits for signatories and non-signatories certainly increase for $K \leq K^1$ and $K \leq K^1 - 1$, whereas $K > K^A$ is a sufficient condition to ensure decreasing profits in K, if $K^A < n$.[37]

The second implication of conditions (A.4–A.5) is that of allowing us to compare (A.2b) and (A.3b) and draw the following conclusions concerning the shape of the profit functions, which in turn determine the behaviour of the stability function, finally showing that in equilibrium $K^* = n$ firms adopt the VA.

(i) *For $K \geq K^\circ$ and $\beta < 1$, i.e. in the interval of K in which $K^\wedge$ can exist, profits of non-signatories start decreasing before profits of signatories.*
This can be seen by looking at the last term in (A.2b) and (A.3b), then noticing that $0 \leq \beta K \psi' / P'(Y^*) < 1$ for the stability condition (1a) and the non-decreasing relationship between output and abatement; finally that output of non-signatories starts decreasing before signatories. The same condition is even more easily verified between $d\pi_A^*(K)$ and $d\pi_{NA}^*(K-1)$, given that the last terms in (A.2b) and (A.3b) have the same root K^1.

(ii) *In the interval $[K^\wedge, n]$, any positive difference between $\pi_A^*(K)$ and $\pi_{NA}^*(K)$ increases with K up to $K = n$.*
In fact, in the interval $[K^\circ, n]$, $\frac{d\pi_A^*}{dK} > \frac{d\pi_{NA}^*}{dK}$. Comparing (A.2b) and (A.3b), this is readily seen when $K \leq K^A$, i.e. the equilibrium production of signatories increases in K. For $K > K^A$, the result is recovered by the presence of the positive term $\left[\frac{\beta K \psi'}{P'(Y^*)} + 1\right] dy_A^* - dy_{NA}^*$ within the curly brackets in (A.3b).

Results (i). and (ii). jointly imply the existence of $K^\wedge$ since the internal stability condition $\pi_A^*(K) \geq \pi_{NA}^*(K-1)$ is satisfied and $L(K)$ is positive for all $K \in [K^\wedge, n]$ for both the following cases:

(a) *the case where $\pi_{NA}^*(K)$ increases with K (so that $\pi_{NA}^*(K) > \pi_{NA}^*(K-1)$) in a neighbourhood of $K^\wedge$* (Fig. 1);
We know that $\pi_{NA}^*(K) > \pi_A^*(K)$ for $K < K^\circ$. However, for $K \geq K^\circ$, the positive difference between $d\pi_A(K)$ and $d\pi_{NA}(K)$ (and $d\pi_{NA}(K-1)$) allows for reversing the relationship between non-signatories' and signatories' profits for some K. *A fortiori*, since $\pi_{NA}^*(K) > \pi_{NA}^*(K-1)$, we also have that $\pi_A^*(K) \geq \pi_{NA}^*(K-1)$, i.e. $K^\wedge$ exists.[38] Conclusion (ii) ensures that any positive difference between profits increases in $[K^\wedge, n]$, hence $L(K)$ is positive up to $K = n$.

[37] This is a sufficient condition since, from (A.2b) and (A.3b), it can be easily seen that profits for signatories and non-signatories could start to decrease even for some smaller Ks in the intervals $[K^1, K^A]$ and $[K^\circ, K^A]$ respectively.

[38] Of course, as in the case of K°, whose value is $\delta/(1-\beta)$, there is no guarantee that $K^\wedge \leq n$.

(b) *the case where* $\pi_A^*(K)$ *and* $\pi_{NA}^*(K)$ *show opposite signs or decrease in K in a neighbourhood of* $K^\wedge$ (Fig. 2).
The fact that $\pi_{NA}^*(K)$ and $\pi_{NA}^*(K-1)$ start decreasing before $\pi_A^*(K)$ ensures that $\pi_{NA}^*(K-1)$ intercepts the profit function of signatories from above at $K^\wedge$, be it in the graphical example of Fig. 2a or when both functions decrease in K(Fig. 2b). Then conclusion (ii) again ensures that $L(K)$ is positive up to $K = n$.

References

Aggeri, F., & Hatchuel, A. (1999). A dynamic model of environmental policies. The case of innovation oriented voluntary agreements. In C. Carraro & F. Lévêque (Eds.), *Voluntary approaches in environmental policy*. Dordrecht: Kluwer Academic Publishers.

Alberini, A., & Segerson, K. (2002). Assessing voluntary programs to improve environmental quality. *Environmental and Resource Economics, 22*, 157–184.

Arora, S., & Gangopadhyay, S. (1995). Toward a theoretical model of voluntary overcompliance. *Journal of Economic Behaviour and Organisation, 28*, 289–309.

Carraro C., & Soubeyran, A. (1996). Environmental taxation, market share and profits in oligopoly. In C. Carraro, Y. Katsoulacos, & A. Xepapadeas (Eds.), *Environmental Policy and Market Structure*. Dordrecht: Kluwer Academic Publishers.

Bagnoli, M., & Watts, S. (2003). Selling to socially responsible consumers: Competition and the private provision of public goods. *Journal of Economics & Management Strategy, 12*, 419–445.

Baranzini, A., & Thalmann, P. (2004). *Voluntary approaches in climate policy*. Cheltenham: E. Elgar.

Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers, 46*, 878–894.

Barrett, S. (2002). *Environment and statecraft*. Oxford: Oxford University Press.

Beath, J., Poyago-Theotoky, J., & Ulph, D. (1998). Organization design and information-sharing in a research joint venture with spillovers. *Bulletin of Economic Research, 50*, 47–59.

Bjørner, T. B., & Jensen, H. H. (2002). Energy taxes, voluntary agreements and investment subsidies: A micro-panel analysis of the effect on Danish industrial companies' energy demand. *Resource and Energy Economics, 24*, 229–249.

Blackman, A., & Boyd, J. (2002). Tailored regulation: Will voluntary site-specific environmental performance standards necessarily improve welfare? *Southern Economic Journal, 169*, 309–326.

Block, F. (2003). Non-cooperative models of coalition formation in games with spillovers. In C. Carraro (Ed.), *The endogenous formation of economic coalitions*. Cheltenham: E. Elgar.

Brau, R., & Carraro, C. (2006). The economic analysis of voluntary approaches to environmental protection. In K. Deketelaere, L. Kreiser, J. Milne, H. Ashiabor, & A. Cavaliere (Eds.), *Critical issues in environmental taxation: International and comparative perspectives*. Oxford: Oxford University Press.

Carraro, C. (Ed.). (2003). *The endogenous formation of economic coalitions*. Cheltenham: E. Elgar.

Carraro, C., & Lévêque, F. (1999). The rational and potential of voluntary approaches. In C. Carraro & F. Lévêque (Eds.), *Voluntary approaches in environmental policy*. Dordrecht: Kluwer Academic Publishers.

Carraro, C., & Marchiori, C. (2003). Stable coalitions. In C. Carraro (Ed.), *The endogenous formation of economic coalitions*. Cheltenham: E. Elgar.

Carraro, C., Marchiori, C., & Oreffice, S. (2009). Endogenous minimum participation in international environmental treaties. *Environmental and Resource Economics, 42*, 411–425.

Carraro, C., & Siniscalco, D. (1993). Strategies for the international protection of the environment. *Journal of Public Economics, 52*, 309–328.

Carraro, C., & Soubeyran, A. (1996). Environmental taxation, market share and profits in oligopoly. In C. Carraro, Y. Katsoulacos, & A. Xepapadeas (Eds.) *Environmental policy and market structure.* Dordrecht: Kluwer Academic Publishers.

Cavaliere, A., & Frontoso Silvestri, F. (2002). Voluntary agreements as information sharing devices: Competition and welfare effects. In D. Piacentino & G. Sobbrio (Eds.), *Intervento pubblico e architettura dei mercati*. Milan: Franco Angeli.

D'Apremont, C., & Jacquemin, A. (1988). Cooperative and noncooperative R&D in an oligopoly with spillovers. *American Economic Review, 78*, 1133–1137.

D'Apremont, C., Jacquemin, A., Gabszewicz, J. J., & Weymark, J. (1983). On the stability of collusive price leadership. *Canadian Journal of Economics, 16*, 17–25.

David, M., & Sinclair-Desgagné, B. (2005). Environmental regulation and the eco-industry. *Journal of Regulatory Economics, 28*, 141–155.

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

Delmas, M. A., & Keller, A. (2005). Free riding in voluntary environmental programs: The case of the U.S. EPA WasteWise program. *Policy Sciences, 38*, 91–106.

Delmas, M. A., & Montes-Sancho, M. J. (2010). Voluntary agreements to improve environmental quality: Symbolic and substantive cooperation. *Strategic Management Journal, 31*, 575–601.

De Muizon, G., & Glachant, M. (2004). The UK's climate change levy agreements: Combining voluntary agreements with tax and emissions trading. In A. Baranzini & P. Thalmann (Eds.), *Voluntary approaches in climate policies*. Cheltenham, UK: Edward Elgar.

Dixit, A. (1986). Comparative statics for oligopoly. *International Economic Review, 27*, 107–122.

Dung, T. H. (1993). Optimal taxation and heterogeneous oligopoly. *Canadian Journal of Economics, 26*, 933–947.

Farzin, Y. H. (2003). The effects of emissions standards on industry. *Journal of Regulatory Economics, 24*, 315–327.

Farzin, Y. H., & Akao, K. (2006). Environmental quality in a differentiated duopoly. Fondazione Eni Enrico Mattei Working Papers No. 138.06

Finus, M. (2003). Stability and design of international environmental agreements: The case of global and transboundary pollution. In H. Folmer & T. Tietenberg (Eds.), *International yearbook of environmental and resource economics, 2002/3*. Cheltenham, UK: Edward Elgar.

Glachant, M. (1999). The cost efficiency of voluntary agreements for regulating industrial pollution: A Coasean approach. In C. Carraro & F. Lévêque (Eds), *Voluntary Approaches in Environmental Policy*. Dordrecht: Kluwer Academic Publishers.

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

Gremminger, M., Laurila, M., & Miersch, G. (2001). The commission defines principles of competition for the packaging waste recovery markets. *EC Competition Policy Newsletter*, 29–33.

Hoel, M., & Schneider, K. (1997). Incentives to participate in an international environmental agreement. *Environmental and Resource Economics, 9*, 153–170.

Kamien, M. I., Muller, E., & Zang, I. (1992). Research joint ventures and R&D cartels. *American Economic Review, 82*, 1293–1306.

Katsoulacos, Y., & Xepapadeas, A. (1996). Emission taxes and market structure. In C. Carraro, Y. Katsoulacos, & A. Xepapadeas (Eds.), *Environmental policy and market structure*. Dordrecht: Kluwer Academic Publishers.

Katz, M.L. (1986). An analysis of cooperative research and development. *Rand Journal of Economics, 17*, 527–543.

Konar, S., & Cohen, M. A. (2001). Does the market value environmental performance? *The Review of Economics and Statistics, 83*, 281–289.

Krarup, S., & Millock, K. (2007). The Danish agreements on industrial energy efficiency. In R. D. Morgenstern & W. A. Pizer (Eds.), *Reality check: The nature and performance of voluntary environmental programs in the United States, Europe and Japan.*. Washington, DC: RFF Press.

Lutz, S., Lyon, T. P., & Maxwell, J. W. (2000). Quality leadership when regulatory standards are forthcoming. *Journal of Industrial Economics, 48*, 331–348.

Lyon, T. P., & Maxwell, J. W. (2002). Voluntary approaches to environmental protection. In M. Franzini & A. Nicita (Eds.), *Economic institutions and environmental policy*. New York: Ashgate Publishing.

Lyon, T. P., & Maxwell, J. W. (2003). Self-regulation, taxation and public voluntary environmental agreements. *Journal of Public Economics, 87*, 1453–1486.

Lyon, T. P., & Maxwell, J. W. (2007). Public voluntary programs reconsidered. *Policy Studies Journal, 35*, 723–750.

Manzini, P., & Mariotti, M. (2003), A bargaining model of voluntary environmental agreements. *Journal of Public Economics, 87*, 2725–2736.

Morgenstern, R. D., & Pizer, W. A. (2007). *Reality check: The nature and performance of voluntary environmental programs in the United States, Europe and Japan*. Washington, DC: RFF Press.

OECD. (1999). *Voluntary approaches for environmental policy in OECD countries*. Paris: OECD.

OECD. (2003). *Voluntary approaches for environmental policy. Effectiveness, efficiency and usage in policy mixes*. Paris: OECD.

Rutz, S. (2001). Minimum participation rules and the effectiveness of multilateral environmental agreements. Working paper 01/22, Centre for Economic Research, SFIT, Zurich.

Seade, J. (1980). The stability of cournot revisited. *Journal of Economic Theory, 23*, 15–27.

Segerson, K., & Miceli, T. (1998). Voluntary environmental agreements: Good or bad news for environmental protection? *Journal of Environmental Economics and Management, 36*, 109–130.

Suzumura, K. (1992). Cooperative and noncooperative R&D in an oligopoly with spillovers. *American Economic Review, 82*, 1307–1320.

ten Brink, P., Morère, M., & Wallace-Jones, J. (2003). Negotiated agreements and climate change mitigation. In C. Carraro & C. Egenhofer (Eds.), *Firms, governments and climate policy: Incentive-based policies for long-term climate change*. Cheltenham, UK: Edward Elgar.

Wu, J. J., & Babcock, B. A. (1999). The relative efficiency of voluntary vs mandatory environmental agreements. *Journal of Environmental Economics and Management, 38*, 158–175.

Xepapadeas, A., & Passa, C. (2004). Participation in and compliance with public voluntary environmental programs: An evolutionary approach. FEEM Nota di Lavoro 67.04, Milan.

Yi, S. S. (1997). Stable coalition structures with externalities. *Games and Economic Behavior, 20*, 201–237.

Yi, S. S. (2003). Endogenous formation of economic coalitions. A survey of the partition function approach. In C. Carraro (Ed.), *The endogenous formation of economic coalitions*. Cheltenham: E. Elgar.

Copyright of Journal of Regulatory Economics is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.