



Regulation via the Polluter-pays Principle

Stefan Ambec, Lars Ehlers

► To cite this version:

Stefan Ambec, Lars Ehlers. Regulation via the Polluter-pays Principle. The Economic Journal, 2016, 126 (593), pp.884-906. 10.1111/ecoj.12184 . hal-02637549

HAL Id: hal-02637549

<https://hal.inrae.fr/hal-02637549>

Submitted on 17 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Regulation via the Polluter-Pays Principle*

Stefan Ambec[†]

Lars Ehlers[‡]

May 22, 2014

Abstract

We consider the problem of regulating an economy with environmental pollution. We examine the distributional impact of the polluter-pays principle which requires that any agent compensates all other agents for the damages caused by his or her (pollution) emissions. With constant marginal damages we show that regulation via the polluter-pays principle leads to the unique welfare distribution that induces non-negative individual welfare change and renders each agent responsible for his or her pollution impact. We extend both the polluter-pays principle and this result to increasing marginal damages due to pollution. We also compare the polluter-pays principle with the Vickrey-Clark-Groves scheme.

JEL classification: C7, D02, D30, D6.

Keywords: Regulation, Polluter-Pays Principle, Fairness, Pollution, Externalities.

*We are especially grateful to anonymous referees and the Editor Martin Cripps for their helpful comments and suggestions. This research started while the second author was visiting Toulouse School of Economics (INRA-LERNA). It received financial support from INRA (France), the ANR (France) through the project ANR-08-JCJC-0111-01 on “Fair Environmental Policies”, the SSHRC (Canada) and the FQRSC (Québec).

[†]Toulouse School of Economics (INRA-LERNA-IDEI), France, and University of Gothenburg, Sweden; e-mail: stefan.ambec@toulouse.inra.fr.

[‡]Département de Sciences Économiques and CIREQ, Université de Montréal, Canada; e-mail: lars.ehlers@umontreal.ca (Corresponding author).

1 Introduction

From water management to air pollution, managing environmental problems efficiently requires well-designed public policies or coordination among stakeholders (Ostrom, 1990). Environmental policies are launched to mitigate the failure of market economy due to the presence of negative externalities. Yet public intervention has an impact not only on the social welfare of the economy as a whole but also on the distribution of welfare. Although the economics literature on the choice of environmental regulations tends to focus on efficiency, the way instruments affect individuals' welfare matters a lot in practice. It determines their success or failure in democratic societies as citizens might oppose regulations that hurt them or that are perceived unfair. Equity is a main determinant of policy options on environmental issues.

Fairness principles are central in the debate on climate change mitigation policies. Several have been invoked during international climate negotiations, leading to conflicting policy recommendations. Developed countries are mostly in favor of the sovereignty principle. Under the premise that all nations have equal rights to the atmosphere, it takes current greenhouse gas emissions levels as the status-quo. It implies that limitations on future emissions should be proportional to current ones. In contrast, developing countries support the responsibility principle. Brazil argued during the Kyoto negotiation that responsibility for compensating environmental damage should be related to the degree of responsibility for its causes. It implies that emission reductions should be proportional to historical contribution (Heyward, 2007).¹

We analyze the fairness properties of welfare distributions implemented (in Nash equilibrium) by regulation schemes in an economic environment with pollution. The model allows for a variety of negative externalities including unilateral or multilateral ones, heterogeneous impacts due to distance or mitigation. It formalizes many complex environmental issues such as water quality management in a river or the reduction of sulfur dioxide or greenhouse gas emissions in an international setting.² We assess the performance of welfare distributions regarding two fairness criteria.

Our first criterion is a lower bound on the change of individual welfare: it should be non-

¹Lange, Löschel, Vogt and Ziegler (2010) provide evidence that the fairness principles invoked by negotiator are consistent with the interest of the country they represent.

²To that respect, it is as rich as the seminal model of Montgomery (1972).

negative (whereby the welfare change is measured against some alternative in which no harm is incurred from pollution). It is a minimal acceptability requirement since an agent with negative welfare change does not benefit from the welfare-enhancing economic activities exhibiting pollution. As long as pollution improves social welfare, nobody should lose compared to the situation of no pollution.

Our second criterion relies on the concept of responsibility in the theory of justice (Fleurbaey, 2008). It makes a polluter responsible for his pollution impact on society. More precisely, if a polluter modifies the environmental impact of his own emissions in the economy, he should get the full return or loss due to this change. For instance, a firm which filters its own emissions to reduce their sulfur content should get the full benefit for the economy of its cleaning investment. A farmer who uses more pesticides and fertilizers leading to dirtier waste water should pay the social cost associated to this pesticide and fertilizer increase.³

The polluter-pays welfare distributions are the welfare distributions implemented by a regulation inspired by a literal interpretation of the polluter-pays (PP) principle. The PP principle states that *the costs of pollution should be borne by the entity which profits from the process that causes pollution*.⁴ It is commonly invoked in practice during policy discussions on environmental issues. The PP-scheme requires that any agent compensates all agents who suffer from his pollution emissions for the damage he causes. When marginal damages due to pollution are constant, the PP welfare distributions are the only ones that satisfy non-negativity and responsibility of pollution impact.

The case of increasing marginal damages raises additional issues, and to address these we need a third criterion, which is an upper bound on individual welfare. It is the maximal welfare an agent (or a single polluter) can obtain subject to compensating all other agents for the damage due to pollution and all other agents being inactive (or emitting no pollution). When

³In practice, it is sometimes difficult to determine the entity which is responsible for pollution. It is a central issue in Coase (1960). For instance, Coase mentions the case of straying cattle which destroy crops on neighboring land. Legal liability for the destroyed crops is unclear. He states that "... it is true that there would be no crop damage without the cattle. It is equally true that there would be no crop damage without the crop." It is implicitly assumed here that the polluter can be clearly identified.

⁴See http://en.wikipedia.org/wiki/Polluter_pays_principle and Principle 16 in the Rio Declaration on Environment and Development.

environmental damages are increasing with pollution concentration, polluters exert negative externalities among them: the fact that some agents are polluting reduces the ability of others to pollute. Since an agent is not responsible for the pollution emitted by others, he would claim the social welfare of his activity if he was the only one to pollute. The single-polluter welfare cannot be assigned to all agents. Thus, by solidarity, single-polluter upper bounds requires that every agent takes up a share of the negative externalities among polluters by enjoying not more than his single-polluter welfare.

When marginal damages are increasing with pollution concentration, the PP principle is not straightforwardly defined because the cost generated by an emitter depends on the other polluters' emissions. We extend the PP principle to this framework by making the polluter pay for the incremental impact of his emissions on the total welfare of the other agents when he does pollute and when he does not pollute (at the efficient levels of pollution). As a result, each polluter receives in the PP welfare distribution the difference of society's welfare when he does pollute and when he does not pollute. Under increasing marginal damages, the PP welfare distributions are the only ones that satisfy the above three criteria: non-negativity, responsibility for pollution impact and single-polluter upper bounds.

To the best of our knowledge, our paper is the first contribution to characterize the welfare distributions induced by the polluter-pays principle via fairness properties. The PP-scheme is by construction feasible because it induces no budget deficit. It is also efficient in the sense that it uniquely implements the allocation of pollution emissions that maximizes total welfare in Nash equilibrium. These features are shared by the PP-scheme and other schemes proposed in the literature on pollution environments. For instance, Duggan and Roberts (2002) introduced a scheme in which each agent chooses his emission and reports the emission of his neighbor. In Montero (2008) each agent reports his inverse demand for any level of emissions. The focus of both papers is the implementation of the efficient allocation under asymmetric information whereas we are interested in the distributional impacts of a scheme that implements the efficient allocation under perfect information. The Vickrey-Clark-Groves (VCG)-scheme, which makes each agent pay for his marginal impact to society, has a similar flavor to our PP-scheme. However, as we explain in Section 6, the two schemes differ. In particular, the VCG-scheme applied to the pollution problem does not satisfy non-negativity

while the PP-scheme does. Furthermore, we show that the PP-scheme is incentive compatible when damages are known (and benefits are private information) whereas the VCG-scheme is incentive compatible when both damages and benefits are private information.

We proceed as follows. Section 2 introduces a simple model of pollution with constant marginal damages. It also provides several real-world examples which fit our framework. Section 3 describes regulation schemes and their induced distribution rules in equilibrium. It also discusses several regulations used in real life. Section 4 introduces the polluter-pays regulation and characterizes its induced distribution rule in terms of non-negativity and responsibility for pollution impact. Section 5 generalizes our model to differentiate pollution and damages in order to allow increasing marginal damages. We generalize the PP principle and extend our main results to this framework. Section 6 compares the PP-scheme with VCG-scheme and discusses preference revelation for the PP-scheme. Section 7 concludes.

2 A Model of Pollution

Consider a set $N = \{1, \dots, n\}$ of agents (countries, cities, farmers, firms, consumers,...). Each agent $i \in N$ is polluting or is polluted or both. Agent i enjoys a benefit $b_i(e_i)$ from production and/or consumption where $e_i \geq 0$ denotes i 's level of economic activity hereafter called "emissions". The benefit function b_i is assumed to be both strictly concave and strictly increasing from 0 to a maximum \hat{e}_i with $b'_i(\hat{e}_i) = 0$ for every $i \in N$, and twice continuously differentiable: for all $i \in N$ and for all $0 \leq e_i < \hat{e}_i$, both $b'_i(e_i) > 0$ and $b''_i(e_i) < 0$. We normalize $b_i(0) = 0$ and assume that the marginal benefit at $e_i = 0$ is high enough (say infinite) so it is optimal for all agents to produce and/or to consume.

Pollution from agent i causes marginal damage $a_{ij} \geq 0$ to agent j . The parameter a_{ij} measures the magnitude of the pollution impact of i 's emission on j . For the moment we consider constant marginal damages. Later we extend our results to environments with convex damages and thus, increasing marginal damages from emissions. A (negative) externality or pollution problem (N, b, a) is defined by a set of agents N , a profile of benefit functions $b = (b_i)_{i \in N}$, and a matrix of externality/pollution marginal impacts $a = [a_{ij}]_{ij \in N \times N}$. When there is no confusion, we write for short a instead of (N, b, a) . Throughout we assume that

$a_{ii} > 0$ for any $i \in N$ with $\sum_{j \in N \setminus \{i\}} a_{ij} > 0$, i.e. if i is polluting other agents, then his pollution causes some damage at his location. Let

$$\mathcal{A} = \{a = [a_{ij}]_{ij \in N \times N} : a_{ij} \geq 0 \text{ for all } ij \in N \times N \text{ \& } a_{ii} > 0 \text{ for all } i \in N \text{ with } \sum_{j \in N \setminus \{i\}} a_{ij} > 0\}$$

denote the set of all problems.

Given $a \in \mathcal{A}$, let $Ri(a) = \{j \in N | a_{ij} > 0\}$ denote the receptors of i 's pollution: the set of agents which are polluted by i . Let $R^0i(a) = \{j \in N \setminus \{i\} | a_{ij} > 0\}$ denote the receptors of i 's pollution excluding i . Let $Si(a) = \{j \in N | a_{ji} > 0\}$ denote the set of agents who pollute agent i . Let $S^0i(a) = \{j \in N \setminus \{i\} | a_{ji} > 0\}$ denote the set of agents who pollute i excluding i . When a is fixed, we write for short Ri , R^0i , Si and S^0i instead of $Ri(a)$, $R^0i(a)$, $Si(a)$ and $S^0i(a)$.

Let $a \in \mathcal{A}$ be a problem. The environmental damage suffered by i in the emission vector $e = (e_i)_{i \in N}$ is therefore

$$d_i = \sum_{j \in Si} a_{ji} e_j.$$

The welfare of agent i with emissions $e = (e_i)_{i \in N}$ is

$$b_i(e_i) - d_i = b_i(e_i) - \sum_{j \in Si} a_{ji} e_j. \quad (1)$$

The first term in (1) is i 's benefit from his own emissions whereas the second term is i 's welfare loss due to pollution.

An *efficient* emission plan $e^* = (e_i^*)_{i \in N}$ maximizes total welfare $\sum_{i \in N} [b_i(e_i) - d_i] = \sum_{i \in N} b_i(e_i) - \sum_{i \in N} \sum_{j \in Si} a_{ji} e_i$. It satisfies the following first-order conditions for every $i \in N$:

$$b'_i(e_i^*) = \sum_{j \in Ri} a_{ij}. \quad (2)$$

Note that our assumptions on the benefit function b_i guarantee that e_i^* is unique because $b'_i(\hat{e}_i) = 0$ and b_i is strictly concave and strictly increasing between 0 and \hat{e}_i . The marginal benefit of pollution emitted by i should be equal to its marginal damage for society. Let

$$W(a) = \sum_{i \in N} b_i(e_i^*) - \sum_{i \in N} \sum_{j \in Si} a_{ji} e_i^*$$

denote the economy's welfare from the efficient emission plan e^* in the problem (N, b, a) .⁵ A welfare distribution for the problem (N, b, a) is a vector $z = (z_i)_{i \in N} \in \mathbb{R}^N$ such that $\sum_{i \in N} z_i \leq$

⁵Although the pollution problem is defined not only by the matrix a but also by the vector b and the set N , we often omit b and N in the notation because we consider only variations of a in the properties.

$W(a)$. A distribution rule ϕ associates with any problem (N, b, a) a welfare distribution $\phi(a)$ for a , i.e. $\phi: \mathcal{A} \rightarrow \mathbb{R}^N$ such that $\sum_{i \in N} \phi_i(a) \leq W(a)$ for all $a \in \mathcal{A}$. Note that a distribution rule identifies for each problem a welfare distribution which the society may wish to implement.

The externality problem (N, b, a) exhibits multilateral externalities if $Si = Ri$ for any $i \in N$. The problem (N, b, a) exhibits unilateral externalities if $S^0i \cap R^0i = \emptyset$ for any $i \in N$. Let $V \subseteq N$ denote the set of agents who do not pollute other agents and only suffer from pollution due to other agents' activities. Formally, for any $i \in V$, $a_{ij} = 0$ for all $j \neq i$ and $a_{ji} > 0$ for at least one $j \neq i$, or equivalently $R^0i = \emptyset$ and $S^0i \neq \emptyset$; and without loss of generality, $\hat{e}_i = 0$.⁶ Similarly, let $P \subseteq N$ denote the set of agents who do not suffer from other agents' pollution: $a_{ij} > 0$ for some $j \neq i$ and $a_{ji} = 0$ for all $j \neq i$, that is $R^0i \neq \emptyset$ and $S^0i = \emptyset$. Note that any agent in $N \setminus V$ is polluting the society from his economic activities.

Below we describe briefly four real-life applications which can be easily accommodated with our model.

In the river pollution problem, the agents (countries, cities or factories) are located along a river. Each agent i emits e_i units of pollution which impact its followers downstream: one unit emitted at i causes marginal damage a_{ij} at j . Symmetrically, agent i suffers from pollution emitted upstream by agents and himself.⁷ In a single canal or one-tributary river, agents are ordered according to their position from downstream to upstream, say $N = \{1, \dots, n\}$, and for any $i \in N$, $Ri = \{1, 2, \dots, i\}$ and $Si = \{i, i + 1, \dots, n\}$. Note that this is a case of unilateral externalities.⁸

In the international greenhouse gas emissions game, the agents are countries. Greenhouse gases emitted into the atmosphere cause global warming that damages countries' economies. The magnitude of global warming depends on total emissions on the earth surface $\sum_{j \in N} e_j$. Suppose that total emissions cause a constant marginal damage of δ_i to country i . In this

⁶If $\hat{e}_i > 0$, then agent i 's activity does not have any impact on society and his activities can be disregarded.

⁷In the case of a river, "linearity is a good approximation up to the point at which the river becomes so overloaded with organic material that oxygen (needed for aerobic bacteriological decomposition) is depleted. At that point, [refereed as the river carrying capacity] the river's capacity to clean itself is greatly diminished." from Kolstad (2000) footnote 2 page 177.

⁸See Ambec and Sprumont (2002) and Ambec and Ehlers (2008) for a rigorous analysis of the river water sharing problem. Demange (2004) considers stability in hierarchies given by trees.

example, $Si = Ri = N$ and $a_{ii} = a_{ij} = \delta_i$ for all $i, j \in N$: all countries exert multilateral externalities on all other countries of the same magnitude.⁹

In the international acid rain game, agents are countries emitting sulfur dioxide (SO₂) by burning coal for power production. This causes acid rain which damages forests and ecosystems in neighboring countries. The parameter a_{ij} captures the marginal impact of country i 's SO₂ emissions to acid rain in country j .¹⁰

In the polluters versus victims problem, agents in V are individuals and those in P are firms, and each agent belongs either to V or to P . Firms emit pollution without incurring any damage: for any $i \in P$, $a_{ji} = 0$ for every $j \neq i$. In contrast, any $i \in V$ does not emit pollution but suffers from pollution: $\hat{e}_i = 0$ for every $i \in V$ and $a_{ji} > 0$ for at least one $j \in P$. In this case, a_{ji} can be interpreted as the marginal damage which each unit of firm j 's pollution causes to person i in term of health or environmental impact. The main difference with the previous examples is that emitters and victims are disjunct sets of agents. This is a case of unilateral externalities.¹¹

3 Regulation Schemes and Distribution Rules

The policy tools used for pollution problems can be characterized as regulation schemes. A regulation scheme $t : \mathbb{R}_+^N \rightarrow \mathbb{R}^N$ specifies for any emissions a vector of payments (or transfers) $t(e) = (t_i(e))_{i \in N}$. It assigns to agent i the transfer $t_i(e)$ for any emission plan $e = (e_i)_{i \in N}$.

⁹Seminal papers on international agreements for greenhouse emission reduction are Chandler and Tulkens (1992), Carraro and Siniscalco (1993) and Barrett (1994).

¹⁰Mäler and De Zeeuw (1998) provide estimations on those parameters for 1990 and 1991 in Europe. Mäler (1989, 1994) considers an acid rain game with such heterogeneous "transportation" parameters and constant marginal damages. This game has been extended by Mäler and De Zeeuw (1998) and Finus and Tjøtta (2003) to environments with convex marginal damages.

¹¹One can easily check that all our results remain valid if we restrict the set of all problems in the following way: let $H = (H_i)_{i \in N}$ be such that for all $i \in N$, (i) $H_i \subseteq N$ and (ii) $i \in H_i$ if $H_i \setminus \{i\} \neq \emptyset$. Now let

$$\mathcal{A}_H = \{a \in \mathcal{A} : Ri(a) = H_i \text{ for all } i \in N\}.$$

Now we can fix H to accommodate any of the applications above. For instance, for the river pollution problem, we set $H_i = \{1, 2, \dots, i\}$ for all $i \in N$.

Given the scheme t and the emission plan e , agent i 's welfare under the vector $t(e)$ is given by

$$b_i(e_i) - d_i + t_i(e) = b_i(e_i) - \sum_{j \in Si} a_{ji}e_j + t_i(e). \quad (3)$$

Of course, each agent i chooses his own emissions and for any problem a , the regulation scheme t induces an “emissions game” given by the game form where each agent's set of strategies consists of all emissions in \mathbb{R}_+ and the outcome function is t . Let $\mathcal{N}(t, a)$ denote the set of (pure) non-cooperative Nash equilibria in the emissions game under the scheme t and the problem a .

In the non-cooperative Nash equilibrium of the externality problem with scheme t , each player i maximizes (3) with respect to e_i given $e_{-i} = (e_j)_{j \in N \setminus \{i\}}$. Let $e^t \in \mathcal{N}(t, a)$ be a Nash equilibrium emission plan. Agent i 's equilibrium welfare under e^t is

$$z_i^t = b_i(e_i^t) - d_i^t + t_i(e^t),$$

where $d_i^t = \sum_{j \in Si} a_{ji}e_j^t$. The total welfare is

$$W^t(a) = \sum_{i \in N} z_i^t = \sum_{i \in N} [b_i(e_i^t) - d_i^t + t_i(e^t)] = \sum_{i \in N} b_i(e_i^t) - \sum_{i \in N} d_i^t + \sum_{i \in N} t_i(e^t),$$

where in the last expression the first term is the total benefit from emission, the second is the total damage and the third is the regulation scheme surplus (or deficit if negative).

Given a distribution rule ϕ and a scheme t , we say that t implements ϕ (in Nash equilibrium) if for all problems $a \in \mathcal{A}$ and all $e^t \in \mathcal{N}(t, a)$, we have

$$\phi_i(a) = z_i^t = b_i(e_i^t) - \sum_{j \in Si} a_{ji}e_j^t + t_i(e^t).$$

A particular regulation scheme is the *laissez-faire* scheme t^{lf} defined by $t_i^{lf}(e) = 0$ for all $i \in N$ and all $e \in \mathbb{R}_+^N$. The laissez-faire scheme represents situations without regulation or where society chooses not to intervene. It implements the emission plan $e^{lf} = (e_i^{lf})_{i \in N}$ satisfying the following first-order conditions,

$$b'_i(e_i^{lf}) = a_{ii},$$

for every $i \in N$. Thus, for each problem a , $\mathcal{N}(t^{lf}, a)$ is unique and implicitly given by the above equalities. In contrast to the efficient emission plan e^* , under laissez-faire each agent i

considers the impact of his emissions only on his own welfare. In particular, $e_i^{lf} = \hat{e}_i$ if $a_{ii} = 0$. As long as $a_{ij} > 0$ for some $j \neq i$, i.e. i 's emissions have an impact on another agent j , then $e_i^{lf} > e_i^*$ and therefore $d_j^{lf} > d_j^*$ for every $j \in Ri$.

Many regulation schemes are used in practice. For instance, consider a norm on pollution emissions scheme, denoted by \bar{t} . It defines upper bounds on emissions $\bar{e}_i \geq 0$ and penalties for exceeding these bounds. Formally, let $\bar{e} = (\bar{e}_i)_{i \in N}$ and for all $e \in \mathbb{R}_+^N$,

$$\bar{t}_i(e) = \begin{cases} 0 & \text{if } e_i \leq \bar{e}_i \\ -F_i(e_i - \bar{e}_i) & \text{if } e_i > \bar{e}_i \end{cases}$$

for every $i \in N$ where $F_i > 0$ is the fine in case of excess pollution (which can be infinite or lump-sum). In case of an uniform norm, $\bar{e}_i = \bar{e}_j$ and $F_i = F_j$ for all $i, j \in N$. If the fine is high enough to be persuasive and the norm is binding in the sense that $e_i^{lf} > \bar{e}_i$ for all $i \in N$, then the unique emission plan implemented in Nash equilibrium by \bar{t} are $e_i^{\bar{t}} = \bar{e}_i$ for all $i \in N$.

The emission fee scheme t^f specifies fees $f = (f_i)_{i \in N}$ on emissions and, therefore, charges the payment $t_i^f(e) = -f_i e_i$ from agent i . Here $f_i > 0$ is polluter i 's tax rate. The Pigouvian fee is $f_i = \sum_{j \in R^0 i} a_{ij}$ for every polluter $i \in N$. It implements the first-best emissions e^* in Nash equilibrium. Alternatively, the fee can be on ambient pollution rather than on emissions. An ambient pollution fee scheme t^F charges $F_j > 0$ per unit of emissions at each receptor j which makes agent i pay $t_i^F(e) = -\left(\sum_{j \in Ri} a_{ij} F_j\right) e_i$. The emission or ambient pollution fee scheme can be associated with a redistribution policy of the money collected, e.g. through lump-sum transfers or subsidies.

A further important regulation instrument that can be embedded in our model is cap-and-trade or tradable emission permits. Agents are endowed with some initial emissions allowances or permits $\bar{e} = (\bar{e}_i)_{i \in N}$ which can be traded in a market. They are not allowed to emit more than the amount of permits they own at the end of a pre-defined phase (trade occurs at the same time as pollution in the European EUTS market). Providing that the permit market is competitive (implying that agents are price takers), the tradable emission permit regulation is as if each agent i faces a transfer scheme $t_i^{tp}(e) = p(\bar{e}_i - e_i)$ where p is the equilibrium price of permits. This price is uniquely determined by the first-order conditions $b_j'(e_j^t) - a_{jj} = p$ for every $j \in N \setminus V$ and the market clearing condition $\sum_{j \in N} \bar{e}_j = \sum_{j \in N} e_j^t$. The initial allocation of permits impacts the level and distribution of welfare. Under grandfathering, each agent is

assigned a share of his or her laissez-faire emission $\bar{e}_i = \alpha e_i^{lf}$ with $0 < \alpha \leq 1$. A lower α means lower emissions in the economy. When permits are auctioned by the government, it is as if those who get the revenue from this auction are endowed with the permits. For instance, if the money is used exclusively to reduce or compensate the damage at agent h 's location, then it is as if agent h obtains all permits and trades them with polluters in a competitive market, i.e. $\bar{e}_h = \sum_{j \in N \setminus V} e_j^t$. Emission allowances can also be defined on receptors emissions, each agent i potentially owning \bar{e}_{ij} emission allowances at receptor j that can be exchange against other emission allowances for the same receptor j .

Given the abundance of different regulation schemes in reality, a society would like to distinguish between them according to desirable criteria. The following will be two very basic requirements any society would like any regulation to comply with.

Efficiency requires that the first-best outcome is implemented in Nash equilibrium.

Efficiency: For all problems a and all $e^t \in \mathcal{N}(t, a)$, we have $e^t = e^*$.

The second property requires that the payments of the scheme induce no budget deficit at Nash equilibrium.

(Budget) Feasibility: For all problems a and all $e^t \in \mathcal{N}(t, a)$, we have $\sum_{i \in N} t_i(e^t) \leq 0$.

A (budget) feasible regulation scheme t where $\mathcal{N}(t, a)$ is a singleton for any a , say $\mathcal{N}(t, a) = \{e^t\}$, induces a distribution rule ϕ^t of the total welfare. For any problem $a \in \mathcal{A}$, the distribution rule implemented by the feasible scheme t is given by

$$\phi_i^t(a) = b_i(e_i^t) - \sum_{j \in Si} a_{ji} e_j^t + t_i(e^t) \text{ for all } i \in N.$$

Any of the above regulation schemes is feasible and has a unique Nash equilibrium, and hence, induces a corresponding distribution rule. We now focus on a particular regulation scheme, the one inspired by the polluter-pays principle.

4 Welfare Properties of the Polluter-Pays Principle

In this section, we first describe the polluter-pays scheme and show two of its properties, namely feasibility and efficiency. Second, we examine the properties of the welfare distribution rules implemented by regulation schemes in Nash equilibrium, and in particular by the polluter-pays welfare distribution rule.

4.1 The Polluter-Pays Scheme

Many countries pay lip-service to the “polluter-pays” (PP) principle as a regulation scheme. It basically renders the polluter responsible for the damage it causes to the environment. It requires that *the costs of pollution should be borne by the entity which profits from the process that causes pollution*.

In order to satisfy the polluter-pays principle, the entity who pollutes should compensate those who suffer from this pollution for the damages it causes. If a victim is not fully compensated, then he or she pays part of the cost of someone else’s pollution. Hence, strictly speaking, the PP principle imposes not only that polluters pay for the damage caused to society, but also that victims are fully compensated for those damages. In our model, an arbitrary agent i who pollutes should compensate every agent $j \in R^0 i$ for the caused damage $a_{ij}e_i$. Agent i pays $a_{ij}e_i$ to every $j \in R^0 i$. Therefore, as a victim of pollution, agent i receives the compensation $a_{ji}e_j$ from each agent $j \in S^0 i$ who pollutes him. Summing up all these side-payments, the polluter-pays principle leads to the regulation scheme $t^{PP}(e)$ defined as follows for any agent $i \in N$:

$$t_i^{PP}(e) = \sum_{j \in S^0 i} a_{ji}e_j - \sum_{j \in R^0 i} a_{ij}e_i = d_i - a_{ii}e_i - \sum_{j \in R^0 i} a_{ij}e_i = d_i - \sum_{j \in Ri} a_{ij}e_i. \quad (4)$$

Agent i receives the net transfer from the cost of pollution he suffers minus the cost of pollution he causes to society. Since the polluter-pays principle involves side-payments among agents, the payments in the PP-scheme sum up to zero. It is therefore (budget) feasible. Agent i ’s welfare under the payments $t^{PP}(e)$ with emission plan e is:

$$b_i(e_i) - \sum_{j \in Ri} a_{ij}e_i \quad (5)$$

Since agent i pays for the marginal damage caused to others and is compensated from the marginal damage caused by others, his welfare under the PP-scheme in (5) is the social benefit from his economic activity. Therefore, agent i has incentive to emit the efficient level e_i^* for any given emissions emitted by other agents. Formally, maximizing (5) with respect to e_i leads to the first-order condition (2) which implies $e_i^t = e_i^*$ for every $i \in N$. This implies that the PP-scheme implements the efficient emission plan e^* in Nash equilibrium, i.e. $\mathcal{N}(t^{PP}, a) = \{e^*\}$. A particular feature of regulation through the PP-scheme with constant marginal damages is that, since any individual's payoffs depend only on the agent's own choice (no externality), the efficient emission plan is a dominant strategy equilibrium, which is an equilibrium concept which is less demanding in terms of cognitive skills than Nash equilibrium. Therefore, the efficient emission plan remains the unique Nash equilibrium when the parameters a are publicly known but the benefit functions are private information.

Remark 1 One can check that the efficient emission plan is robust to collusion, i.e. it remains the unique equilibrium in the PP-scheme if we allow coalitions to jointly change their emissions. Formally, given scheme t and $a \in \mathcal{A}$, $e' \in \mathbb{R}_+^N$ is a coalition-proof Nash equilibrium if there exists no $\emptyset \neq U \subseteq N$ and $e''_U = (e''_i)_{i \in U} \in \mathbb{R}_+^U$ such that for all $i \in U$,

$$b_i(e''_i) - \left(\sum_{j \in Si \cap U} a_{ji} e''_j + \sum_{j \in Si \setminus U} a_{ji} e'_j \right) + t_i(e''_U, e'_{N \setminus U}) > b_i(e'_i) - \sum_{j \in Si} a_{ji} e'_j + t_i(e').$$

Now for any $a \in \mathcal{A}$, in the PP-scheme e^* is the unique coalition-proof Nash equilibrium because for any non-empty coalition $U \subseteq N$, $(e_j^*)_{j \in U}$ solves $\max_{(e_j)_{j \in U} \geq 0} \sum_{i \in U} [b_i(e_i) - (\sum_{j \in Si \cap U} a_{ji} e_j + \sum_{j \in Si \setminus U} a_{ji} e_j^*) + t_i^{PP}(e_U, e_{N \setminus U}^*)]$.

We will denote by ϕ^{PP} the polluter-pays (PP) distribution rule associating with each problem a the polluter-pays welfare distribution $\phi^{PP}(a)$: for any $i \in N$, agent i 's equilibrium welfare is given by

$$\phi_i^{PP}(a) = b_i(e_i^*) - \sum_{j \in Ri} a_{ij} e_i^*. \quad (6)$$

The result below follows straightforwardly from our discussion.

Proposition 1 *The polluter-pays scheme is an efficient and feasible regulation scheme implementing the polluter-pays distribution rule.*

4.2 The Polluter-Pays Distribution Rule

The following are two desirable criteria a society would like to be satisfied by the welfare distributions implemented via a regulation scheme. The first criterion postulates that any agent should receive a non-negative payoff. This corresponds to the requirement that any agent's welfare change should be non-negative because we implicitly set any agent's welfare equal to zero in the absence of pollution or emission activities (which is interpreted as the status quo).

Non-Negativity: For all problems a and all $i \in N$, $\phi_i(a) \geq 0$.

Non-negativity simply requires that nobody should be worse off under pollution than without pollution. The second criterion renders the polluter responsible to any change of his pollution impact on the economy.

Responsibility for Pollution Impact (RPI): Consider any arbitrary agent $i \in N$. Suppose that agent i 's pollution impact is reduced from $a = (a_{ij})_{j \in N}$ to $a' = (a'_{ij})_{j \in N}$ with $a_{ij} \geq a'_{ij}$ for all $j \in N$, and all other pollution impacts being unchanged ($a'_{lj} = a_{lj}$ for all $l \in N \setminus \{i\}$ and all $j \in N$). The distribution rule ϕ renders agents responsible for their pollution impact if for any $i \in N$, any reduction a' of i 's pollution impact from a ,

$$\phi_i(a') - \phi_i(a) = W(a') - W(a).$$

Responsibility for pollution impact (RPI) requires to assign to any agent the full return or loss of any change of his own pollution impact.

In addition to being a fairness principle, RPI has attractive incentive properties. Suppose that an agent is able to reduce his pollution impact at some cost by switching to a greener technology, reducing or cleaning its wastes, improving energy efficiency or using less toxic inputs. By assigning the full return of this pollution reduction, RPI provides efficient incentives to invest in pollution impact reduction. Symmetrically, if an agent benefits from increasing his pollution impact per unit of emissions (e.g. using higher sulfur content coal), RPI assigns to this agent the economic cost of this extra pollution.

Among the above regulations, the Pigouvian fee regulation scheme is efficient. The welfare

distribution rule associated with the Pigouvian fee regulation scheme satisfies RPI. The Pigouvian fee regulation scheme is feasible if the collected revenue is redistributed among agents. The implemented welfare distribution rule does not satisfy non-negativity unless the money collected is used to cover agents' damages: it then leads to the PP welfare distribution when marginal damage are constant as assumed here. An emission norm $\bar{e}_i = e_i^*$ with a persuasive fine (e.g. infinite) is efficient and feasible but its welfare distribution does not satisfy RPI and non-negativity. A cap-and-trade system (with tradable pollution allowances) for pollution at each receptor with grandfathering is efficient and feasible but its associated welfare distribution rule does not satisfy non-negativity since victims are not entirely compensated. It might or might not satisfy RPI depending on the initial allocation of permits.

Theorem 1 *The polluter-pays distribution rule is the unique distribution rule that satisfies non-negativity and responsibility for pollution impact.*

Proof. First, we show that if a distribution rule satisfies non-negativity and responsibility for pollution impact, then it must be the polluter-pays distribution rule ϕ^{PP} . Consider another distribution rule ϕ and let a be a problem. Let $\phi(a) = \tilde{z}$ and $\phi^{PP}(a) = z^{PP}$. Let $\sum_{i \in N} \tilde{z}_i = \tilde{W}$. Suppose that $\tilde{z} \neq z^{PP}$. Since $\sum_{i \in N} z_i^{PP} = W(a)$ and \tilde{z} is a welfare distribution, we have $\tilde{W} \leq W(a)$. Thus, $\sum_{i \in N} \tilde{z}_i \leq \sum_{i \in N} z_i^{PP}$ which combined with $\tilde{z} \neq z^{PP}$ forces $\tilde{z}_i < z_i^{PP}$ for some $i \in N$. Note that for all $j \in V$, $z_j^{PP} = 0$ and by non-negativity of ϕ , $\tilde{z}_j \geq 0$. Thus, we must have $i \in N \setminus V$ and both $a_{ii} > 0$ and $\hat{e}_i > 0$. Let a' be such that a is a pollution impact reduction for agent i from a' such that $a_{ii} < a'_{ii}$ and everything else remains identical, i.e. $a'_{lj} = a_{lj}$ for all $l, j \in N$ such that $lj \neq ii$. Pick a'_{ii} sufficiently large such that

$$b_i(e_i^{lf}) < z_i^{PP} - \tilde{z}_i \quad (7)$$

where $\mathcal{N}(t^{lf}, a') = \{e^{lf}\}$. Let $\phi(a') = \tilde{z}'$ and $\phi^{PP}(a') = z'^{PP}$ denote the distributions chosen by ϕ and ϕ^{PP} for the problem (N, b, a') . By responsibility for pollution impact,

$$\tilde{z}_i - \tilde{z}'_i = z_i^{PP} - z'^{PP}_i$$

Rearranging terms and using the definition of z_i^{PP} this leads to

$$z_i^{PP} - \tilde{z}_i = b_i(e_i^{l*}) - \sum_{j \in Ri} a'_{ij} e_i^{l*} - \tilde{z}'_i, \quad (8)$$

where e'^* denotes the efficient emission plan for (N, b, a') . By non-negativity of ϕ , $\tilde{z}'_i \geq 0$. Now since $b_i(e_i^{lf}) \geq b_i(e_i'^*)$, $a'_{ij} \geq 0$ for all $j \in Ri$, and $\tilde{z}'_i \geq 0$, we obtain from (7),

$$z_i^{PP} - \tilde{z}_i > b_i(e_i^{lf}) \geq b_i(e_i'^*) - \sum_{j \in Ri} a'_{ij} e_i'^* - \tilde{z}'_i,$$

which contradicts (8).

Second, we show that ϕ^{PP} satisfies non-negativity and responsibility for pollution impact. For non-negativity,

$$z_i^{PP} = b_i(e_i^*) - \sum_{j \in Ri} a_{ij} e_i^* = \max_{e_i \geq 0} \left(b_i(e_i) - \sum_{j \in Ri} a_{ij} e_i \right) \geq b_i(0) - \sum_{j \in Ri} a_{ij} \times 0 = 0,$$

where the inequality follows from the fact that agent i can always choose $e_i = 0$ (no emission or production).

For responsibility for pollution impact, for any agent i , consider any reduction of i 's pollution impact from a to a' : $a_{ij} \geq a'_{ij}$ for all $j \in N$ and $(a_{kj})_{j \in N} = (a'_{kj})_{j \in N}$ for any $k \neq i$. Let $\phi^{PP}(a) = z^{PP}$ and $\phi^{PP}(a') = z'^{PP}$. Let $W(a)$ and $W(a')$ denote the corresponding total welfare in (N, b, a) and (N, b, a') , respectively. Note that by efficiency of t^{PP} , we have both $W^{PP}(a) = W(a)$ and $W^{PP}(a') = W(a')$. Similarly, denote by e^* and e'^* the efficient emission plan of (N, b, a) and (N, b, a') , respectively. By definition,

$$z'^{PP} - z^{PP} = b(e'^*) - \sum_{j \in Ri} a'_{ij} e'^*_{ij} - \left(b(e_i^*) - \sum_{j \in Ri} a_{ij} e^*_{ij} \right). \quad (9)$$

Since $a_{kj} = a'_{kj}$ for every $k \neq i$, the efficient emission levels are not affected by the change of matrix of pollution impacts from a to a' which implies $e_k^* = e'^*_k$ for every $k \in N \setminus \{i\}$. Therefore, we have:

$$W(a') - W(a) = b(e'^*) - \sum_{j \in Ri} a'_{ij} e'^*_{ij} - \left(b(e_i^*) - \sum_{j \in Ri} a_{ij} e^*_{ij} \right)$$

which, combined with (9), leads to $z'^{PP} - z^{PP} = W(a') - W(a)$. \square

Because for any problem a , $\phi^{PP}(a)$ is an efficient welfare distribution, Theorem 1 shows that non-negativity and responsibility for pollution impact imply efficiency, i.e. for any problem the total welfare is distributed among the agents.

5 Generalization to Increasing Marginal Damages

We now consider the polluter-pays principle with convex damage functions which requires a slight modification of the model. We differentiate emissions from pollution and damage. The emission plan e generates a pollution level p_i at i 's location (to receptor i) defined by

$$p_i = \sum_{j \in Si} a_{ji} e_j. \quad (10)$$

The matrix a defines now the *transfer coefficients* that translates emissions of i into pollution of j (e.g. waste water released by i into water pollution concentration on j). Pollution at level p_i causes damages $d_i(p_i)$ to i with d_i being increasing and convex: $d_i(0) = 0$, $d'_i(p_i) > 0$ and $d''_i(p_i) \geq 0$ for every $p_i \in \mathbb{R}_+$ and $i \in N \setminus P$.¹² The welfare of agent i with emissions $e = (e_i)_{i \in N}$ is

$$b_i(e_i) - d_i(p_i), \quad (11)$$

where p_i is defined by (10). A pollution problem is now described by (N, b, a, d) .

The first-order conditions that characterize the unique efficient emission plan e^* (which maximizes the total welfare $\sum_{i \in N} [b_i(e_i) - d_i(p_i)]$) are for every $i \in N$ ¹³

$$b'_i(e_i^*) = \sum_{j \in Ri} a_{ij} d'_j(p_j^*) = \sum_{j \in Ri} a_{ij} d'_j \left(\sum_{l \in Sj} a_{lj} e_l^* \right). \quad (12)$$

The marginal benefit of agent i 's emission should be equal to its marginal cost for society which depends on its marginal impact on pollution a_{ij} and the marginal damage of pollution at each receptor $j \in Ri$. Each unit of emission from agent i leads to a_{ij} units of pollution at receptor j which causes marginal damages evaluated to $a_{ij} d'_j(p_j^*)$. The total welfare with the efficient emission plan e^* is:

$$W(a) = \sum_{i \in N} [b_i(e_i^*) - d_i(p_i^*)] = \sum_{i \in N} \left[b_i(e_i^*) - d_i \left(\sum_{j \in Si} a_{ji} e_j^* \right) \right].$$

¹²Recall that P is the set of only polluter agents.

¹³The existence of the efficient emission plan e^* is guaranteed by Brouwer's fixed point theorem: define $g : \times_{i \in N} [0, \hat{e}_i] \rightarrow \times_{i \in N} [0, \hat{e}_i]$ by $g(e) = ((b'_i)^{-1}(\sum_{j \in Ri} a_{ij} d'_j(\sum_{l \in Sj} a_{lj} e_l)))_{i \in N}$. Since b_i is strictly concave, b'_i tends to infinity at zero, and b'_i tends to zero at \hat{e}_i , g is a well defined function. Our assumptions on damages ensure that g is continuous. Now since $\times_{i \in N} [0, \hat{e}_i]$ is compact and convex, Brouwer's fixed point theorem implies that the function g must have a fixed point which is a solution to (12). Uniqueness of e^* follows from strict concavity of benefits and convexity of damages.

In contrast with constant marginal damages (i.e. the first-order condition in (2)), with increasing marginal damage the efficient level of i 's emission (the first-order condition in (12)) depends on what is emitted by the other polluters of j with j being a receptor of i 's pollution ($j \in Ri$). Marginal damage being increasing with pollution concentration, agent i 's emission has more impact on damages at j when pollutant emitted by other polluters in R^0j increases. Because a polluter's marginal impact depends on pollution concentration due to other polluters, applying the polluter-pays principle in this framework is not straightforward. One needs to define each polluter's responsibility on the damage caused to society when computing the "cost of pollution of one entity on others". With only one single polluter i , agent i should simply pay the damage $d_j(a_{ij}e_i)$ to victim j . However, with more than one polluter at receptor j , say i and k , the PP principle does not tell us how to share $d_j(a_{ij}e_i + a_{kj}e_k)$ (the overall cost at j) among i and k . If polluter i is held responsible for the first $a_{ij}e_j$ units of pollution, he has to pay $d_j(a_{ij}e_i)$. If polluter i is responsible for the last ones, he has to pay $d_j(a_{ij}e_i + a_{kj}e_k) - d_j(a_{kj}e_k)$ which is larger than $d_j(a_{ij}e_i)$ by convexity of d_j . It is also increasing with the other polluter k 's emissions. One can think about several ways to share the damage $d_j(p_j)$. For instance, it could be assigned proportionally to a polluter's share on total pollution, each polluter i paying $\frac{a_{ij}e_i}{p_j}d_j(p_j)$ to j for every $i \in Rj$.

Such a division of the damage is defined for given emissions by i and k . Yet, since emissions are substitutes for receptor j , the presence of i 's emissions at j leads to a reduction of k 's emissions e_k at the first-best. The inter-connection of polluters' efficient emissions with convex damage creates a further cost of pollution on society: i 's emission do not only cause damage at j , it also encroaches on k 's emission at the first-best.

In this framework, we interpret the PP principle of making paying the "cost of pollution of one entity on others" by charging a polluter the incremental impact of his emissions on other agents. Due to increasing marginal damage, we can distinguish between two impacts. A first one is an increase of damage at each receptor $j \in Ri$. The second one is due to the substitution between polluters' emissions for each receptor j : if i emits more pollution, then each polluter $k \in Sj$ should emit less at the first-best. We also interpret the PP principle by compensating each agent exactly for the damage caused by others' emissions in absence of his emission. Let us denote by e^{0i} the efficient emission plan *without* i 's emission for every $i \in N$

(with fixing $e_i^{0i} = 0$). Notice that e^{0i} is the efficient plan of the economy *without* i 's emission but *with* i 's damage function d_i (i.e. agent i is then a “victim only”). It maximizes the total welfare of the problem (N, b_{-i}, a, d) where by b_{-i} implicitly means that agent i becomes a victim. The polluter-pays regulation scheme t^{PP} is defined for every $i \in N$ by

$$t_i^{PP}(e) = d_i(p_i^{0i}) - \sum_{j \neq i} [b_j(e_j^{0i}) - d_j(p_j^{0i}) - (b_j(e_j) - d_j(p_j))]. \quad (13)$$

The transfer is decomposed in two terms. The left-hand term is agent i 's damage at the first-best without i 's emissions. The summation is the economic loss due to i 's emission for all other agents $j \neq i$. For a polluter only agent $j \in P$, the change is simply the loss of benefit $b_j(e_j^{0i}) - b_j(e_j)$. For a polluter j who is a victim of i 's pollution it is the change of welfare including damage $b_j(e_j^{0i}) - d_j(p_j^{0i}) - (b_j(e_j) - d_j(p_j))$.

The PP-scheme yields to agent i a total welfare of (noting $b_i(e_i^{0i}) = b_i(0) = 0$):

$$b_i(e_i) - d_i(p_i) + t_i^{PP}(e) = \sum_{j \in N} [b_j(e_j) - d_j(p_j) - (b_j(e_j^{0i}) - d_j(p_j^{0i}))]. \quad (14)$$

Agent i 's welfare under the PP regulation scheme is his emission's contribution to total welfare for any emission plan. Since each agent i internalizes the impact of his own emissions on total welfare given the other agent's emissions, the PP principle implements the efficient emission plan e^* . Indeed, given other agent's emissions e_{-i}^t , the maximization of agent i 's welfare

$$b_i(e_i) - \sum_{j \in Ri} d_j \left(a_{ij}e_i + \sum_{l \in Sj \setminus \{i\}} a_{lj}e_j^t \right) + \sum_{j \in N \setminus \{i\}} b_j(e_j^t) - \sum_{j \in N \setminus Ri} d_j(p_j^t) - \sum_{j \in N} (b_j(e_j^{0i}) - d_j(p_j^{0i}))$$

with respect to e_i leads to the first-order conditions in (12) of the efficient emission plan e^* . Therefore, $e^t = e^*$ for any $e^t \in \mathcal{N}(t, a)$.¹⁴ Thus, by (14), agent i 's equilibrium welfare is

$$\phi_i^{PP}(a) = b_i(e_i^*) - d_i(p_i^*) + t_i^{PP}(e^*) = W(a) - \sum_{j \in N} (b_j(e_j^{0i}) - d_j(p_j^{0i})). \quad (15)$$

where $W(a) = \sum_{j \in N} (b_j(e_j^*) - d_j(p_j^*))$. Similarly as before, we call ϕ^{PP} the polluter-pays distribution rule (induced by t^{PP} for convex damages). Agent i 's welfare is the incremental contribution of his emissions at the first-best. For a victim only agent $i \in V$, it simplifies to zero since he is fully compensated for the damage $d_i(p_i^*)$. A polluter only agent $i \in P$ obtains his first-best benefit $b_i(e_i^*)$ net of the negative impact of his emissions on society

¹⁴Note that if $\mathcal{N}(t, a)$ is not a singleton, then there would exist several efficient emission plans.

$\sum_{j \neq i} [b_j(e_j^*) - d_j(p_j^*) - (b_j(e_j^{0i}) - d_j(p_j^{0i}))]$. Note that since $e_j^{0i} = e_j^*$ with constant marginal damages for every $j \neq i$ the PP-scheme defined in (13) is a generalization of the one defined in (4) to convex damage functions. The next proposition shows that t^{PP} induces no budget deficit.

Proposition 2 *The polluter-pays scheme is an efficient and feasible regulation scheme implementing the polluter-pays distribution rule.*

Proof. Since t^{PP} is efficient, it suffices to show $\sum_{i \in N} t_i^{PP}(e^*) \leq 0$. Note that since e^{0i} is an efficient emission plan of the problem (N, b_{-i}, a, d) while the emission plan $(e_{-i}^*, 0)$ can be implemented in (N, b_{-i}, a, d) , we have

$$\sum_{j \in N} [b_j(e_j^{0i}) - d_j(p_j^{0i})] \geq -d_i(p_i^* - a_{ii}e_i^*) + \sum_{j \neq i} [b_j(e_j^*) - d_j(p_j^* - a_{ij}e_i^*)].$$

Multiplying both sides with -1, we combine the above inequality with the definition of $t^{PP}(e)$ in (13) at the first-best and use $b_i(e_i^{0i}) = b_i(0) = 0$ and $a_{ij} = 0$ for $j \notin Ri$, and obtain:

$$t_i^{PP}(e^*) \leq d_i(p_i^* - a_{ii}e_i^*) - \sum_{j \in R^{0i}} [d_j(p_j^*) - d_j(p_j^* - a_{ij}e_i^*)]$$

Summing up all transfers t_i^{PP} leads to:

$$\sum_{i \in N} t_i^{PP}(e^*) \leq \sum_{i \in N} (d_i(p_i^* - a_{ii}e_i^*) - \sum_{j \in R^{0i}} [d_j(p_j^*) - d_j(p_j^* - a_{ij}e_i^*)])$$

Rearranging terms yields:

$$\sum_{i \in N} t_i^{PP}(e^*) \leq \sum_{i \in N} (d_i(p_i^* - a_{ii}e_i^*) - \sum_{j \in S^{0i}} [d_i(p_i^*) - d_i(p_i^* - a_{ji}e_j^*)]). \quad (16)$$

Consider any $i \in N$. Without loss of generality, let $Si = \{1, \dots, s\}$. Since $p_i^* = \sum_{j \in Si} a_{ji}e_j^*$, we can rewrite $d_i(p_i^*)$ by:

$$d_i(p_i^*) = \sum_{k=1}^s [d_i(p_i^* - \sum_{j=1}^{k-1} a_{ji}e_j^*) - d_i(p_i^* - \sum_{j=1}^k a_{ji}e_j^*)] \quad (17)$$

Note that for any $k = 1, \dots, s$, $p_i^* - \sum_{j=1}^{k-1} a_{ji}e_j^* - (p_i^* - \sum_{j=1}^k a_{ji}e_j^*) = a_{ki}e_k^* = p_i^* - (p_i^* - a_{ki}e_k^*)$

Thus, by convexity of d_i , for any $k = 1, \dots, s$,

$$d_i(p_i^* - \sum_{j=1}^{k-1} a_{ji}e_j^*) - d_i(p_i^* - \sum_{j=1}^k a_{ji}e_j^*) \leq d_i(p_i^*) - d_i(p_i^* - a_{ki}e_k^*) \quad (18)$$

Combining (17) with (18) for any $k = 1, \dots, s$ leads to:

$$d_i(p_i^*) \leq \sum_{k=1}^s [d_i(p_i^*) - d_i(p_i^* - a_{ki}e_k^*)] = \sum_{j \in Si} [d_i(p_i^*) - d_i(p_i^* - a_{ji}e_j^*)].$$

By $Si = S^0i \cup \{i\}$, this is equivalent to:

$$d_i(p_i^* - a_{ii}e_i^*) \leq \sum_{j \in S^0i} [d_i(p_i^*) - d_i(p_i^* - a_{ji}e_j^*)].$$

The last inequality combined with (16) yields the desired conclusion. \square

Notice that as long as two polluters impact the same receptors, the PP distribution rule does not distribute total welfare in the sense that $\sum_{i \in N} \phi_i^{PP}(a) < W(a)$. To see that, suppose that $N = \{1, 2, 3\}$ with polluters 1 and 3 polluting only victim 2, i.e. $a_{i2} > 0$ for $i = 1, 3$. Then polluter 1 has to pay the incremental damage at 2, formally $d_2(a_{12}e_1^* + a_{32}e_3^*) - d_2(a_{32}e_3^{01})$ as well as the loss of benefit for 3, that is $b_3(e_3^{01}) - b_3(e_3^*)$. Similarly polluter 3 has to pay for the increment damages at 2 and benefit loss for 1 due to his emissions. The victim 2 receives a compensation equals to the damage $d_2(p_2^*)$. Yet, the total payment by 1 and 3 more than offsets the compensation to 2: $t_1^{PP}(e^*) + t_3^{PP}(e^*) + t_2^{PP}(e_2^*) < 0$ because $-t_1^{PP}(e^*) - t_3^{PP}(e^*) > t_2^{PP}(e_2^*) = d_2(a_{12}e_1^* + a_{32}e_3^*)$. The PP regulation scheme exhibits a financial surplus and, therefore, the PP regulation rule distributes strictly less than total welfare.

To characterize the PP distribution rule ϕ^{PP} with marginal increasing damages, we introduce a further fairness principle, called single-polluter (social welfare) upper bounds. Its motivation relies on polluters' claims on the welfare improvement due to their economic activity as explained in the introduction or, equivalently, on their claim on payments when applying the PP principle. To minimize his payment, a polluter would claim responsibility only on the damage impact due to his own emission in absence of any other pollution at each receptor $j \in Ri$ (including himself). We call it agent i 's single-polluter welfare. Under this interpretation of the PP principle each agent i would enjoy an individual welfare of $\max_{e_i \geq 0} [b_i(e_i) - \sum_{j \in Ri} d_j(a_{ij}e_i)]$. On the other hand, applying the PP principle requires to fully compensate any agent j for the damage $d_j(p_j)$. With (strictly) convex damage functions we have $\sum_{i \in Sj} d_j(a_{ij}e_i) < d_j(p_j) = d_j(\sum_{i \in Sj} a_{ij}e_i)$ whenever $|Sj| > 1$ and $p_j > e_i$, and such an interpretation of the polluter pays principle would lead to unbalanced transfers. One way

to reconcile a distribution rule (or feasible transfers) with the above claims is to impose that, by solidarity, no agent should get more than the claimed single-polluter welfare.

Single-Polluter Upper Bounds: For all problems a and all $i \in N$, $\phi_i(a) \leq \max_{e_i \geq 0} [b_i(e_i) - \sum_{j \in Ri} d_j(a_{ij}e_i)]$.

A second, and more fundamental, justification of solidarity upper bounds finds its roots in Moulin's notion of group externality (Moulin, 1990). Under increasing marginal damage, the presence of pollution from other sources might reduce the ability of a polluter to emit. Formally, let us denote by e_i^{0-i} polluter i 's efficient emission when i is the only polluter to emit ($e_j = 0$ for every $j \neq i$). It is the efficient emission plan to the problem (N, b_i, a, d) (where b_i means that all agents in $N \setminus \{i\}$ become victims). It is also the solution to the maximization problem in the single-polluter upper bounds property. Note that if there exist $j \in Ri$ and $k \in Sj$ with $k \neq i$, then $e_i^{0-i} > e_i^*$: agent i could pollute more in the absence of k . Doing so, under the PP regulation scheme t^{PP} , he could enjoy a welfare of $b_i(e_i^{0-i}) - d_i(p_i^{0-i}) + t_i^{PP}(e_i^{0-i}) = \max_{e_i \geq 0} [b_i(e_i) - \sum_{j \in Ri} d_j(a_{ij}e_i)]$, which is higher than his welfare under the emission plan e^* . In Moulin's terms, the presence of other polluters exhibits a negative group externality on polluter i . Single-polluter upper bounds require that every polluter who creates this negative group externality should take up a share of it. For victim only polluters $i \in V$, the single-polluter upper bound is equal to zero which is their welfare under the PP-scheme.¹⁵

We now provide our characterization of the PP principle generalized to increasing marginal damages.

Theorem 2 *The polluter-pays distribution rule is the unique distribution rule that satisfies non-negativity, responsibility for pollution impact and the single-polluter upper bounds.*

Proof. Let ϕ be a distribution rule satisfying non-negativity, responsibility for pollution impact and the solidarity upper bounds. Let a be a problem, $\phi(a) = z$ and $\phi^{PP}(a) = y^*$.

¹⁵It is worth to notice that for linear damages the single-polluter upper bounds of any agent i coincide with his welfare under $\phi^{PP}(a)$. This is because the social impact of a polluter does not depend on emissions by others: it is the same whatever the others are emitting.

Suppose that $z \neq y^*$. Note that for all $i \in V$, by non-negativity and solidarity upper bounds, $z_i = 0 = y_i^*$. Thus, there exists $i \in N \setminus V$ such that $z_i \neq y_i^*$.

Let $a'_{ii} > a_{ii}$. Consider the problem where a_{ii} changes to a'_{ii} and everything else remains identical, i.e. $a' = (a_{-ii}, a'_{ii})$. Let $\phi(a') = z'$, $\phi^{PP}(a') = y'^*$, and e'^* denote the efficient emission plan for a' .

By RPI,

$$z_i - z'_i = W(a) - W(a') = y_i^* - y'^*_i.$$

Now we may take limits, i.e.

$$\begin{aligned} \lim_{a'_{ii} \rightarrow +\infty} z_i - z'_i &= \lim_{a'_{ii} \rightarrow +\infty} W(a) - W(a') \\ &= \lim_{a'_{ii} \rightarrow +\infty} y_i^* - y'^*_i, \end{aligned}$$

and we obtain

$$\begin{aligned} z_i - \lim_{a'_{ii} \rightarrow +\infty} z'_i &= W(a) - \lim_{a'_{ii} \rightarrow +\infty} W(a') \\ &= y_i^* - \lim_{a'_{ii} \rightarrow +\infty} y'^*_i. \end{aligned}$$

Note that $\lim_{a'_{ii} \rightarrow +\infty} \max_{e_i \geq 0} [b_i(e_i) - \sum_{j \in Ri} d_j(a'_{ij} e_i)] = 0$. Therefore, by non-negativity and solidarity upper bounds, both $\lim_{a'_{ii} \rightarrow +\infty} z'_i = 0 = \lim_{a'_{ii} \rightarrow +\infty} y'^*_i$. But now we obtain

$$z_i = W(a) - \lim_{a'_{ii} \rightarrow +\infty} W(a') = y_i^*,$$

which contradicts $z_i \neq y_i^*$.

Second, we show that ϕ^{PP} satisfies RPI, non-negativity and solidarity upper bounds. From (15) it is straightforward that ϕ^{PP} satisfies RPI because e^{0i} is optimal for both (N, b_{-i}, a, d) and (N, b_{-i}, a', d) whenever i 's pollution impact is reduced (with $a'_{lj} = a_{lj}$ for all $l \in N \setminus \{i\}$ and all $j \in N$). Since e^{0i} can be implemented as an emission plan in the problem (N, b, a, d) , $W(a) \geq \sum_{j \in N} (b_j(e_j^{0i}) - d_j(p_j^{0i}))$ and, therefore, by (15), ϕ^{PP} satisfies non-negativity. For solidarity upper bounds, first note that by convexity of d_j ,¹⁶ we have for any $e_i \geq 0$,

$$d_j(a_{ij} e_i) \leq d_j(a_{ij} e_i + p_j^{*-i}) - d_j(p_j^{*-i}),$$

¹⁶Note that $d_j(0) = 0$ and therefore, d_j is superadditive: $d_j(u) + d_j(v) \leq d_j(u + v)$ for any $u, v \in \mathbb{R}_+$.

where $p_j^{*-i} = \sum_{k \in Rj \setminus \{i\}} a_{kj} e_k^*$. Therefore, for any $i \in N$ and $e_i \geq 0$,

$$b_i(e_i) - \sum_{j \in Ri} \left(d_j(a_{ij} e_i + p_j^{*-i}) - d_j(p_j^{*i}) \right) \leq b_i(e_i) - \sum_{j \in Ri} d_j(a_{ij} e_i).$$

Maximizing both sides of the inequality with respect to e_i leads to:

$$b_i(e_i^*) - \sum_{j \in Ri} (d_j(p_j^*) - d_j(p_j^* - a_{ij} e_i^*)) \leq \max_{e_i \geq 0} [b_i(e_i) - \sum_{j \in Ri} d_j(a_{ij} e_i)]. \quad (19)$$

Second, since e^{0i} maximizes $-d_i(p_i) + \sum_{j \in N \setminus \{i\}} (b_j(e_j) - d_j(p_j))$ while $(0, e_{-i}^*)$ is a possible emission plan for (N, b_{-i}, a, d) , it yields a higher total welfare:

$$-d_i(p_i^{0i}) + \sum_{j \in N \setminus \{i\}} (b_j(e_j^{0i}) - d_j(p_j^{0i})) \geq -d_i(p_i^* - a_{ii} e_i^*) + \sum_{j \in N \setminus \{i\}} (b_j(e_j^*) - d_j(p_j^* - a_{ij} e_i^*)).$$

Multiplying both sides by -1 , adding $W(a)$ to both sides, and using the definition of ϕ^{PP} in (15) yields:

$$\phi_i^{PP}(a) \leq b_i(e_i^*) - \sum_{j \in Ri} (d_j(p_j^*) - d_j(p_j^* - a_{ij} e_i^*)).$$

The last inequality combined with (19) shows that solidarity upper bounds holds for all $i \in N$. \square

For linear damages, by Theorem 1, non-negativity and responsibility for pollution impact imply that for any problem the total welfare is distributed among the agents. For increasing marginal damages, non-negativity, responsibility for pollution impact (RPI) and solidarity upper bounds imply that *not* for any problem the total welfare is distributed among the agents. Here RPI does not imply efficiency.

6 PP versus VCG

6.1 VCG and Pollution Emissions

We compare the PP-scheme with the Vickrey-Clark-Groves (VCG)-scheme applied to economies with externalities.¹⁷ A VCG-scheme would make each agent pay or receive his impact on total

¹⁷In a setting with a finite set of alternatives and where agents reveal their utilities, Moulin (1986) characterized VCG-schemes by strategy-proofness (agents reveal their true utilities) and other properties (see also

welfare. Given that agents are choosing emission levels, the VCG-scheme t^{VCG} assigns to every $i \in N$,

$$t_i^{VCG}(e) = \sum_{j \neq i} (b_j(e_j) - d_j(p_j)) + k_i,$$

where k_i is a constant which does not depend on agent i 's emissions.¹⁸ It leads to the VCG-distribution rule ϕ^{VCG} defined by

$$\phi_i^{VCG}(a) = W(a) + k_i.$$

The PP principle can be seen as a special case of the VCG-scheme where, for every $i \in N$, the constant terms are¹⁹

$$k_i = d_i(p_i^{0i}) - \sum_{j \neq i} (b_j(e_j^{0i}) - d_j(p_j^{0i}))$$

Another special case of the VCG-scheme is the pivotal scheme t^{piv} which sets the constant parameters as

$$k_i = - \max_{e_{-i}} \left\{ \sum_{j \neq i} (b_j(e_j) - d_j(p_j)) \right\},$$

for every $i \in N$ (where $p_j = \sum_{l \in S_j \setminus \{i\}} a_{lj} e_l$). The lump-sum transfer paid by i is the total welfare at the first-best *without* i . Note that “without i ” here means both without i 's emission and without i 's damage. Let e^{-i*} denote the solution to the above maximization program, that is the efficient emission plan of the pollution problem without i denoted $(N, b_{-i}, a_{-i}, d_{-i})$ for every $i \in N$. The pivotal scheme t^{piv} assigns to every $i \in N$,

$$t_i^{piv}(e) = \sum_{j \neq i} (b_j(e_j) - d_j(p_j)) - \sum_{j \neq i} (b_j(e_j^{-i*}) - d_j(p_j^{-i*})).$$

Under the pivotal scheme, each agent i obtains the total welfare net of the welfare without i at the first-best. Therefore, agent i internalizes the impact of his emission on society which Thomson (1976) for the case of two alternatives). Note that here agents simply choose their emission (like in real life) and do not report utility functions and the set of emissions is infinite. The case of agents reporting their preferences (benefits and damages) is briefly analyzed in Section 6.2.

¹⁸Note that, in the general model with convex damages, efficiency (or RPI) requires k_i cannot depend on the emissions by other agents e_{-i} for every $i \in N$. In contrast, with linear damages, since agents' best reply functions are orthogonal, e_i^* is independent of e_{-i}^* for every $i \in N$.

¹⁹Note that the constant k_i depends on i 's preferences via his damage function d_i which impacts e^{0i} .

means that the pivotal scheme is efficient, i.e. $\mathcal{N}(t^{piv}, a) = \{e^*\}$. It leads to the pivotal distribution rule ϕ^{piv} defined for every $i \in N$ by:

$$\phi_i^{piv}(a) = W(a) - \sum_{j \neq i} (b_j(e_j^{-i*}) - d_j(p_j^{-i*})).$$

Each agent $i \in N$ obtains the difference between the welfare with and without him at the first-best. In case of unilateral externalities, for a victim-only agent $i \in V$, $\phi_i^{piv}(a) < 0$ because i 's presence in the economy reduces total welfare. Therefore, non-negativity of the distribution rule induced by t^{piv} is violated. Indeed, agent i does not only bring his damage d_i to the economy which diminishes welfare, it also forces polluters $j \in Si$ to reduce their emissions. Hence, in addition to not being compensated for the damage $d_i(p_i^*)$, a victim i has to pay for the loss of welfare which his presence causes to the polluters, namely $\sum_{j \in Si} [(b_j(e_j^{-i*}) - d_j(p_j^{-i*})) - (b_j(e_j^*) - d_j(p_j^*))]$. For a polluter-only agent $i \in P$, the PP and pivotal welfare coincide because $e_j^{0i} = e_j^{-i*}$ for every $j \in N \setminus \{i\}$ while $d_i = 0$ for any $i \in P$. Therefore $\phi_i^{PP}(a) = \phi_i^{piv}(a)$ for any $i \in P$. With multilateral externalities pollution is a public bad and the pollution problem is closer to the public good provision framework in which the pivotal scheme has been put forward. Although the pivotal scheme satisfies responsibility for pollution impact (RPI) and the single-polluter upper bounds, it fails to satisfy non-negativity. An agent i adds both new emission e_i and new damage d_i to the welfare. Agent i pollutes other agents and forces in addition them to reduce their own emissions from e_j^{-i*} to e_j^* for every $j \in Si$ and $j \in Sk \setminus \{i\}$ where $k \in Ri$ for convex damage function d_i . Therefore, under multilateral externalities, we may have $t_i^{piv}(e^*) < 0$. It is easy to find examples in which the negative impact of his presence $t_i^{piv}(e^*)$ to society is not compensated by i 's net benefit $b_i(e_i^*) - d_i(p_i^*)$ at the first-best meaning that $\phi_i^{piv}(a) < 0$ for every $i \in N$.²⁰ This is a major drawback of the pivotal scheme. Under the PP principle, agents pay for the negative impact of their emissions on society, and they are compensated for their incurred damage.

6.2 Preference Revelation and Incentive Compatibility

While the focus is here on agents choosing emission levels, one could take an alternative approach whereby agents reveal preferences and afterwards an emission plan is chosen which is

²⁰A detailed example is available from the authors upon request.

efficient for the revealed preferences. Transfers are determined according to revealed preferences. Agent i 's preference is given by his benefit b_i and his damage d_i . Let $(b, d) = (b_i, d_i)_{i \in N}$. A general VCG-distribution rule is defined by transfers

$$t_i^{VCG}(b, d) = \sum_{j \neq i} (b_j(e_j^*) - d_j(p_j^*)) + h_{-i}(b_{-i}, d_{-i}), \quad (20)$$

where e^* is the efficient emission plan for (b, d) and h_{-i} is an arbitrary function. Then the distribution rule (using (b, d) as arguments) based on t^{VCG} is defined for all $i \in N$ by

$$\phi_i^{VCG}(b, d) = b_i(e_i^*) - d_i(p_i^*) + t_i^{VCG}(b, d).$$

Because e^* is an efficient emission plan and revealing arbitrary (b_i, d_i) induces changes of the emissions, revealing the true (b_i, d_i) weakly dominates revealing any other (b'_i, d'_i) . In other words, the VCG-scheme is incentive compatible because truth-telling is a weakly dominant strategy. Formally, a general distribution rule ϕ is *incentive compatible* if for all (b, d) , all $i \in N$ and all (b'_i, d'_i) we have

$$\phi_i(b, d) \geq \phi_i((b'_i, b_{-i}), (d'_i, d_{-i}))$$

(where all benefits and damages are supposed to satisfy our conditions). It is well known that in our quasi-linear setting any incentive compatible distribution rule must be a general VCG-distribution rule with some functions h_{-i} .

If damages are complete information, setting

$$h_{-i}^{piv}(b_{-i}, d_{-i}) = d_i(p_i^{0i}) + \sum_{j \neq i} (b_j(e_j^{0i}) - d_j(p_j^{0i})),$$

the pivotal distribution rule using transfers t^{piv} (for fixed damages) is identical with the PP-scheme. Then the PP-scheme is incentive compatible for revealing true benefit functions. Once damages are not known, PP is not incentive compatible because the function h_{-i} cannot depend on the damage d_i . Therefore the pivotal scheme and PP-scheme differ when preferences (benefits and damages) are private information.

7 Conclusion

Most of the economic literature on the choice of policy instruments to tackle environmental issues focuses on efficiency. In contrast here, we analyze the fairness properties of welfare

distributions implemented by environmental regulations. To do so, we rely on a general model in which agents can be polluters, victims of pollution or both. We assume that the benefit from emitting pollutants is separated from the damage due to pollution. This allows disentangling two effects of pollution emissions on individuals: the direct externality on the victims and the indirect externality on other polluters. Pollution does not only increase damages, but it also reduces the amount of pollution that other polluters emit at first-best. In Moulin (1990)'s terms, the presence of other polluters creates a negative group externality. Although the group externality is absent with constant marginal damages, it is present when marginal damages are increasing with pollution. We introduce welfare bounds that mitigate the negative impact of the two externalities. The non-negativity lower bounds limits the pollution externality: it makes sure that nobody is worse off with pollution than without. The single-polluter upper bounds share the cost of the group externality among polluters: no polluter obtains more than what he would achieve without group externality. In addition to the above two bounds on welfare, we introduce a fairness criterion called responsibility for pollution impact (RPI) that accounts for both externalities. When a polluter reduces his pollution impact on society, this does not only reduce the damages he causes. It also increases the benefit of other pollutants by allowing them to emit more at the first-best. RPI requires that a polluter obtains exactly the full return of his reduction, which includes the effect of both externalities. It turns out that the only welfare distribution rule satisfying the three criteria is the one implemented by the polluter-pays principle.

Although our focus was on fairness properties, the scheme based on the polluter-pays principle has also interesting properties when pollution emissions or benefits from emissions are not observed by the regulator. It is indeed incentive compatible for choosing the efficient emission levels or for revealing the true benefit functions. To that respect, it can be seen as a special case of the VCG-scheme, which is not the pivotal scheme. Yet when both benefits and damages are private information, unlike the VCG-scheme, the polluter-pays scheme cannot be used because information on damages is necessary to determine the transfers of the PP-scheme.

Our analysis can be extended in many directions.²¹ First, instead of increasing marginal damages, one may consider decreasing marginal damages. Although this assumption is not

²¹We thank a referee for raising these interesting issues.

common in environmental economics, such problems might arise in previously-unspoiled natural environments, where the first action causes the greatest damage, and subsequent actions may matter rather less. Then there may exist multiple efficient emission plans which induce different welfare distributions and one may have to identify additional criteria to select one efficient emission plan.

Second, we did not consider behavioral changes of victims from pollution. In many cases, victims have several actions available via which they can protect themselves from harmful impacts of pollution. They can move further away from the source of pollution, invest in air-cleaning devices, better isolate their homes against noise or filtrate unsafe water. Such responses are often called “defense activities”. It is well-known that compensating fully victims of pollution causes underinvestment (or no investment at all) in defensive activities (see Baumol and Oates (1988, p21) for a discussion on this issue). With defensive activities, our literal interpretation of the polluter-pays principle would lead to an inefficient outcome. To improve efficiency, the responsibility for pollution impact needs to be shared among all agents: polluters and victims of pollution. The relative share assigned to each agent is not straightforwardly defined and is context specific. Whether efficiency can be achieved together with some fairness properties is an open question. We leave this for future research.

References

- Ambec, S., and Y. Sprumont (2002): “Sharing a River,” *Journal of Economic Theory* 107, 453–462.
- Ambec, S., and L. Ehlers (2008): “Sharing a River among Satiabile Agents,” *Games and Economic Behavior* 64, 35–50.
- Barret, S. (1994): “Self-enforcing international environmental agreements” *Oxford Economic Papers* 46, 878–894.
- Baumol, W.J., and W.E. Oates (1975): *The Theory of Environmental Policy*. Prentice Hall, Englewood Cliffs, NJ.
- Carraro, C. and D. Siniscalco (1993): “Strategies for the International Protection of the Environment,” *Journal of Public Economics* 52, 309–328.
- Chander P. and H. Tulkens (1997): “The Core of an Economy with Multilateral Environmental Externalities,” *International Journal of Game Theory* 26, 379–401.
- Coase, R.H. (1960): “The Problem of Social Cost,” *Journal of Law and Economics* 3, 1–44.
- Demange, G. (2004): “On Group Stability in Hierarchies and Networks,” *Journal of Political Economy* 112, 754–778.
- Duggan, J. and J. Robert (2002): “Implementing the Efficient Allocation of Pollution,” *American Economic Review* 92, 1070–1078.
- Finus, M. and S. Tjøtta (2003): “The Oslo Protocol on sulfur reduction: the great leap forward?,” *Journal of Public Economics* 87, 2031–2048.
- Fleurbaey, M. (2008): *Fairness, responsibility and welfare*, Oxford University Press, Oxford.
- Heyward, M. (2007): “Equity and International Climate Change Negotiations: a Matter of Perspective,” *Climate Policy* 7, 518–534.
- Kolstad, C. (2000): *Environmental Economics*, Oxford University Press, Oxford.

- Lange, A., A. Löschel, C. Vogt and A. Ziegler (2010): “On the Self-Serving Use of Equity Principles in International Climate Negotiation,” *European Economic Review* 54, 359–375.
- Mäler, K.-G. (1989): “The Acid Rain Game,” In: Folmer H. and van Ierland E. (eds.), *Valuation Methods and Policy Making in Environmental Economics*, Elsevier, Amsterdam, p.231–252.
- Mäler, K.-G. (1994): “Acid Rain in Europe: A Dynamic Perspective on the Use of Economic Incentives,” In: van Ierland, E.C.(ed) *International Environmental Economics*, Elsevier, Amsterdam, p.351–371.
- Mäler, K.-G. and A. de Zeeuw (1998): “The acid differential game,” *Environmental and Resource Economics* 12, 167–184.
- Montero, J.-P. (2008): “A Simple Auction Mechanism for the Optimal Allocation of the Commons,” *American Economic Review* 98, 496–518.
- Montgomery, W.D. (1972): “Markets in Licenses and Efficient Pollution Control Programs,” *Journal of Economic Theory* 5, 395–418.
- Moulin, H. (1986): “Characterizations of the Pivotal Mechanism,” *Journal of Public Economics* 31, 53–78.
- Moulin, H. (1990): “Uniform externalities, two axioms for fair allocation.” *Journal of Public Economics* 43, 305–326.
- Ostrom, E. (1990): *Governing the commons*. Cambridge University Press, Cambridge.
- Thomson, W. (1976): *Incentives and Information*, Ph.D. Dissertation, Stanford University.