WORKING PAPERS IN ECONOMICS

No. 17/02

SJUR DIDRIK FLÅM

BALANCED ENVIRONMENTAL GAMES

Department of Economics

UNIVERSITY OF BERGEN

Balanced Environmental Games

Sjur Didrik Flåm*

Sept. 23, 2002

Abstract. Focus is here on coalitional games among economic agents plagued by aggregate pollutions of diverse sorts. Defecting players presumably pollute more than others. Then, granted convex preferences and technologies, the core is proven nonempty. In fact, under natural assumptions, a specific, computable core solution comes in terms of shadow prices on the said aggregates. Such prices may, in large part, implement the cooperative treaty by clearing a competitive market for emissions.

1. Introduction

Environmental degradation - and the prospect of climate change - has motivated many game theoretic studies, often focused on cooperation and core solutions. This note pursues that line of research. It adds to the results of Helm (2001), Chander and Tulkens (1997) by allowing technological externalities, more general utility functions, and several pollutants. More important, it treats the aggregate discharge from defecting players axiomatically, presuming merely that these agents will, on the average, pollute most.

Pollutants are "uniformly dispersed" in global commons. So, given the total emission (say, of greenhouse gases), it does not matter for external effects who contributed how much.[1] Then, granted convexity in preferences and production, the core proves nonempty. Thus, in principle, the prospects for efficient and stable cooperation may be rather good. That point is reinforced here in two ways: First, under natural assumptions, a specific core solution can be computed. Second, that solution, being determined by shadow prices on aggregate emissions, seems implementable by trades in competitive emission markets.

2. The Game

Accommodated here is a finite, fixed set I, each member i being a consumer, producer and polluter - all at the same time. Correspondingly, the consumption and emission bundles of these agents will come into focus next. Such bundles belong to ordered, real vector spaces $(X, \leq)$ and $(E, \leq)$, respectively.[2] Since goods and "bads" are

*Economics Department, Bergen University 5007 Norway; sjur.flaam@econ.uib.no. Thanks for financial support is due Ruhrgas and Røwdes fond.

[1] Clearly, if dispersion is tilted towards some receiving regions, then polluters are not on equal footing. Acid rain is an important case. See [7], [11].

[2] A vector space V is *ordered* by a binary relation $\leq$ if there exists a convex cone $K \subset V$ such that $v \leq v' \Leftrightarrow v' - v \in K$. For example, the nonnegative orthant defines the customary order in the ambient Euclidean space.

manifold, and differentiated by their availability in location or time, the spaces X and E could have large dimensions.[3]

Agent $i \in I$ contemplates consuming a commodity vector $x_i \in X$ and emitting a vector $e_i \in E$ of pollutants. Thereby he obtains real-valued payoff (or transferable utility) $\pi_i(x_i, e_I)$ where $e_I := \sum_{i\in I} e_i$ denotes the total emission. In autarchy i would face the technological constraint $x_i \leq f_i(e_i, e_I)$.

As customary, we take $f_i : E \times E \to X$ and $\pi_i : X \times E \to R \cup \{-\infty\}$ to be concave functions, both increasing in the first argument and decreasing in the second. Thus, each player is directly and adversely affected by the total emission e_I.[4] The extreme payoff $\pi_i(x_i, e_I) = -\infty$ serves here as a fictitious, but convenient "death penalty." By indicating violation of underlying constraints this simple device saves us repeated and explicit mention of evident restrictions (such as nonnegativity or capacity limits).

Our concern is with the prospects of cooperation. Can the grand coalition form? Can it secure efficiency and split the potential gains to satisfy every party? More precisely, for a suitable characteristic function, *is the core empty*?

While addressing that issue, Chander and Tulkens (1997) defined the *γ-worth of coalition $S \subseteq I$* by considering *a noncooperative game against S*. Specifically, using shorthand expressions $e_S := \sum_{i\in S} e_i$ and $e_{-S} := \sum_{i\in I \diagdown S} e_i$, then, in that game,
• S acts as *one* player with objective $\sum_{i\in S} \pi_i(x_i, e_S + e_{-S})$ and constraints $\sum_{i\in S} x_i \leq \sum_{i\in S} f_i(e_i, e_S + e_{-S}), \sum_{i\in S} e_i \leq e_S$;
• each outsider $i \in I \diagdown S$ plays with similar objective $\pi_i(x_i, e_i + e_{-i})$ and constraint $x_i \leq f_i(e_i, e_i + e_{-i})$.

Definition (The γ-characteristic function). *The worth $v^\gamma(S)$ of coalition S is the Nash equilibrium payoff it obtains in the described game against S.* □

Clearly, this definition is somewhat demanding. To see how, let $\tilde{\mathbf{e}} = (\tilde{e})_{i\in I}$ denote a complete list of emissions and define, for any coalition S, its *best reply* (correspondence) $B_S(\tilde{\mathbf{e}})$ to comprise each emission profile $(e_i)_{i\in S}$ that together with a suitable consumption pattern $(x_i)_{i\in S}$ would solve the problem:

$$\text{maximize} \sum_{i\in S} \pi_i(x_i, e_S + \tilde{e}_{-S}) \text{ s.t. } \sum_{i\in S} x_i \leq \sum_{i\in S} f_i(e_i, e_S + \tilde{e}_{-S}), \sum_{i\in S} e_i \leq e_S \qquad (1)$$

Then, $\mathbf{e}$ is a Nash equilibrium (in the game against S) iff $e_i \in B_{\{i\}}(\mathbf{e})$ for all $i \in I \diagdown S$ and $(e_i)_{i\in S} \in B_S(\mathbf{e})$.

One may reasonably ask whether $v^\gamma(S)$ is well defined in this manner. That is: does equilibrium exists in all these games? Is it always unique? Does the definition invariably provide a unique value $v^\gamma(S)$?[5]

[3]For simplicity one may think of one consumption good and one pollutant. Then X and E are one-dimensional with the usual order.

[4]Put differently: everybody falls victim to spillovers produced by the others. Such spillovers are discharged in global commons.

[5]An equilibrium *exists* if each pair (x_i, e_i) must belong to a nonempty compact convex set $K_i \subset$

We shall circumvent these difficulties via a more axiomatic approach. For the statement recall that, given a collection $\mathcal{C}$ of coalitions, a corresponding real-valued mapping $S \in \mathcal{C} \mapsto \delta_S \geq 0$ is declared a *balanced collection of weights* iff for each $i \in I$ we have $\sum \{\delta_S : i \in S \in \mathcal{C}\} = 1$. For simplicity we write henceforth $\mathcal{C}_i := \{S \in \mathcal{C} : i \in S\}$. Thus, $\sum_{S \in \mathcal{C}_i} \delta_S = 1$ for all i. Using this notion we now make a key

Hypothesis (about anticipation, best reply, and free riding)
(i) (Foreseeable emissions and best reply) *Any coalition S, if it were to form, would face a foreseeable total emission $\tilde{e}_{-S}$ from the outsiders. In response, S would be worth the optimal value $v(S)$ of problem (1).*
(ii) (External agents are excessive free-riders) *For any balanced collection of weights and any agent i it holds that*

$$\sum_{S \in \mathcal{C} \setminus \mathcal{C}_i} \delta_S e_S \leq \sum_{S \in \mathcal{C}_i} \delta_S \tilde{e}_{-S}. \quad \square \tag{2}$$

We shall assume, of course, that the optimal value $v(S)$ in (1) be finite. That value need, however, not be attained. Clearly, our definition of worth ignores some problems by tacitly presuming informational symmetry. Also, it simplifies many conflictual issues by coaching cooperation merely in monetary terms [12].

Inequality (2) is crucial but hard to interpret: Suppose each agent i, whenever he belongs to a coalition S, takes part to the degree δ_S. Thereby he falls victim to proportional emission $\sum_{S \in \mathcal{C}_i} \delta_S \tilde{e}_{-S}$ from defecting outsiders. (2) says the latter item exceeds the aggregate discharge produced by the coalitions to which i does not belong. Broadly speaking, if coalition memberships are balanced, each agent will experience that defectors pollute more than do other contracting parties. To justify (2) we consider next an important case:

Proposition 1. (Excessive free-riding) *Suppose there exists a particular emission pattern $\bar{\mathbf{e}} = (\bar{e}_i)_{i \in I}$ such that*

$$e_S \leq \bar{e}_S \quad \text{and } \bar{e}_{-S} \leq \tilde{e}_{-S} \quad \text{for every coalition } S. \tag{3}$$

More generally, suppose that for every balanced collection of weights $\delta_S, S \in \mathcal{C}$, and agent i there exists a particular emission pattern $\bar{\mathbf{e}} = (\bar{e}_i)_{i \in I}$ such that

$$\sum_{S \in \mathcal{C} \setminus \mathcal{C}_i} \delta_S e_S \leq \sum_{S \in \mathcal{C} \setminus \mathcal{C}_i} \delta_S \bar{e}_S \quad \text{and} \quad \sum_{S \in \mathcal{C}_i} \delta_S \bar{e}_{-S} \leq \sum_{S \in \mathcal{C}_i} \delta_S \tilde{e}_{-S}. \tag{4}$$

Then (2) *holds.*

$X \times E$, and all functions are continuous, properly defined on the ensuing sets. Moreover, given Euclidean spaces, equilibrium will be *unique* if the marginal payoffs $m_i := \partial \pi_i / \partial(x_i, e_i)$ constitute a strictly monotone map $m = (m_i)$ on the product set of all K_i.

Proof. Fix any balanced collection of weights $\delta_S, S \in \mathcal{C}$, and consider some agent i. Let $\bar{\mathbf{e}}$ be an emission pattern that satisfies (4). Then

$$\begin{aligned}\sum_{S\in\mathcal{C}\setminus\mathcal{C}_i}\delta_S e_S &\leq \sum_{S\in\mathcal{C}\setminus\mathcal{C}_i}\delta_S\bar{e}_S = \sum_{S\in\mathcal{C}}\delta_S\sum_{j\in S}\bar{e}_j - \sum_{S\in\mathcal{C}_i}\delta_S\bar{e}_S = \sum_j\sum_{S\in C_j}\delta_S\bar{e}_j - \sum_{S\in\mathcal{C}_i}\delta_S\bar{e}_S \\ &= \bar{e}_I - \sum_{S\in\mathcal{C}_i}\delta_S\bar{e}_S = \sum_{S\in\mathcal{C}_i}\delta_S\bar{e}_I - \sum_{S\in\mathcal{C}_i}\delta_S\bar{e}_S = \sum_{S\in\mathcal{C}_i}\delta_S\bar{e}_{-S} \leq \sum_{S\in\mathcal{C}_i}\delta_S\tilde{e}_{-S}. \quad \square\end{aligned}$$

One may think of the emission pattern $\bar{\mathbf{e}} = (\bar{e}_i)_{i\in I}$ in (3) as one that would emerge under total lack of cooperation. For example, it could stem from a Nash equilibrium in the noncooperative game having atomistic player set I. Free-riding would then explain $e_S \leq \bar{e}_S$. In addition, the tragedy of the commons would entail $\bar{e}_{-S} \leq \tilde{e}_{-S}$. Chander and Tulkens (1997) bring out such results formally. They used a version of ratio equilibrium (Kaneko 1977) to find a core solution. We rather follow Helm (2001)[6] in proving that the game is balanced:

Theorem 1. (A balanced game) *The coalitional game has a nonempty core.*

Proof. Pick any balanced collection of weights $\delta_S, S \in \mathcal{C}$. By the Bondareva-Shapley theorem [10] it suffices to verify that $v(I) \geq \sum_{S\in\mathcal{C}}\delta_S v(S)$. Fix any number $\varepsilon > 0$ and let $\varepsilon_S := \varepsilon/\sum_{S\in\mathcal{C}}\delta_S$. For any coalition $S \in \mathcal{C}$ let $(x_i^S, e_i^S)_{i\in S}$ solve problem (1) up to ε_S-optimality, and denote by $\tilde{e}_{-S}$ the aggregate emission produced by the outsiders. Define for each agent i a particular choice $(x_i, e_i) := \sum_{S\in\mathcal{C}_i}\delta_S(x_i^S, e_i^S)$. The plan so constructed is feasible. Indeed, since production is everywhere decreasing in total emissions, (4) entails

$$\begin{aligned}\sum_{i\in I}x_i &= \sum_{i\in I}\sum_{S\in\mathcal{C}_i}\delta_S x_i^S = \sum_{S\in\mathcal{C}}\delta_S\sum_{i\in S}x_i^S \leq \sum_{S\in\mathcal{C}}\delta_S\sum_{i\in S}f_i(e_i^S, e_S^S + \tilde{e}_{-S}) \\ &= \sum_{i\in I}\sum_{S\in\mathcal{C}_i}\delta_S f_i(e_i^S, e_S^S + \tilde{e}_{-S}) \leq \sum_{i\in I}f_i(e_i, \sum_{S\in\mathcal{C}_i}\delta_S e_S^S + \sum_{S\in\mathcal{C}_i}\delta_S\tilde{e}_{-S}) \\ &\leq \sum_{i\in I}f_i(e_i, \sum_{S\in\mathcal{C}_i}\delta_S e_S^S + \sum_{S\in\mathcal{C}\setminus\mathcal{C}_i}\delta_S e_S^S) = \sum_{i\in I}f_i(e_i, e_I).\end{aligned}$$

Also note that (4) implies

$$\begin{aligned}e_I &= \sum_{i\in I}e_i = \sum_{i\in I}\sum_{S\in\mathcal{C}_i}\delta_S e_i^S = \sum_{S\in\mathcal{C}}\delta_S e_S^S = \sum_{S\in\mathcal{C}_i}\delta_S e_S^S + \sum_{S\in\mathcal{C}\setminus\mathcal{C}_i}\delta_S e_S^S \\ &\leq \sum_{S\in\mathcal{C}_i}\delta_S e_S^S + \sum_{S\in\mathcal{C}_i}\delta_S\tilde{e}_{-S} = \sum_{S\in\mathcal{C}_i}\delta_S(e_S^S + \tilde{e}_{-S}).\end{aligned}$$

[6] Helm considers single commodities x_i and e_i. He uses quasi-linear utility: $\pi_i(x_i, e_I) = x_i - d_i(e_I)$ and production technology free from externalities: $f_i(e_i, e_I) = f_i(e_i)$.

Consequently, $\pi_i(x_i, e_I) \geq \pi_i(x_i, \sum_{S\in\mathcal{C}_i} \delta_S(e_S^S + \tilde{e}_{-S}))$ for each i. So, $v(I) \geq$

$$\begin{aligned}\sum_{i\in I}\pi_i(x_i,e_I) &\geq \sum_{i\in I}\pi_i(\sum_{S\in\mathcal{C}_i}\delta_S x_i^S, \sum_{S\in\mathcal{C}_i}\delta_S(e_S^S+\tilde{e}_{-S})) \geq \sum_{i\in I}\sum_{S\in\mathcal{C}_i}\delta_S\pi_i(x_i^S, e_S^S+\tilde{e}_{-S}) \\ &= \sum_{S\in\mathcal{C}}\delta_S\sum_{i\in S}\pi_i(x_i^S,e_S^S+\tilde{e}_{-S}) \geq \sum_{S\in\mathcal{C}}\delta_S\left[v(S)-\varepsilon_S\right] = \sum_{S\in\mathcal{C}}\delta_S v(S)-\varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary the desired conclusion follows. □

3. A Reduced Game with Emission Rights and Trade

Suppose here that an overall emission profile $(\check{e}_i)_{i\in I}$ has been agreed upon. More precisely, suppose total emission has been fixed, by contract, at $\check{e}_I$ together with a consistent allocation of property rights $\check{e}_i, i \in I$, $\sum_{i\in I}\check{e}_i = \check{e}_I$. For example, $\check{e}_I$ could be the total emission that solves problem (1) for $S = I$, and $(\check{e}_i)_{i\in I}$ might result from some principle of equity (or grandfathering).

In this simplified setting, where everybody holds emission rights and regards aggregate emissions as given, coalition S could achieve worth

$$\check{v}(S) := \sup\left\{\sum_{i\in S}\pi_i(x_i,\check{e}_I) : \sum_{i\in S}x_i \leq \sum_{i\in S}f_i(e_i,\check{e}_I), \sum_{i\in S}e_i \leq \check{e}_S\right\}.$$

Then, how can a core solution be found by a decentralized procedure? To that end suppose consumption and emission bundles could be purchased at fixed nonnegative *prices* x^* and e^*. By prices we understand real-valued, linear mappings $X \ni x \mapsto \langle x^*, x\rangle$, $E \ni e \mapsto \langle e^*, e\rangle$ such that $0 \leq x \Rightarrow 0 \leq \langle x^*, x\rangle$, and $0 \leq e \Rightarrow 0 \leq \langle e^*, e\rangle$. While regarding total emission $\check{e}_I$ and own endowment $\check{e}_i$ as given, let

$$u_i := \sup_{x_i,e_i}\left\{\pi_i(x_i,\check{e}_I) + \langle x^*, f_i(e_i,\check{e}_I) - x_i\rangle - \langle e^*, e_i\rangle\right\} + \langle e^*, \check{e}_i\rangle$$

denote the highest profit agent i could aim at under price-taking behavior. Note that the supremal term in u_i does not depend on the distribution of property rights. With u_i as just defined, we declare $x^* \geq 0, e^* \geq 0$ a pair of **shadow prices** if

$$\check{v}(I) \geq \sum_{i\in I}u_i.$$

Theorem 2. (Core solutions defined by shadow prices) *Suppose x^*, e^* are shadow prices. Then, for established property rights $\check{e}_i, i \in I$, $\check{e}_I = \sum_{i\in I}\check{e}_i$, the payment scheme $u_i, i \in I$, belongs to the core of the game that has characteristic function $S \mapsto \check{v}(S)$.*

Proof. Since

$$\check{v}(S) \leq \sup\sum_{i\in S}\left\{\pi_i(x_i,\check{e}_I) + \langle x^*, f_i(e_i,\check{e}_I) - x_i\rangle - \langle e^*, e_i\rangle + \langle e^*, \check{e}_i\rangle\right\} = \sum_{i\in S}u_i,$$

no coalition $S \subseteq I$ should reasonably block the proposed scheme of payments. In particular, $\check{v}(I) \leq \sum_{i \in I} u_i$. So, invoking the preceding assumption, we get $\check{v}(I) = \sum_{i \in I} u_i$ whence Pareto efficiency also obtains. This proves that the core is nonempty. $\square$

Given our convexity assumptions, existence of shadow prices (i.e.of so-called Lagrange multipliers) is ensured under standard qualifications. Also, when the agents are many and minor, these assumptions become relatively less important; see [2].

REFERENCES

[1] P. Chander and H Tulkens, A core of an economy with multilateral environmental externalities, *Int. J. Game Theory* 26, 379-401 (1997).

[2] I. V. Evstigneev and S. D. Flåm, Sharing nonconvex cost, *J. Global Optimization* 20, 257-271 (2001).

[3] Y, Funaki and T. Yamato, The core of an economy with a common resource pool: A partition function form approach, *Int. J. Game Theory* 28 (2) 157-171 (1999).

[4] N. Hanley, J. F. Shogren and B. White, *Environmental Economics*, Macmillan, London (1997).

[5] C. Helm, On the existence of a cooperative solution for a coalitional game with externalities, *Int. J. Game Theory* 30, 141-146 (2001).

[6] M. Kaneko, The ratio equilibria and the core of the voting game $G(N, V)$ in a public goods economy, *Econometrica* 45 (7) 1589-1594 (1977).

[7] K.G. Mäler, The acid rain game, in H. Folmer and E. van Ierland (eds.) *Valuation and Policy Making in Environmental Economics*, North-Holland, Amsterdam (1989).

[8] D. Ray and R. Vohra, Equilibrium binding agreements, *J. Economic Theory* 73, 30-78 (1997).

[9] R. W. Rosenthal, External economies and cores. *J. Economic Theory* 3, 182-188 (1971).

[10] L. S. Shapley, On balanced sets and cores, *Naval Research Logistics Quarterly* 14, 453-460 (1967).

[11] O. Tahvonen, V. Kaitala and M. Pohjola, A Finnish-Soviet acid rain game: noncooperative equilibria, cost efficiency and sulphur agreements, *J. Environmental Economics and Management* 24, 87-99 (1994).

[12] A. Underdal, The concept of regime effectiveness, *Cooperation and Conflict* 27, 3, 227-240 (1992).