TSE working paper series

12-340

Research Group: Environmental Economics and Natural Resources

October 2012

"Solving the GlobalWarming Problem: Beyond Markets, Simple Mechanisms May Help!"

Wilfried Sand-Zantman

and

David Martimort

Toulouse School of Economics (TSE) - Manufacture des Tabacs
Aile Jean-Jacques Laffont - 21, allée de Brienne - 31015 TOULOUSE Cedex 6
Tél : +33 (0) 5 61 12 85 89 - Fax : +33 (0) 5 61 12 86 37 - **www.tse-fr.eu** - contact@tse-fr.eu

Toulouse School of Economics

Solving the Global Warming Problem: Beyond Markets, Simple Mechanisms May Help![*]

David Martimort[†] and Wilfried Sand-Zantman[‡]

October 2012

Abstract: This paper discusses the feasibility and performances of simple mechanisms to implement international environmental agreements in the multilateral externalities context of global warming. Asymmetric information and voluntary participation by sovereign and heterogenous countries are key constraints on the design of those agreements. Mechanisms must prevent two sorts of free-riding problems - free riding in effort provision and free riding in participation. As markets might fail to solve simultaneously those two problems, we construct instead a simple menu of options that trades off the *provision of incentives for participating countries* and the *provision of incentives to participate*.With such mechanism, all countries voluntary contribute to a fund, although at different intensities, but only the most efficient ones effectively reduce their pollution below its "business as usual" level.

Keywords: Free-riding, environmental agreements, asymmetric information, mechanism design.

1 Introduction

Over the past twenty years, the long run consequences of human activities on environmental fundamentals and more specifically the anthropogenic origin of global warming have been a cause of raising concern. That period has seen flourish worrying reports which have repeatedly insisted not only on the environmental urgency to reach a worldwide agreement in controlling emissions, but also on the irreversible costs of not doing so.[1] Nevertheless, very little advances have been encountered in persuading

[*]We thank participants at the *Congress of the Canadian Economic Association* in Calgary where this paper was presented as a *State of the Art Lecture*. All errors are ours.

[†]Paris School of Economics-EHESS. Email: david.martimort@parisschoolofeconomics.eu

[‡]Toulouse School of Economics (GREMAQ-IDEI). Email: wsandz@tse-fr.eu

[1]See the Stern Review (2006).

the most polluting countries to enter into binding agreements that would push them to credibly reduce their emissions. The objective of this paper is to offer an economists' look at this major issue, highlight the fundamental difficulties in reaching such agreements and propose a simple mechanism that might help under some circumstances.

For economists, pollution is often refereed to as the prime textbook example of multilateral externalities. In the case of global warming of interest for this paper, each country may indeed benefit from not reducing its own emissions while at the same time it certainly suffers from the insufficient global level of emissions reached with those individual efforts. Environmental economics textbooks are full of solutions to reach efficiency in such contexts, either by means of markets, quotas, taxes or subsidies, under various circumstances. Yet, this body of works which brings rather optimistic news stands in sharp contrast with the difficulties to enforce an agreement on a particular mechanism in practice. This suggests that something might be missing in the textbook logic.

Indeed, a close look at the record of repeated failures in climate negotiations also suggests that the addition of three specific ingredients might hinder the design of environmental agreements. These ingredients are: heterogeneity, sovereignty and private information. First, and this is certainly a standard feature of most public good problems, the costs and (under rare circumstances the possible) benefits of global warming are unevenly distributed among countries. For some countries, decreasing emissions will come with hindered growth, reallocation costs within and across sectors and much internal political haggling. For others, implementation of restrictive policies may turn out to be easier because of less powerful vested interests in national political arenas or, taking an optimistic perspective, because it might open new growth opportunities (for instance by means of innovation in green technologies).[2] Heterogeneity matters.

Second, agreements are drafted in contexts surrounded by significant uncertainty. Beyond the fundamental uncertainty that underlies the physical processes involved in climate change and that affects equally all parties in their expectations about future evolutions, another source of uncertainty lies in the mere heterogeneity we just pointed out. Indeed, each participant to an agreement may retain private information on its exact costs of decreasing pollution and improving environmental conditions. In such contexts, countries might take advantages from others' incomplete information and exaggerate how costly they find to implement a given profile of emissions reductions. Incentives matter.

Lastly, a key constraint on feasible agreements is that negotiating parties are sovereign countries. This feature, which is certainly more specific to the provision of transna-

[2]Similarly, the costs associated to climate change may depend on the location of the countries. As an example, the tribute paid by the Maldives and Switzerland following a one meter increase of the sea level is hardly comparable

tional public goods,[3] has two immediate implications. Not only each country must find it optimal to voluntary participate to an agreement but it is free to adopt a "business as usual" (thereafter BAU) route otherwise. To be accepted worldwide, an agreement must yield to joining countries greater payoffs than these BAU options. Participation matters.

Our Contribution. Taking heterogeneity, incentives and participation as key constraints on agreements, we discuss the performances of two institutions. The first institution, which has received much attention in practice, is of course the *market mechanism*. Opening a market for pollution rights is indeed a natural solution to the multilateral externalities problem in some contexts. Think for instance of a national government willing to induce optimal abatements from polluting firms and standing ready to make participation to such mechanism compulsory. Starting from an initial allocation of rights to emissions among firms, the market mechanism certainly reaches an efficient allocation. Those firms ready to reduce emissions below their BAU levels sell some rights while those eager to expand emissions buy those rights. This textbook solution, although attractive as it looks, is unfortunately of little help and sometimes even unfeasible in the contexts of international negotiations where voluntary participation must be induced and private information is a concern. Of course, the price mechanism may align the private incentives to reduce pollution with their socially optimal value. Free riding in effort provision might be solved. However, private information makes it also difficult and, let us insist, sometimes impossible, to find an initial allocation of rights that cannot be contingent on countries' private information and that would end up being satisfactory for all of them. Indeed, consider the situation where all countries are given the same initial rights. Then, countries which are efficient at reducing pollution will benefit from this system and may sell some rights while those countries with higher costs of depollution might start with too few rights to find this market solution attractive. They may just prefer to opt out the mechanism and free-ride on the joining ones.

To solve this important trade-off between *providing incentives to participants* and *providing incentives to participate*, we propose an alternative simple mechanism. The key additional ingredient that characterizes this mechanism is its ability to "screen" countries so that information on their preferences gets revealed through the negotiation process. On top of inducing efforts towards depollution and generating participation, this mechanism induces information revelation (even if it is in a rather coarse manner); a third objective that the market mechanism does not necessarily achieve. In the tradition of the mechanism design literature, this proposal consists in a menu of two options

[3] See Sandler (1992) for some perspectives on the specificities of such transnational public goods and Laffont and Martimort (2005) for a formal modeling.

among which countries are allowed to choose. The first "flat" option sees countries keeping their emissions at their BAU level but also paying a fixed amount that goes into a *green fund*. That contribution makes those countries indifferent between joining in or breaking the whole agreement to reach payoffs at their BAU level. If a country chooses instead the second 'incentives" option, it must also contribute upfront to the fund, but of course, and screening purposes, at a different level. On top, this country benefits also from a subsidy for each unit of effort. When the market mechanism fails, the menu mechanism optimally balances the dual objectives of *providing incentives to participants* and *providing incentives to participate*. Indeed, conditionally on the participation of each country, it would be optimal to propose a large subsidy per unit to induce the right actions - and solve the free-riding problem at the intensive margin. However, this would require to ask participants to pay large up-front contributions which deter participation of some less efficient countries. This would exacerbate the free-riding problem at the extensive margin. As a consequence of this trade-off, the optimal subsidy is reduced below the "market price." The optimal menu mechanism entails inefficient subsidies but facilitates participation.

A quick overview of the literature The literature on environmental agreements is by now significantly broad. It has spanned a whole array of topics with some applied papers discussing the properties of specific mechanisms which have some appeal from an implementation viewpoint and more theoretical contributions using game theory concepts to predict the outcomes of international negotiations in more abstract environments. Given space constraints, this review will just address a few contributions that are particularly meaningful to situate our own work.

As far as practical concerns are up, Bradford (2005) provides a very interesting proposal to organize climate agreements, the so called *"global public good purchase"* (GPGP) that can be viewed as an alternative to more traditional cap and trade systems. This mechanism roughly consists in building a fund out of countries' voluntary contributions and using this fund to reward efforts above the BAU trajectory. The interesting aspect of this proposal is that once the system is set up, all countries are rewarded though this fund for any abatement they do above the BAU level and so they all necessarily gain from the mechanism once in place. The negative side of this GPGP approach is that it does not say much on how the cost of the mechanism is shared among countries. Therefore, it leaves pending the free-riding problem on participation.

Inducing worldwide participation has precisely been the core of the theoretical debate. A traditional approach to climate agreements (see Hardin and Baden, 1977) would view climate as a standard public good where, as a result of a prisoner's dilemma, free riding cannot be avoided. A way out this negative result has been to suggest that repeated relationships may enlarge the set of equilibrium outcomes (see Cheikbossian and Sand-Zantman, 2011). This has led to the characterization of situations where all countries decide to join the agreement by fear of future retaliation (see Barret, 1994, for an in-depth discussion).

Taking into account the sovereign right for each country to enter (or not) into a binding agreement and studying how their behavior depends on conjectures on others' retaliation if they don't, Chander and Tulkens (1995, 1997) develop an alternative approach showing that efficiency may be compatible with a worldwide coalition. Their contribution shows that efficiency can be reached in a context of full information and under well specified conjectures.[4] Another alley for studying cooperation in the management of global public goods assumes that countries choose their abatement policies after an initial stage -part of a metagame- where each country decides whether to play cooperatively or not. Following this idea, Carraro and Siniscalco (1993) study the profitability and stability of partial international agreements in a framework where identical countries generate transboundaries emissions. They show that the commitment to use a cooperative strategy is crucial to broaden participation.[5]

There are surprisingly very few papers considering asymmetric information on the costs and benefits of mitigating climate change and almost no work addressing climate agreements taking a mechanism design perspective. Helm and Wirl (2011) is one of the very few exceptions. The authors consider a two-country setting where bargaining power is asymmetrically distributed and the uninformed country designs a mechanism controlling collective emissions. They show that the uninformed party must jointly use subsidies and his own emissions to incentivize the informed party and ensure its participation. To address informational issues in the multi-agent contexts of environmental problems without assuming *a priori* a particular institution, Baliga and Maskin (2003) have advocated the use of the mechanisms design tools. Although their contribution stresses the role of free riding in effort provision, it remains silent on free-riding problems in participation because they import the mechanism design methodology[6] developed for standard public good problems without taking care of the specific participation constraints that sovereignty imposes. This is the road we follow in a companion paper (Martimort and Sand-Zantman, 2012). That paper characterizes optimal mechanisms without making any restriction on the complexity of menus. Doing so, we develop a more complete analysis of the trade-off between free ridings on effort provision and on participation. We also compare the performances of this optimal mechanism with a simple menu of options where the incentive options in-

[4]In particular, Chander and Tulkens obtain their efficiency result by imposing unanimous agreement to join. We also consider this conjecture in the present paper, but in a setting of asymmetric information.

[5]Carraro (2005) discusses the influence of institutional rules (minimal participation rules, negotiation linkage, regional versus global treaties) on the outcome of environmental negotiations and studies the incentives to adopt those rules.

[6]See the seminal contributions by Laffont and Maskin (1982) and Mailath and Postlewaite (1990).

duces first-best incentives. This comparison stresses the small welfare costs associated with simplicity. The simple menu mechanism that we propose in the present paper improves on the menu mechanism proposed in Martimort and Sand-Zantman (2012) because we allow here to play on the slope of the incentive option. This extra degree of freedom is important in view of trading on incentives and participation in the most efficient way. The present paper differs also in stressing the performances of market mechanisms under informational requirements.

Organization of the paper. Section 2 presents the model as well as two useful benchmarks. Section 3 assesses the performances of a simple market mechanism and shows its limits when participation is voluntary. Section 4 proposes an alternative menu mechanism that reaches an optimal trade-off between the *provision of incentives to participants* and the *provision of incentives to participate*. Section 5 concludes. Proofs are relegated to an Appendix.

2 The Model and Two Useful Benchmarks

Our objective is to study the feasibility of international agreements. To do so, we assume that the world is populated by a continuum of countries of unit mass. Those countries, if they want to mitigate their pollution emissions, can exert a specific effort. More precisely, we assume that the effort e_i exerted by country i yields both local benefits of size αe_i (where $\alpha \in [0, 1)$) and global benefits, worth $(1-\alpha)e_i$, which accrue worldwide. The parameter α represents the magnitude of the local consequences of efforts in curbing pollution.[7] An alternative and broader rationale is that implementing a restrictive policy e_i impacts negatively on local production and welfare so that the local value of increasing effort is less than the social one.

To model heterogeneity, a key feature of our analysis, we assume that countries differ by their marginal cost of exerting effort, namely $C(e_i, \theta_i) = \frac{e_i^2}{2\theta_i}$, where θ_i is an efficiency parameter. Those costs should be understood in a broad sense, including not only technological costs but are also meant for opportunity costs (in terms of foregone growth or in terms of internal political haggling).[8] We thus write country i's utility function over payment-effort pairs (t_i, e_i) as:

$$U_i(t_i, e_i, \mathcal{E}, \theta_i) = t_i + \alpha e_i + (1-\alpha)\mathcal{E} - \frac{e_i^2}{2\theta_i}$$

where $\mathcal{E}$ represents the "aggregate" effort taken worldwide.

[7]For instance, CO_2 is known as having a global impact whereas other greenhouse gazes like SO_2 or NO_x have also significant local impacts.

[8]With the latter interpretation, it is not clear whether developed or developing countries are the ones with the smallest abatement costs.

Country i has private information on its efficiency parameter θ_i while its effort in mitigating pollution is observable. As the θ_i's are not only technological but also represent opportunity costs, assuming private information is certainly reasonable. Similarly, the assumption that effort is observable, contrary to other public good issues, fits the current situation where new technologies (by means of satellite observations) allow a very precise control of the behavior of each pollution unit.

Efficiency parameters θ_i are independently drawn from the same cumulative distribution $F(\cdot)$ with support $\Theta = [\underline{\theta}, \overline{\theta}]$ (with $\underline{\theta} > 0$) and an everywhere positive and atomless density $f(\theta) = F'(\theta)$. Let denote by $E_\theta(\cdot)$ the expectation operator with respect to θ. For technical reasons, we also assume that the following monotonicity condition holds:

$$\frac{d}{d\theta}\left(\frac{1-F(\theta)}{\theta f(\theta)}\right) \leq 0 \quad \forall \theta \in \Theta. \tag{1}$$

Benchmarks. Two natural benchmarks are interesting to look at: the first best allocation and the BAU scenario where no agreement is signed. To look at the first case, observe that worldwide welfare can be expressed as:

$$\mathcal{W} = \int_i U_i di = \int_i \left(t_i + \alpha e_i + (1-\alpha)\mathcal{E} - \frac{e_i^2}{2\theta_i}\right) di.$$

Using the budget balance condition, $\int_i t_i di = 0$, leads to

$$\mathcal{W} = \int_i \left(e_i - \frac{e_i^2}{2\theta_i}\right) di \equiv \int_\Theta \left(e(\theta) - \frac{e(\theta)^2}{2\theta}\right) f(\theta) d\theta.$$

At the first best, worldwide welfare is maximized for the following effort profile

$$e^{FB}(\theta) = \theta \quad \forall \theta \in \Theta$$

and its level is thus simply

$$\mathcal{W}^* = \frac{E_{\tilde{\theta}}(\tilde{\theta})}{2}. \tag{2}$$

Let us turn now to the BAU scenario. Now, countries do not internalize the impact of their own effort on overall welfare, and efforts are too low. More precisely, the corresponding effort of a country with type θ is:

$$e_N(\theta) = \arg\max_e \alpha e - \frac{e^2}{2\theta} + (1-\alpha)E_{\tilde{\theta}}(e_N(\tilde{\theta})) = \alpha\theta \quad \forall \theta \in \Theta.$$

Individual payoffs are then given by:

$$U_N(\theta) = \frac{\alpha^2}{2}\theta + (1-\alpha)\alpha E_{\tilde{\theta}}(\tilde{\theta}).$$

Social welfare becomes

$$\mathcal{W}_N = \frac{E_{\tilde{\theta}}(\tilde{\theta})}{2}\alpha(2-\alpha) < \mathcal{W}^*. \tag{3}$$

It is easy to see here that the gap between the BAU effort and the first-best level is naturally decreasing in α, and the same is true when looking at welfare levels. Indeed, the more local are the consequences of environmental policy, the smaller is the free-riding problem in exerting effort. Nevertheless, we will show next that the inefficiency of market mechanisms arises only when α is not too small.[9]

3 An Impossibility Result and the Inefficiency of the Market Mechanism

In this Section, we assess the performances of a simple market mechanism that allows countries to trade their abatement duties, or equivalently their rights for emissions. In the asymmetric information context that we consider, those performances must be assessed not only in terms of their incentive properties but also with an eye on whether this mechanism generates enough participation.

In this set-up, an initial allocation of those rights is chosen at the outset and each country is first endowed with some initial duties to fulfill. For simplicity and although the results below would hold with more generality,[10] we assume that duties are distributed uniformly. Let denote by E_0 this common value applying to all countries worldwide. Starting from that initial allocation, countries can trade those rights on a worldwide market. We denote by p the market price.

A country with type θ wants to trade rights so as to maximize the following expression:

$$U_0(\theta, p, E_0) = \max_e \alpha e + p(e - E_0) + (1-\alpha)\mathcal{E}_0 - \frac{e^2}{2\theta}, \tag{4}$$

where $\mathcal{E}_0 = E_{\tilde{\theta}}(e(\tilde{\theta}))$ denotes the overall level of abatements worldwide and $e(\tilde{\theta})$ is the equilibrium level of effort exerted by a country with type θ under this market scenario. The maximization of the above objective immediately yields the following expression of the "gross demand function" for rights:

$$e(\theta, p) = (\alpha + p)\theta. \tag{5}$$

[9] At the extreme, $\alpha = 1$ would nevertheless correspond to a degenerate case were BAU is efficient.

[10] The reader accustomed with the mechanism design literature already knows from the work of Cramton, Gibbons and Klemperer (1987) that efficiency can be achieved under asymmetric information in many contexts provided that the status quo payoff stipulates rights which are not too asymmetric. Although this latter paper was developed in the context of a partnership to divide a tradable good, its insights are quite general.

The market clearing condition finally gives the expression of the prevailing price p_0:

$$E_{\tilde{\theta}}(e(\tilde{\theta}, p_0)) = E_0 \Leftrightarrow p_0 = \frac{E_0}{E_{\tilde{\theta}}(\tilde{\theta})} - \alpha. \tag{6}$$

Through this market mechanism, a country with type θ ends up exerting an equilibrium effort worth:

$$e(\theta, p_0) = \frac{E_0}{E_{\tilde{\theta}}(\tilde{\theta})}\theta. \tag{7}$$

The market mechanism is thus an efficient way of inducing efforts if the uniform level of duties that applies worldwide is efficient "on average":

$$E_0 = E_{\tilde{\theta}}(\tilde{\theta}). \tag{8}$$

The prevailing market price is thus such that each extra unit of abatements beyond the initial allocation is paid at its social value:

$$p_0 = 1 - \alpha. \tag{9}$$

So doing, each country internalizes the impact of its own choice of effort on aggregate welfare and has the right marginal incentives to exert effort:

$$e(\theta, p_0) \equiv e^{FB}(\theta) \quad \forall \theta \in \Theta.$$

The remaining question is whether this market mechanism induces participation of all sovereign countries. Of course, checking that participation condition requires to compute the fall-back payoffs of countries that decide not to join the agreement. These payoffs in turn depend on the various conjectures that this country will entertain on the behavior of the ratifying ones. This issue is in fact a difficult one that we study at more length in Martimort and Sand-Zantman (2012). For the purpose of this paper, we follow Chander and Tulkens (1995, 1997) and assume that the approval of all countries is necessary to implement the market solution. Otherwise, the BAU scenario prevails worldwide.

In this context, inducing participation requires thus to satisfy:

$$U_0(\theta, p_0, E_0) \geq U_N(\theta) \quad \forall \theta \in \Theta. \tag{10}$$

Elaborating on this condition yields the following result:

Proposition 1 *The market mechanism cannot be efficient and induce voluntary participation worldwide when:*

$$\alpha > \alpha_1 = \frac{\underline{\theta}}{2E_{\tilde{\theta}}(\tilde{\theta}) - \underline{\theta}} \in (0, 1). \tag{11}$$

Condition (11) certainly holds when the parameter α is close enough to one (the case of a weak externality) or when heterogeneity on the productivity type θ is large enough so that $E_{\tilde{\theta}}(\tilde{\theta})$ is sufficiently above $\underline{\theta}$. In the case of global warming, this is of course this second interpretation that we will favor.

To understand this result, one must figure out the impact that choosing a uniform allocation of rights $E_0 = E_{\tilde{\theta}}(\tilde{\theta})$ has both on incentives to exert effort and on incentives to participate. From the market-clearing condition (6), this choice certainly induces the "right" market price, i.e., an implicit subsidy per unit of abatement which is Pigovian. All countries have the "right" incentives to exert effort. In the vocabulary coined in Martimort and Sand-Zantman (2012), the market mechanism induces *no free-riding at the intensive margin.*

Of course, the drawback of such a choice is that the market mechanism forces the less efficient countries (especially, those with types θ less than $E_{\tilde{\theta}}(\tilde{\theta})$) to start buying rights to abate less than their required duties. This is very costly and make their fall-back option much more attractive. *Free-riding at the extensive margin* becomes now the main concern and the market mechanism is ill equipped to address this issue.

A possible way out this trade-off could simply be to impose lower duties on inefficient countries at the outset. This would certainly relax free riding at the extensive margin. However, asymmetric information makes it impossible to make such type-dependent allocation of rights.

Of course, when trading off those costs and benefits of an initial allocation of rights which is "on average" efficient, one has to keep in mind that both the incentives and the participation effects depend in a non-trivial way on the size of the externality α. When α is small, positive externalities are significant and the fall-back option yields low payoffs. Participation constraints are *a priori* relaxed by this first effect that bites on the right-hand side of (10). On the other hand, and from an incentives viewpoint, countries do not care much about the local impact of their effort. To induce effort, the market price must be very large which means that inefficient types must pay a lot to abate less than their duties. This second-effect reduces the left-hand side of (10) which makes participation much harder. Condition (11) shows that α must be low enough to ensure that the first effect dominates.

Alternative implementation. When the market mechanism is efficient and induces worldwide participation, an alternative implementation of the final outcome would be for all countries to agree on a per unit subsidy for abatements cum a fixed contribution to a fund. Indeed, the countries' payoffs and their behavior are the same than in the market mechanism when facing the Pigovian subsidy $p_0 = 1 - \alpha$ and the fixed contribution T_0 such that:

$$T_0 = p_0 E_0 = (1 - \alpha) E_{\tilde{\theta}}(\tilde{\theta}). \tag{12}$$

Observe that this scheme is obviously budget balanced.

In the remainder of the paper, we will favor this more "centralized" approach in circumstances where efficiency is no longer feasible. We will bear a particular attention on how such centralized mechanisms must be modified in a second-best environment.

Remark 1 *Proposition 1 certainly casts also some doubts on the efficiency of a market for carbon if initial allocations are allocated by grand-fathering (i.e., as functions of past realizations that may be only partially correlated with future opportunity costs and thus "uniform" in a certain sense). An alternative interpretation of our findings is thus that the proposal to establish such market can never find support from all countries, especially those with the highest opportunity costs of abatements. Advocates of such market solution should certainly re-assess its virtues in view of the fundamental impossibility result presented in Proposition 1.*

4 Ensuring Voluntary Participation and Incentives with a Simple Menu

The previous Section has left us with a pessimistic view of what can be achieved with a worldwide market unless gains from cooperation are sufficiently sizable to overcome informational constraints. The key difficulty with a simple market mechanism is that the least efficient countries may want to stay out of the agreement rather than being imposed duties which are certainly too costly for them. Second-best mechanisms must be constructed to address this participation problem.

To do so, we propose a simple menu with two options. Those options allow self-selection of the different countries according to their opportunity costs of effort. Roughly speaking each of those options responds to the specific problem raised by one kind of the free-riding problems. The first option consists of a fixed contribution $\bar{T}$ and subsidy p per unit of abatement. This incentive option is targeted to the countries which are the most able to exert effort. With that option, those efficient countries have incentives to expand effort beyond the BAU level and free riding at the intensive margin diminishes for those types. The second option only entails a fixed contribution $\underline{T}$. Inefficient countries opt for that scheme which is designed to induce their participation to the mechanism. Free riding at the extensive margin is fought with that option.

Countries are now sorted according to their efficiency parameter θ. The most inefficient ones, with types on an interval $[\underline{\theta}, \theta^*]$ (thus a fraction $F(\theta^*)$) opt for the fixed contribution. *A contrario*, the most efficient countries whose types belong to $[\theta^*, \overline{\theta}]$ choose the incentive option. With such menu, all countries ratify the mechanism. If any country refuses to do so, we assume as before that the whole agreement breaks down with the BAU's payoffs being again the fall-back options.

By adopting the "incentive option" $(\bar{T}, p)$, a country with type θ reaches a payoff worth (where we make explicit its dependence on the mechanism):

$$U_1(\theta, p, \bar{T}) = \max_e (\alpha + p)e - \frac{e^2}{2\theta} + (1-\alpha)\mathcal{E}_1 - \bar{T} \tag{13}$$

where $\mathcal{E}_1$ whose expression will be clarified below is the expected effort worldwide under this scenario. Of course, maximizing the above objective immediately yields an effort supply which is still given by (3).

When opting for the fixed contribution, a country with type θ still exert its BAU effort level, namely $e_N(\theta)$. Such country benefits however from the greater effort exerted by more efficient ones. This is reflected in its payoff which is now given by:

$$U_1(\theta, \underline{T}) = \max_e \alpha e - \frac{e^2}{2\theta} + (1-\alpha)\mathcal{E}_1 - \underline{T}. \tag{14}$$

Taking into account that only a mass $F(\theta^*)$ of countries still exert their BAU effort $e_N(\theta)$ while a mass $1 - F(\theta^*)$ expands effort beyond this BAU level, we immediately find that the worldwide effort is worth:

$$\mathcal{E}_1 = \int_{\underline{\theta}}^{\theta^*} e_N(\theta) f(\theta) d\theta + \int_{\theta^*}^{\bar{\theta}} e(\theta, p) f(\theta) d\theta > \mathcal{E}_N = \int_{\underline{\theta}}^{\bar{\theta}} e_N(\theta) f(\theta) d\theta. \tag{15}$$

Equipped with this specification, we can write down the conditions that the mechanism $(\underline{T}, \bar{T}, p)$ must satisfy.

- *Incentive compatibility.* The cut-off type θ^* must be indifferent between the two options proposed:
$$U_1(\theta^*, p, \bar{T}) = U_1(\theta^*, \underline{T}). \tag{16}$$

- *Budget balance.* The overall fixed contributions contribute to a fund that is used to give subsidies under the incentive option. This budget balance condition can be written as:
$$F(\theta^*)\underline{T} + (1 - F(\theta^*))\bar{T} = p \int_{\theta^*}^{\bar{\theta}} e(\theta, p) f(\theta) d\theta. \tag{17}$$

- *Participation.* Insuring participation by the least efficient countries is obtained by making their fixed contribution "pay for the positive externality" exerted by most efficient ones when expanding their effort beyond the BAU level. This immediately gives the following condition:

$$\underline{T} = (1-\alpha) \int_{\theta^*}^{\bar{\theta}} (e(\theta, p) - e_N(\theta)) f(\theta) d\theta. \tag{18}$$

The general analysis that we develop in Martimort and Sand-Zantman (2012) shows that those three properties apply to more general (nonlinear) mechanisms. Yet, those simple menus with two options already highlight some important economic properties.

Proposition 2 *Suppose that conditions (1) and (11) both hold. The optimal menu* $(\underline{T}_1, \bar{T}_1, p_1)$ *has the following properties.*

- *The subsidy in the incentive option is lower than the Pigovian level:*

$$p_1 < p_0 = 1 - \alpha. \tag{19}$$

- *Not all countries always opt for this incentive option. Indeed, we have:*

$$\theta^* \in (\underline{\theta}, \overline{\theta}) \tag{20}$$

when

$$1 \geq \alpha \geq \frac{E_{\tilde{\theta}}(\tilde{\theta})}{\underline{\theta}} \alpha_1. \tag{21}$$

In a second-best environment, i.e., under the condition of Proposition 2 that prevents efficiency and full participation if the market mechanism prevails, the trade-off between the two free-riding problems at the extensive and at intensive margin is solved by giving up a bit of efficiency. This implies that the most efficient countries will receive a reduced subsidy and won't fully internalize the impact of their effort choice on worldwide welfare. Reducing this subsidy facilitates participation to the agreement by the least efficient countries. Indeed, those countries ratify the mechanism because they are just indifferent between joining in or not. They are ready to pay the positive externality that the most efficient ones bring with their increased effort beyond the BAU level. With an incentive option that becomes less powerful, less subsidies have to be paid and fixed contributions to the fund diminish which helps to solve free riding at the extensive margin.

When the tension between those two free-riding problems is strong (i.e., when (21) which is a strengthened version of (11) holds), each option within the menu attracts a positive measure of countries. Instead, when the tension is less pronounced, all countries choose the incentive option. Still, the subsidy remains lower than its Pigovian level.

5 Concluding Remarks

In the search for policies that could efficiently mitigate global warming, much emphasis has recently been put on the use of market mechanisms. The source of such interest should certainly be found in the optimistic view expressed by practitioners and scholars that the successes of cap and trade systems that were experienced in several regions of the world could carry over at a worldwide level. This paper casts some doubt on such optimistic stance. Market mechanisms may fail in getting to an efficient allocation of effort worldwide when participants are quite heterogenous and their voluntary participation must be induced. Market mechanisms certainly improve the free-riding problem in effort provision but may fail to induce participation. For worldwide agreements, nations are sovereign parties at the negotiation table; they cannot be viewed as firms subject to compulsory regulations in a domestic context. Because of this striking difference, the possible failure of markets should receive more consideration in political debates. Economists and practitioners should take with a word of caution the success of instruments and regulations developed in national arenas when it comes to broaden their scope to an international setting.

This paper takes seriously this unavoidable trade-off between incentives and participation. Yet, we have demonstrated that a simple menu mechanism with two options can reach unanimous participation and provide "almost" right incentives. With the first option, countries keep their emissions at their BAU level but contribute to a fund. With the second one, countries not only contribute to this fund but also benefit from the subsidies linked to the level of abatements they choose.

The approach we developed above might be criticized for being excessively static. As such, our model is necessarily silent on the relationship between what should be the current allocation of rights in a market solution and the past behaviors of countries. Although tailoring the level of permits left to a country to its past emissions may be attractive, especially for screening purposes, it might also introduce the possibility of gaming and, beyond, open the door to much politicking in deciding upon reference points. Second, the absence of dynamics renders our simple model of little value to account for issues of ethics and horizontal fairness among countries with different degrees of development that decisions on such allocation of initial rights may encounter. These are important questions that would deserve further works.

References

Baliga, S. and E. Maskin, 2003, "Mechanism Design for the Environment," in K.-G. Mler and . R. Vincent eds., *Handbook of Environmental Economics*, Elsevier.

Barrett, S., 1994, "Self-Enforcing International Environmental Agreements," *Oxford Economic Papers*, 46: 878-894.

Bradford, D., 2008. "Improving on Kyoto: GReenhouse Gas Control as the

Purchase of a Global Public Good", in R. Guesnerie and H. Tulkens eds., *The Design of Climate Policy*, CESIfo Seminar Series, MIT Press.

Carraro, C., 2005, "Institution Design for Managing Global Commons," in G. Demange and M. Wooders eds., *Group Formation in Economics*, Cambridge University Press

Carraro, C. and D. Siniscalco, 1993, 'Strategies for International Protection of the Environment," *Journal of Public Economics*, 52: 309-328.

Chander, P. and H. Tulkens, 1995, "A Core-Theoretic Solution for the Design of Cooperative Agreements and Transfrontier Pollution", *International Tax and Public Finance*, 2: 279-294.

Chander, P. and H. Tulkens, 1997, "The Core of an Economy with Multilateral Environment Externalities", *International Journal of Game Theory*, 26: 379-401.

Chander, P. and H. Tulkens, 2008, "Cooperation, Stability, Self-Enforcement in Agreements", in R. Guesnerie and H. Tulkens eds., *The Design of Climate Policy*, MIT Press.

Cheikbossian, G. and W. Sand-Zantman, 2011, "Dynamic Cooperation in Local Public Goods Supply with Imperfect Monitoring," *Annals of Economics and Statistics*, 101-102: 327-340.

Cramton, P., R. Gibbons and P. Klemperer, 1987, "Dissolving a Partnership Efficiently," *Econometrica*, 55: 615-632.

Hardin, G. and J. Baden, 1977, *Managing the Commons*, New York: Freeman and Co.

Laffont, J.J. and D. Martimort, 2005, "The Design of Transnational Public Good Mechanisms for Developing Countries," *Journal of Public Economics*, 89: 159-196.

Laffont, J.J. and E. Maskin, 1982, "The Theory of Incentives: An Overview," in *Advances in Economic Theory*, ed. W. Hildenbrand. Cambridge University Press.

Mailath, G. and A. Postlewaite, 1990, "Asymmetric Information Bargaining Problems with Many Agents," *Review of Economic Studies*, 57: 351-368.

Martimort, D. and W. Sand-Zantman, 2012, "A Mechanism Design Approach to Climate Agreements", Working Paper IDEI and PSE.

Sandler, T., 1992, *Collective Action: Theory and Applications,* University of Michigan Press.

Stern, N., 2006, *The Economics of Climate Change: The Stern Review*, Cambridge University Press.

Appendix

Proof of Proposition 1. Condition (10) amounts to:

$$\frac{1-\alpha^2}{2}\theta \geq \alpha(1-\alpha)E_{\tilde{\theta}}(\tilde{\theta}) \quad \forall\theta.$$

Taking $\theta = \underline{\theta}$ and taking into account that $\alpha < 1$ yields (11). ■

Proof of Proposition 2. Altogether (16), (17) and (18) define three equations with $(\underline{T}, \bar{T}, p)$ and the cut-off θ^* as the unknowns. After some simplifications using (3), (16), (17) and (18) yield respectively:

$$\theta^* p(p+2\alpha) = 2(\bar{T} - \underline{T}), \tag{A1}$$

$$\underline{T}F(\theta^*) + \bar{T}(1-F(\theta^*)) = (\alpha+p)p\int_{\theta^*}^{\bar{\theta}} \theta f(\theta)d\theta, \tag{A2}$$

and

$$\underline{T} = (1-\alpha)p\int_{\theta^*}^{\bar{\theta}} \theta f(\theta)d\theta. \tag{A3}$$

Combining the above conditions leads to the following expression of the cut-off θ^* as a function of p only:

$$\frac{\theta^*(p)(1-F(\theta^*(p)))}{\int_{\theta^*(p)}^{\bar{\theta}} \theta f(\theta)d\theta} = \frac{\theta^*(p)}{E_{\tilde{\theta}}(\tilde{\theta}|\tilde{\theta} \geq \theta^*(p))} = 2\left(1 - \frac{1}{p+2\alpha}\right). \tag{A4}$$

Let us now study the properties of $\theta^*(p)$. Before then, we prove the following Lemma.

Lemma A.1

$$\frac{d}{d\theta}\left(\frac{\theta}{E_{\tilde{\theta}}(\tilde{\theta}|\tilde{\theta} \geq \theta)}\right) > 0. \tag{A5}$$

Proof. Differentiating, we get:

$$\frac{d}{d\theta}\left(\frac{\theta}{E_{\tilde{\theta}}(\tilde{\theta}|\tilde{\theta} \geq \theta)}\right) = \frac{(1-F(\theta)-\theta f(\theta))\int_{\theta}^{\bar{\theta}} xf(x)dx + \theta^2(1-F(\theta))f(\theta)}{\left(\int_{\theta}^{\bar{\theta}} xf(x)dx\right)^2}.$$

Observe that the numerator on the right-hand side has the same sign as:

$$\Psi(\theta) = \left(\frac{1-F(\theta)}{\theta f(\theta)} - 1\right)\int_{\theta}^{\overline{\theta}} x f(x)dx + \theta(1-F(\theta)).$$

We have also $\Psi(\overline{\theta}) = 0$ and, from Assumption 1,

$$\dot{\Psi}(\theta) = \frac{d}{d\theta}\left(\frac{1-F(\theta)}{\theta f(\theta)}\right)\int_{\theta}^{\overline{\theta}} x f(x)dx < 0.$$

Therefore, we get $\Psi(\theta) \geq 0$ for all θ (withe quality only at $\overline{\theta}$). Hence, (A5) hold. ■

From Lemma A.1, we get:

$$\dot{\theta}^*(p) = \frac{2}{(p+2\alpha)^2 \frac{d}{d\theta}\left(\frac{\theta}{E_{\tilde{\theta}}(\tilde{\theta}|\tilde{\theta}\geq\theta)}\right)|_{\theta=\theta^*(p)}} > 0.$$

Hence, (A4) uniquely defines $\theta^*(p)$ which is interior when p is in the range defined by the next two inequalities:

$$\frac{\underline{\theta}}{E_{\tilde{\theta}}(\tilde{\theta})} \leq 2\left(1 - \frac{1}{p+2\alpha}\right) \leq 1$$

$$\Leftrightarrow 2\left(\frac{E_{\tilde{\theta}}(\tilde{\theta})}{\underline{\theta}}\alpha_1 - 1\right) = \underline{p} \leq p \leq \overline{p} = 2(1-\alpha). \tag{A6}$$

Let us now define worldwide welfare as:

$$\mathcal{W}(p) = \int_{\underline{\theta}}^{\theta^*(p)} \left(e_N(\theta) - \frac{e_N^2(\theta)}{2\theta}\right) f(\theta)d\theta + \int_{\theta^*(p)}^{\overline{\theta}} \left(e(\tilde{\theta},p) - \frac{e(\tilde{\theta},p)^2}{2\tilde{\theta}}\right) f(\theta)d\theta,$$

or, using (3), as

$$\mathcal{W}(p) = \left(\alpha - \frac{\alpha^2}{2}\right)\int_{\underline{\theta}}^{\theta^*(p)} \theta f(\theta)d\theta + \left(\alpha + p - \frac{(\alpha+p)^2}{2}\right)\int_{\theta^*(p)}^{\bar{\theta}} \theta f(\theta)d\theta. \tag{A7}$$

Differentiating with respect to p yields:

$$\mathcal{W}'(p) = (1-\alpha-p)\int_{\theta^*(p)}^{\bar{\theta}} \theta f(\theta)d\theta - \dot{\theta}^*(p)\theta^*(p)f(\theta^*(p))p\left(1-\alpha-\frac{p}{2}\right).$$

We evaluate this derivative at $p_0 = 1-\alpha$. First, observe that

$$\frac{\underline{\theta}}{E_{\tilde{\theta}}(\tilde{\theta})} < \frac{\theta^*(p_0)}{E_{\tilde{\theta}}(\tilde{\theta}|\tilde{\theta} \geq \theta^*(p_0))} = \frac{2\alpha}{1+\alpha} < 1$$

where the left-hand side inequality follows from (11) and the right-hand side inequality from $\alpha < 1$. From Lemma A.1, we thus get $\underline{\theta} < \theta^*(p_0) < \bar{\theta}$. Therefore, we find:

$$\mathcal{W}'(p_0) = -\dot{\theta}^*(p_0)\theta^*(p_0)f(\theta^*(p_0))\frac{(1-\alpha)^2}{2} < 0. \tag{A8}$$

From this, (19) immediately follows since $\dot{\theta}^*(p) > 0$ for all p.

Tedious computations yield also:

$$\mathcal{W}'(\underline{p}) = E_{\tilde{\theta}}(\tilde{\theta})(\alpha - \alpha_1) - \frac{(2E_{\tilde{\theta}}(\tilde{\theta}) - \underline{\theta})}{\Psi(\underline{\theta})}\underline{\theta}(E_{\tilde{\theta}}(\tilde{\theta}) - \underline{\theta})\left(\frac{E_{\tilde{\theta}}(\tilde{\theta})}{\underline{\theta})}\alpha_1 - \alpha\right)$$

and

$$\mathcal{W}'(\underline{p}) > 0 \tag{A9}$$

when (21) holds.

From (A8) and (A9), it follows that the optimal price p_1 satisfies:

$$\underline{p} < p_1 < p_0.$$

Finally, (20) follows because $\dot{\theta}^*(p) > 0$ for all p. ■