International Environmental Agreements as Evolutionary Games

Matthew McGinty

Abstract   This paper applies evolutionary game theory to international environmental agreements (IEAs). Contrary to stage game models (Barrett in J Theor Politics 11:519–541, 1999, Eur Econ Rev 45:1835–1850, 2001), in an evolutionary equilibrium (EE) no signatory prefers to be outside the IEA and the EE is robust to trembles. With two populations, there is a unique interior EE when there is decreasing returns to abatement and small asymmetry in the externality differences across populations. At the interior EE, transfers from the poor to the rich can increase payoffs for all nations, but come at the expense of greater payoff inequality. Transfers can also eliminate the basin of attraction for the payoff inferior EE with decreasing returns to abatement and large asymmetry. Thus IEAs, such as the Kyoto Treaty, predicated on the polluter-pays and ability-to-pay principles may result in Pareto inferior outcomes.

Keywords   International environmental agreements · Evolutionary games · Externalities · Transfers

JEL Classification   C73 · C72 · D62 · D78 · D63

1 Introduction

Two main challenges in designing international environmental agreements (IEAs) are that participation is voluntary and that any punishment must be a credible threat. The incentive to join depends on the terms of the agreement, and the participation of others. An IEA can specify zero-sum transfers to increase participation when nations differ. However, both paying the transfer and imposing any punishment for defection must be a best-response to be credible. This paper is the first to address these challenges in an evolutionary game setting.

M. McGinty (✉)
Department of Economics, University of Wisconsin-Milwaukee,
PO Box 413, Milwaukee, WI 53201, USA
e-mail: mmcginty@uwm.edu

The early IEA literature (Barrett 1994; Carraro and Siniscalco 1993) typically finds a Nash equilibrium with partial participation for a one-shot game. There is a payoff advantage to IEA participation when participation is low, and a payoff advantage to free-riding when participation is high. IEA membership is modeled with sequential accession, using internal and external stability adopted from the cartel literature (d'Aspremont et al. 1983). With this approach Barrett (1994) shows that with significant gains to cooperation, there is only an incentive to join an IEA when participation is very low (less than 4). McGinty (2007) relaxes the assumption of symmetry in Barrett (1994) and finds greater participation when transfers are implemented in the form of tradable pollution permits. However, simulations are required to determine stability when non-linear functional forms are used.

Barrett (1999, 2001) considers a linear IEA modeled as a stage-game. In Barrett (1999) identical nations have either constant or increasing returns to abatement. Barrett (2001) introduces benefit asymmetry (across but not within populations) and considers constant returns to abatement. Barrett (2001) transforms a Prisoner's Dilemma with a dominant strategy of "pollute" into a stage game. In the first stage nations choose to participate or not in an IEA. In stage 2 signatories collectively choose pollute or abate, and non-signatories individually choose pollute or abate. The equilibrium is "lynchpin" in the sense that should any signatory defect and pollute, then all other signatories pollute and the agreement collapses. The punishment for defecting is pollute by all players, the Nash equilibrium of the simultaneous move game. A key assumption is that signatories choose their actions to maximize collective payoff, and comply with this decision. This result differs dramatically from that obtained in Barrett (1994), where one signatory choosing pollute does not result in all other nations playing pollute. The lynchpin equilibrium is found in previous IEAs (Chander and Tulkens 1995), however, the majority of the literature argues the non-cooperative outcome is not the relevant threatpoint when partial participation is the Nash equilibrium of the simultaneous move IEA (Barrett 1994, 1997; McGinty 2007; Weikard 2009). The question remains what types of IEAs will be obtained in Barrett's (1999, 2001) simple linear framework if the underlying game is that of partial participation, the returns to abatement vary, the externalities differ both within and across populations, and the timing assumptions of the stage game are relaxed.

The equilibrium in Barrett (2001) has two characteristics which this paper seeks to eliminate. First, signatories to the IEA earn a lower payoff than other members of their population outside the IEA. While the payoff is greater than when there is no IEA participation, each signatory would prefer to be a non-signatory. Second, the "lynchpin" equilibrium is not robust to a tremble, intentional or otherwise, by a signatory. If one IEA member chooses pollute rather than abate, thus earning a higher payoff, the agreement collapses. Clearly, the IEA is not very robust.

The main point of Barrett (2001) is to show how transfers (side-payments) among signatories can improve the IEA. A remarkable equilibrium of full participation by low benefit nations and greater participation by high benefit nations can be obtained when "cooperation is for sale." High benefit nations "purchase" the cooperation of low benefit nations via the transfer. With strong asymmetry quite substantial gains can be achieved, however, participation and compliance issues remain. International cooperation may be for sale, but no one wants to be the one buying it.

This paper allows for all three types of linear evolutionary games; Prisoner's Dilemma, Hawk-Dove and Coordination (Weibull 1995), and allows for increasing, decreasing or constant returns to abatement. All three classes of games maintain the fundamental characteristics of international public goods provision. Specifically, the aggregate payoff is monotonically increasing in the number of nations that abate, and there are positive externalities from abatement that accrue to both those who abate and those who pollute. Unlike Barrett (1999, 2001),

this paper models the IEA as an evolutionary game, and allows the externalities to differ both within and across populations.

This paper shows that transfers are not a credible way to increase participation in a single population evolutionary game. A two population framework is presented to allow for asymmetry. Nations differ in the externalities their abatement generates, and the cost of abatement.[1] Both Hawk-Dove (decreasing returns to abatement) and Coordination (increasing returns to abatement) games result in an interior Nash equilibrium. However, only when there is a Hawk-Dove with small asymmetry is the interior Nash an evolutionary equilibrium (EE). Essentially, an EE is an Nash equilibrium refinement requiring dynamic stability (Friedman 1998). With asymmetry, transfers can be credibly made in the two population model. At the interior EE the optimal transfer drives the system to a corner solution where all nations are better off. However, this credible transfer is made from the "poor" to the "rich" increasing payoff inequality across populations. For a Hawk-Dove game with large asymmetry credible transfers can eliminate the basin of attraction for the Pareto inferior EE. Both types of transfers increase each nation's payoff and can be agreed to with unanimity.

Evolutionary games have been used to model both international and environmental issues (Friedman and Fung 1996; Fisher and Kakkar 2004; Dijkstra and de Vries 2006), however, this paper is the first evolutionary game IEA. Evolutionary games contain dynamic adjustment, but differ from repeated games in important ways. In an evolutionary game players respond to current payoff differentials, and the strategy selection process ensures that actions with a payoff advantage increase in prevalence. Evolutionary games are not explicitly forward looking. By contrast, a repeated game strategy is a complete contingency plan (Mas-Colell et al. 1995) laid out before play begins. A strategy specifies an action to be taken at each point in time for all possible realizations of the history of play. Repeated games often rely on the threat of future punishments (i.e. trigger strategies) to maintain cooperation, where evolutionary games rely on actual punishment.

Most importantly for the present paper, evolutionary stability requires an equilibrium to be robust to a tremble by a current player, or to an equilibrium entrant. Players choose actions that are a best-response to the current distribution of the population. An EE is a simultaneous best-response by all players, and thus is a Nash equilibrium refinement based on dynamic stability. This differs from the classic evolutionary game equilibrium concept of an evolutionary stable strategy (ESS), due to Maynard Smith (1974). An ESS is a simultaneous best-response given situations where two players are randomly drawn from the population and "pairwise matched." Pairwise matching fits well in situations where players interact with a random draw, however, in many economic applications (i.e., firms in markets) each player interacts with all other players. An EE differs from an ESS in the scope of the best-response. Each player is matched against the entire (finite) population in an EE, and all players are best-responding. The finite population of players gradually adopt strategies with a payoff advantage.

In an evolutionary game model of IEAs, the EE are robust to trembles (i.e., a player choosing a different action) and all members of a population earn the same payoff. No signatory would prefer to be outside the IEA. Evolutionary stability may eliminate some Nash equilibria leaving only those that are robust, and do not rely on future punishment phases to maintain cooperation. The framework allows for a sequential analysis from any given level of participation.

Transfers are zero-sum and generated only by the underlying payoff functions. Previous literature (Folmer et al. 1993; Folmer and van Mouche 1994) has shown that transfers

---

[1] An example of an asymmetric IEA is the Kyoto Treaty, where members face different costs of participation and abatement requirements.

generated by linking two issues can dramatically improve the outcome of either game in isolation. This paper considers two populations, but only one game, thus issue linkage is not considered. Transfers are only considered credible if the post-transfer payoff does not put that player at a disadvantage relative to other members of their population.

Finally, the linear model presented in this paper is clearly a restrictive simplifying assumption. The theoretical IEA literature typically assumes decreasing returns, since abatement is a global public good (Finus 2003). Empirical work is consistent with this notion of declining marginal benefit and increasing marginal cost of abatement (Nordhaus and Yang 1996). However, it is possible that the returns vary with the level. Abatement may have increasing returns at low levels and decreasing returns beyond a threshold.

The remainder of the paper is organized as follows. The single population model is presented in Sect. 2, detailing the payoff functions, externalities, and the role of transfers. Section 3 presents the two population model and derives conditions under which an interior Nash equilibrium is obtained, and when it can be maintained as an EE. Section 4 details the use of credible transfers to increase global payoff. Section 5 concludes and discusses policy implications.

2 Model and Classification

Following Barrett (1999, 2001, 2003) the most simple public goods model from the IEA literature is considered. Barrett (1999, 2001) presents a linear Prisoner's Dilemma with n nations and models the IEA as a stage game. The payoff to a nation depends on its own abatement (q_i) and abatement by other players $Q_{-i} = \sum_{j \neq i}^{n} q_j$. For simplicity, abatement is a binary choice: $q_i = \{0, 1\}$, where q_i is zero if a nation chooses pollute and one if it chooses abate. The individual payoffs from choosing pollute (π_p) and abate (π_a) are:

$$\pi_p = b(Q_{-i} + q_i)$$
$$\pi_a = d(Q_{-i} + q_i) - cq_i \tag{1}$$

The parameter c is the cost of abatement. Denote the proportion of nations (n) choosing abate (a) as $s \in [0, 1]$, with the remaining $(1 - s)n$ nations choosing pollute (p).[2] With this notation the payoff functions (1) become:

$$\pi_p = bsn$$
$$\pi_a = -c + dsn \tag{2}$$

The parameter b is the marginal benefit to a polluter from each nation that abates, d is the marginal benefit to a nation that abates from their own and others' abatement. The sum of the marginal benefits (MB) from abatement is: $b(1 - s)n + dsn$ and the rate of change is: $\frac{\partial \sum_i^n MB_i}{\partial s} = n(d - b)$. If $d > b$ ($d < b$) then there are increasing (decreasing) returns to abatement in s. Barrett (1999) assumes $d \geq b$, while $d = b$ in Barrett (2001). The payoff is normalized to zero if no nation chooses abate. Extending Barrett's framework to allow for different returns to abatement (b relative to d) yields three types of evolutionary games.

Three assumptions that are maintained for all three types of games are positive externalities from abatement, the social optimum is when all nations abate and the gains to cooperation are positive. Positive externalities are insured if $b > 0$ and $d > 0$ since $\frac{\partial \pi_p}{\partial s} = bn > 0$ and $\frac{\partial \pi_a}{\partial s} = dn > 0$. In this linear framework, full participation is socially optimal when the

---

[2] Integer problems are ignored for expositional clarity, allowing the results to be expressed in s.

aggregate payoff $\Pi \equiv sn\pi_a + (1 - s)n\pi_p$ is monotonically increasing in s.

$$\Pi = ds^2n^2 - csn + bs(1 - s)n^2$$
$$\frac{\partial \Pi}{\partial s} = 2sn^2(d - b) - cn + bn^2 \tag{3}$$

The condition $\frac{\partial \Pi}{\partial s} > 0$ for all $s \in [0, 1]$, is $c < 2sn(d - b) + bn$. Given the linear structure, the endpoints are sufficient to truncate the allowable ranges of the parameters. $\frac{\partial \Pi}{\partial s}\big|_{s=0}$ implies $c < bn$ and $\frac{\partial \Pi}{\partial s}\big|_{s=1}$ implies $c < n(2d - b)$. Finally, the aggregate gains from full participation $G(s)$ are required to be positive:

$$G(s) \equiv n\left[\pi_a(1) - \pi_p(0)\right] = n\left[-c + dn\right] \tag{4}$$

The gains are positive when the cost of abatement per nation is less than the global benefit of that abatement: $c < dn$. Together, these three criteria require:

$$\frac{c}{n} < \begin{cases} 2d - b \\ b > 0 \\ d > 0 \end{cases} \tag{5}$$

The sign of the parameter c, and the returns to abatement (relative magnitudes of b and d) determine the classification of the three types of games. A Prisoner's Dilemma may have any type of returns to abatement and still retain pollute as a dominant strategy Nash equilibrium. The Coordination game has increasing returns $d > b$, and a Hawk-Dove has decreasing returns $b > d$. Both Coordination and Hawk-Dove generate a mixed strategy Nash equilibrium. Define the payoff advantage to abate as:

$$\gamma(s) \equiv \pi_a - \pi_p = -c + sn(d - b) \tag{6}$$

Solving $\gamma(s) = 0$ yields the interior Nash equilibrium s^*, if one exists.

$$s^* = \frac{c}{n(d - b)} \tag{7}$$

The dynamics of the state variable s are given by standard evolutionary game theory. Actions with a payoff advantage become more prevalent, as nations adjust their actions to increase their payoffs. This adjustment reflects "survival of the fittest" strategy, with some inertia. The dynamics follow a continuous time deterministic framework, $\dot{s} \equiv ds/dt = H(s, x)$, where H is a dynamic written as a function of the state s and a vector x of the remaining parameters (b, c, d, n). The dynamics are compatible (Friedman 1991) since abate displaces pollute when abate yields a higher payoff: $\dot{s} \lessgtr 0 \iff \gamma(s) \lessgtr 0$. The Nash equilibrium occurs when $\gamma(s) = 0$, or when $s = \{0, 1\}$ if $\gamma(s) \neq 0$.

An EE is a stable Nash equilibrium. Stability may be broadly interpreted as robustness to an equilibrium entrant, a mutation, or trembling-hand perfection. Stability of the interior Nash equilibrium s^* is determined by the slope of $\gamma(s)$. The slope, $\frac{\partial \gamma(s)}{\partial s} = n(d - b)$, is positive if there are increasing returns to abatement $d > b$ (Coordination game) and negative if there are decreasing returns $d < b$ (Hawk-Dove). The Prisoner's Dilemma may have any slope, but requires $\gamma(s) < 0 \; \forall s \in [0, 1]$. Evaluating the endpoints, $\gamma(0) = -c$ and $\gamma(1) = -c + n(d - b)$, allows a classification of the three types of games (Table 1).

The types of IEAs that can be sustained as evolutionary equilibria are as follows.

Table 1 Single population classification

Type	$\gamma(0) = -c$	$\gamma(1) = -c + n(d-b)$	$\frac{\partial \gamma(s)}{\partial s}$
(1) Prisoner's dilemma	$\gamma(0) < 0 \Rightarrow c > 0$	$\gamma(1) < 0 \Rightarrow \frac{c}{n} > d - b$	$\frac{\partial \gamma(s)}{\partial s} \gtreqless 0 \Rightarrow d \gtreqless b$
(2) Coordination	$\gamma(0) < 0 \Rightarrow c > 0$	$\gamma(1) > 0 \Rightarrow \frac{c}{n} < d - b$	$\frac{\partial \gamma(s)}{\partial s} > 0 \Rightarrow d > b$
(3) Hawk-Dove	$\gamma(0) > 0 \Rightarrow c < 0$	$\gamma(1) < 0 \Rightarrow \frac{c}{n} > d - b$	$\frac{\partial \gamma(s)}{\partial s} < 0 \Rightarrow d < b$

Proposition 1 *The IEA evolutionary equilibria are*: (i) *no participation for a Prisoners Dilemma,* (ii) *both full and no participation for Coordination game with basins of attraction separated by s* and* (iii) *partial participation given by s* for a Hawk-Dove game.*

Proof (i) For a Prisoner's Dilemma $\gamma(s) < 0 \; \forall s \in [0, 1]$, thus any initial condition $s_0 \in (0, 1]$ will converge to $s = 0$ by compatibility of the dynamic H, i.e., $\dot{s} \lessgtr 0 \iff \gamma(s) \lessgtr 0$. (ii) There are three Nash equilibria for a Coordination game, two of which are evolutionary equilibria: $s = 0$, and $s = 1$. The interior Nash equilibrium s^* is unstable. At s^* consider a tremble by a nation intending to abate but choosing pollute. Then $s = s^* - \epsilon$, where $\epsilon \equiv \frac{1}{n}$, and $\frac{\partial \gamma(s)}{\partial s} > 0$ implies the EE is $s = 0$. Similarly, a tremble by a nation intending to pollute implies $s = s^* + \epsilon$ and $\frac{\partial \gamma(s)}{\partial s} > 0$ results in the EE $s = 1$. (iii) The unique EE for a Hawk-Dove game is s^*. Any initial condition $s_0 \in [0, 1]$ will converge to s^* by compatibility of the dynamic H given $\frac{\partial \gamma(s)}{\partial s} < 0$. □

A Prisoner's Dilemma may have any type of returns ($d \gtreqless b$), and retain pollute as a dominant strategy. The EE $s = 0$ in Fig. 1b is labeled with a solid circle.

A Coordination game occurs when $\gamma(0) < 0$ and $\gamma(1) > 0$. This implies the restrictions: $c > 0$ and $c < n(d - b)$, or increasing returns to abatement. The interior (mixed) strategy Nash equilibrium is not stable, leaving the two pure strategy Nash as EE. By monotonicity, full participation is the payoff dominant equilibrium. In Fig. 2a the risk-dominant equilibrium ($s = 0$) has a larger basin of attraction. In this case there needs to be a large "threshold" level of cooperation for the IEA to obtain full participation. One example of a threshold IEA is the Kyoto Treaty which specified a minimum level of participation before the treaty entered into force.

With a Hawk-Dove game there are gains to IEA participation when participation is low but, with decreasing returns to abatement, the gains quickly turn negative. The vast majority of the IEA literature fits this situation. For example, the seminal work of Barrett (1994) finds the IEA consists of three out of 100 nations when the gains to cooperation are large (Fig. 3).

Next, we turn to the issue of credible transfers in a single population game.

Proposition 2 *Transfers cannot increase IEA abatement in a single population evolutionary equilibrium.*

Proof (i) A Prisoner's Dilemma is defined as: $\gamma(s) < 0 \; \forall s \in [0, 1]$, with EE $s^* = 0$. Playing a and accepting the transfer τ is a best-response if $\gamma(s) + \tau \geq 0$. Thus, the minimum transfer is: $\tau^{\min}(s) = -\gamma(s) = c + sn(b - d) > 0 \; \forall s \in [0, 1]$. $\tau^{\min}(s)$ is a credible transfer by the $sn - 1$ other signatories when $\gamma(s) - \frac{t^{\min}(s)}{(sn-1)} \geq 0$. Therefore, the transfer is not credible since $\gamma(s) - \frac{t^{\min}(s)}{(sn-1)} < 0 \; \forall s \in [0, 1]$. (ii) Both Coordination and Hawk-Dove games have an interior Nash equilibrium $s^* = \frac{c}{n(d-b)}$ with payoffs: $\pi_a(s^*) = \pi_p(s^*) = \frac{bc}{d-b}$. With an additional signatory the state becomes $s^* + \epsilon$, where $\epsilon = \frac{1}{n}$ and the payoff advantage to
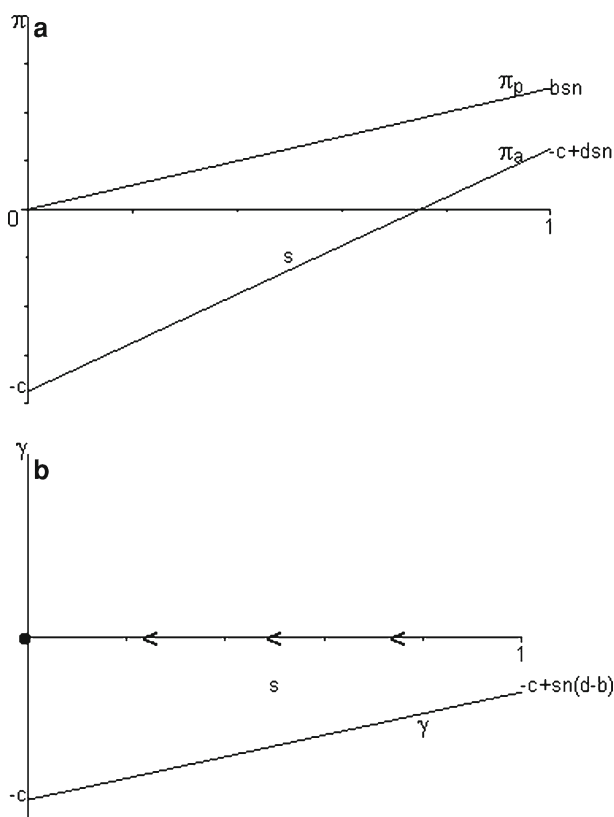
Fig. 1 **a** Prisoner's dilemma payoffs; **b** Prisoner's dilemma payoff difference

participation is: $\gamma(s^* + \epsilon) = d - b$. In a Coordination game $d - b > 0$ and no transfer is needed to increase participation. Any $s < s^*$ is in the $s = 0$ EE basin of attraction. With unanimity the Pareto dominant EE $s = 1$ can be agreed upon without transfers. However, there is no credible transfer in the $s = 0$ EE basin since $\gamma(s + \epsilon) - \frac{\tau^{\min}(s)}{sn} < 0 \ \forall s < s^*$. (iii) If the game is Hawk-Dove then $d - b < 0$ and the minimum transfer such that IEA participation is individually rational is: $\tau^{\min}(s^*) = -\gamma(s^* + \epsilon) = b - d > 0$. For the transfer to be credible: $\gamma(s^* + \epsilon) - \frac{\tau^{\min}(s^*)}{s^* n} = d - b + \frac{(b-d)^2}{c} > 0$. However, this is clearly negative since $b > d$, and $c < 0$. □

The difference with the transformed Prisoner's Dilemma game in Barrett (1999, 2001) is that in the evolutionary equilibrium *all* nations earn the same payoff, making a transfer a non-credible commitment. In Barrett's stage game the signatories to an IEA move first and choose their actions to maximize their collective payoff. Even so, in equilibrium non-signatories earn a higher payoff than signatories.

3 Two Population Model

The two population model considers externalities that differ both within and across populations. Both the externalities generated by participation, and the cost, depend on which type of
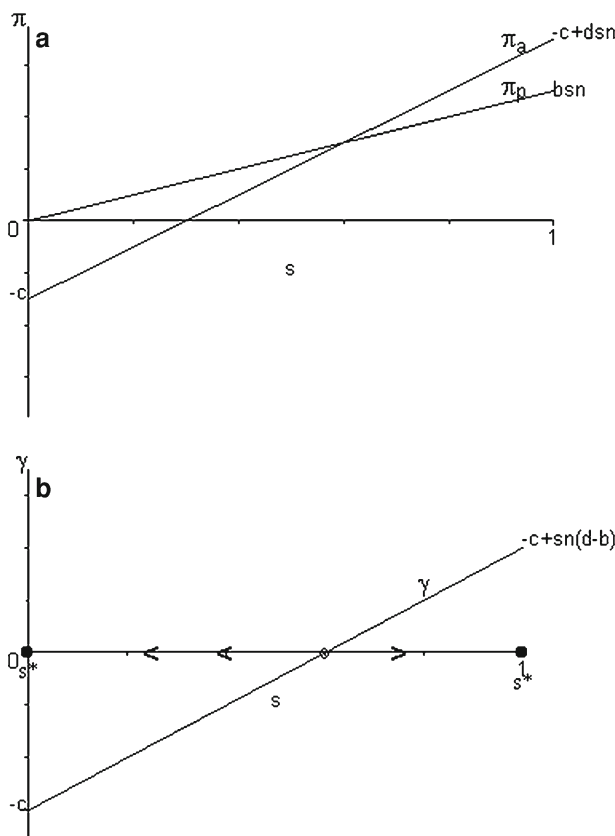
Fig. 2 **a** Coordination game payoffs; **b** coordination game payoff difference

nation joins the IEA, thus $d_1 \neq d_2$ and $c_1 \neq c_2$. The outside payoff also varies by population, hence $b_1 \neq b_2$. Defining abatement by population i as Q_i, Eq. (1) becomes:[3]

$$\pi_i^p = b_i(Q_1 + Q_2), i = 1, 2$$
$$\pi_i^a = d_1 Q_1 + d_2 Q_2 - c_i q_i, i = 1, 2 \qquad (8)$$

Under the Kyoto Treaty only the developed nations (Annex I) are subject to binding abatement requirements. Therefore, their participation generates a greater externality than a non-Annex I nations. Denote the Annex I nations as population 1 with the proportion choosing abate as s_1. Non-Annex I nations are population 2 with proportion s_2 choosing abate. Let ϕ be the proportion of nations in population 1, and n the total number of nations, so $n_1 \equiv \phi n, n_2 \equiv (1 - \phi)n$. Abatement is again a binary choice $q_i = \{0, 1\}$, so the two-population payoffs are:

$$\pi_i^p(s_1, s_2) = b_i (s_1 n_1 + s_2 n_2), i = 1, 2$$
$$\pi_i^a(s_1, s_2) = d_1 s_1 n_1 + d_2 s_2 n_2 - c_i, i = 1, 2 \qquad (9)$$

---

[3] This is the most simple specification that allows for the relevant asymmetry across populations. The alternative specification: $\pi_i^p = b_1 Q_1 + b_2 Q_2, i = 1, 2$, does not allow the pollute payoff to differ across populations. Similarly, the specification $\pi_i^a = d_i (Q_1 + Q_2), i = 1, 2$ does not allow the externality generated by participation to differ across population. In either case, the γ_i functions (10) are linearly dependent, so the payoff advantages to abate are identical up to a scalar.
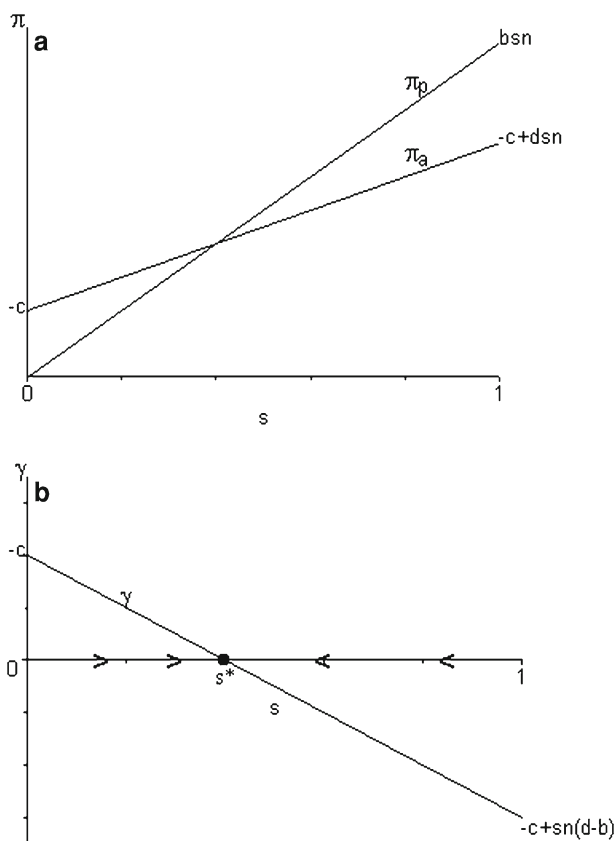
Fig. 3 a Hawk-Dove payoffs; **b** Hawk-Dove payoff difference

The sum of the marginal benefits from abate is: $b_1(1-s_1)n_1+b_2(1-s_2)n_2+d_1s_1n_1+d_2s_2n_2$ and returns to cooperation are: $\frac{\partial \sum_i^n MB_i}{\partial s_1} = n_1(d_1 - b_1)$ and $\frac{\partial \sum_i^n MB_i}{\partial s_2} = n_2(d_2 - b_2)$. Four possible externality orders are considered, but the returns to cooperation are maintained across populations. With increasing returns to cooperation a Prisoner's Dilemma or a Coordination game results. Increasing returns and small asymmetry across populations yields: $b_2 < b_1 < d_2 < d_1$, and with large asymmetry $b_1 < b_2 < d_2 < d_1$. If there are decreasing returns to abatement then a Prisoner's Dilemma or Hawk-Dove game with small $d_2 < d_1 < b_2 < b_1$ or large asymmetry $d_1 < d_2 < b_2 < b_1$ is obtained.

Parameter restrictions ensure that the fundamental aspects of IEAs remain intact. There are positive externalities, global payoff is strictly increasing in participation, and the gains to cooperation are positive. These conditions are a straightforward generalization of the single population case and are given in the Appendix. The two population payoff advantage to abate functions $\gamma_i \equiv \pi_i^a(s_1, s_2) - \pi_i^p(s_1, s_2)$ are:

$$\gamma_i = s_1n_1(d_1 - b_i) + s_2n_2(d_2 - b_i) - c_i \tag{10}$$

Solving for the $\gamma_i = 0$ loci we get s_2 as an implicit function of s_1:

$$\gamma_i = 0 \Rightarrow s_2(s_1) = \frac{s_1n_1(d_1 - b_i) - c_i}{n_2(b_i - d_2)} \tag{11}$$

The $\gamma_i = 0$ loci have slope: $\frac{\partial s_2(s_1)}{\partial s_1} = \frac{n_1(d_1-b_i)}{n_2(b_i-d_2)}$, vertical intercept: $s_2(0) = \frac{-c_i}{n_2(b_i-d_2)}$ and horizontal intercept $s_1(0) = \frac{-c_i}{n_1(b_i-d_1)}$. Both loci have a negative slope in Hawk-Dove and Coordination games when the returns to abatement is preserved across, as well as within, populations.

3.1 Interior Nash Equilibrium

An interior equilibrium occurs when there is partial participation in both populations. All other situations result in corner equilibria. The simultaneous solution to the $\gamma_i = 0$ loci, valid at an interior Nash equilibrium is:

$$s_1^* = \frac{c_1 (d_2 - b_2) + c_2 (b_1 - d_2)}{n_1 (b_1 - b_2) (d_1 - d_2)}$$
$$s_2^* = \frac{c_1 (b_2 - d_1) + c_2 (d_1 - b_1)}{n_2 (b_1 - b_2) (d_1 - d_2)} \tag{12}$$

The payoffs at the interior Nash equilibrium are:

$$\pi_1^a \left(s_1^*, s_2^*\right) = \pi_1^p \left(s_1^*, s_2^*\right) = \frac{b_1(-c_1 + c_2)}{b_1 - b_2}$$
$$\pi_2^a \left(s_1^*, s_2^*\right) = \pi_2^p \left(s_1^*, s_2^*\right) = \frac{b_2(-c_1 + c_2)}{b_1 - b_2} \tag{13}$$

This leads directly to Proposition 3.

Proposition 3 *The interior Nash equilibrium payoffs are increasing in the difference in abatement costs and decreasing in the difference in the pollute externality. The population with the larger pollute externality receives a higher payoff.*

Proof Equation (13) and Table 3 in the Appendix. □

Proposition 3 indicates that it is the pollute externality that determines payoff, not the externality generated by choosing abate. This is due to payoffs being equalized across actions (*a* or *p*) in the EE. Table 3 details the parameter restrictions necessary for an interior equilibrium.

Proposition 4 *There exists a unique interior Nash equilibrium when the ratio of abatement costs (c_1/c_2) is contained in the ratio of the externality differences, and each population is sufficiently large.*

Proof Table 3 details the parameter restrictions required for an interior Nash Equilibrium. The table shows the restrictions such that both s_1 and s_2 are positive fractions. □

The global payoff at an interior equilibrium is:

$$\Pi \left(s_1^*, s_2^*\right) = \frac{(n_1 b_1 + n_2 b_2)(-c_1 + c_2)}{b_1 - b_2} \tag{14}$$

The two-population dynamics are a straightforward generalization of the single population case. The adjustment dynamics are $\dot{s}_i \equiv ds_i/dt = H(s_1, s_2, x)$, where H is function of the state and x is a vector of parameters $(b_1, b_2, c_1, c_2, d_1, d_2, \phi, n)$. Compatibility of H in terms of the payoffs requires: $\dot{s}_i \lessgtr 0 \iff \gamma_i(s) \lessgtr 0$.

Table 2 Evolutionary stability

	Coordination: $b_i < d_j, i, j = 1, 2$		Hawk-Dove: $b_i > d_j, i, j = 1, 2$	
	(i) Small asym	(ii) Large asym	(i) Small asym	(ii) Large asym
	$b_2 < b_1 < d_2 < d_1$	$b_1 < b_2 < d_2 < d_1$	$d_2 < d_1 < b_2 < b_1$	$d_1 < d_2 < b_2 < b_1$
Trace J	+	+	−	−
Determinant J	+	−	+	−
Eigenvalues	+, +	+, −	−, −	−, +

The eigenvalues determine evolutionary stability of an interior Nash. The first-order partials of γ_i form the Jacobian (J):

$$J = \begin{bmatrix} \frac{\partial \gamma_1}{\partial s_1} & \frac{\partial \gamma_1}{\partial s_2} \\ \frac{\partial \gamma_2}{\partial s_1} & \frac{\partial \gamma_2}{\partial s_2} \end{bmatrix} = \begin{bmatrix} n_1(d_1 - b_1) & n_2(d_2 - b_1) \\ n_1(d_1 - b_2) & n_2(d_2 - b_2) \end{bmatrix} \tag{15}$$

The trace is the sum of the eigenvalues: *Trace J* $= n_1(d_1 - b_1) + n_2(d_2 - b_2)$, and the determinant is their product: *Determinant J* $= n_1 n_2 (b_1 - b_2)(d_1 - d_2)$. The trace is positive (negative) for increasing (decreasing) returns to abatement, and the determinant is positive (negative) for small (large) asymmetry. The system is dynamically stable for the Hawk-Dove game with small asymmetry and the interior Nash equilibrium is an EE. Otherwise, the EE occurs along an edge of the unit-simplex. Table 2 summarizes these results.

The stability of the unique interior Nash equilibrium is characterized as follows:

Proposition 5 *The Evolutionary Equilibrium is partial participation in both populations for a Hawk-Dove game* (*decreasing returns to abatement*) *with small asymmetry. A Hawk-Dove game with large asymmetry or a Coordination game* (*increasing returns to abatement*) *result in an EE along an edge of the unit-simplex.*

Proof The Nash equilibrium (12) is a global attractor when both eigenvalues are negative making it the unique EE. The Nash equilibrium (12) is a source otherwise, resulting in a EE along an edge of the unit-simplex. □

To illustrate evolutionary stability consider an example similar to the Kyoto Treaty satisfying the parameter restrictions in the Appendix. The example is an interior Nash equilibrium with decreasing returns to abatement for both small and large asymmetry. Let $n = 100$ nations and $\phi = 0.4$ since there are approximately 40 Annex I nations subject to abatement requirements. For small asymmetry the parameters are: $-c_1 = 160, -c_2 = 100, d_2 = 1, d_1 = 2, b_2 = 3, b_1 = 4$. The EE is: $s_1^* = 0.50, s_2^* = 0.67$, with payoffs $\pi_1^a(s_1^*, s_2^*) = \pi_p^1(s_1^*, s_2^*) = 240, \pi_2^a(s_1^*, s_2^*) = \pi_2^P(s_1^*, s_2^*) = 180$. Global payoff is: 20,400, or 77% of the full participation payoff of 26,400 from $\pi_1^a(1, 1) = 300, \pi_2^a(1, 1) = 240$. There are potential gains of 60 to all players from achieving full participation from the EE. Fig. 4 shows the unique, interior EE.

However, if we reverse the values of d_1 and d_2 and increase $-c_1 = 170$ (at $-c_1 = 160$ the EE is not interior) asymmetry is large, changing the sign of the determinant. The interior Nash equilibrium is unstable, resulting in two EE: {0.42,1} and {1,0.33}. Figure 5 illustrates the large asymmetry case.

Payoffs at {0.42,1} are: $\pi_1^P = \pi_1^a = 306.7, \pi_2^a = 236.7$, and at {1,0.33} payoffs are: $\pi_1^P = 250, \pi_2^a = \pi_2^P = 180$. Clearly, all players prefer {0.42,1}. Global payoff is: 26,467 at
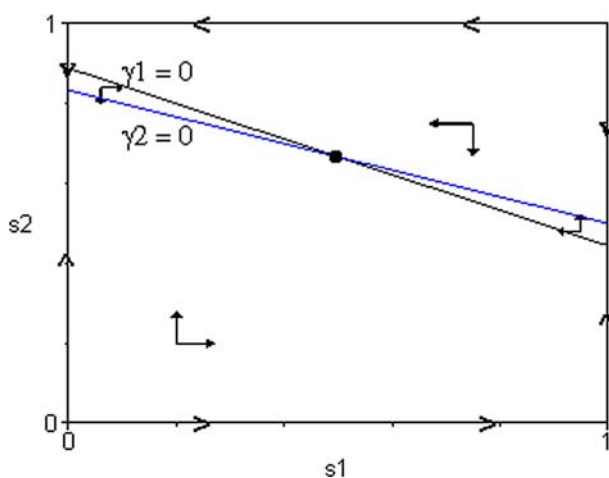
Fig. 4 Hawk-Dove small asymmetry
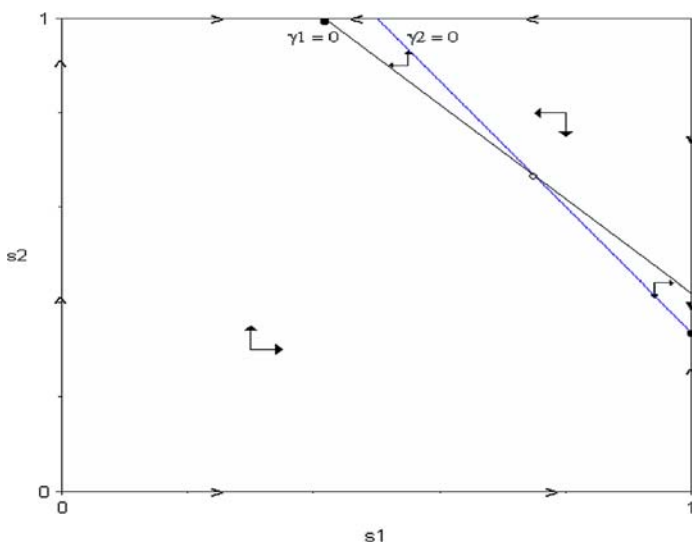
Fig. 5 Hawk-Dove large asymmetry

$\{0.42,1\}$ and 20,800 at $\{1,0.33\}$, compared to the full participation payoff $\{1,1\}$ of 28,800 with $\pi_1^a = 330$, $\pi_2^a = 260$. A Hawk-Dove game has become fundamentally the same as a Coordination game. With unanimity all nations can agree that $\{0.42,1\}$, obtaining 92% of the full participation payoff, is superior to $\{1,0.33\}$ and 72% of the full participation payoff. No transfers are needed for this coordination, however, as the next section shows, transfers can ensure the payoff dominant EE is obtained by eliminating the basin of attraction for the inferior EE.
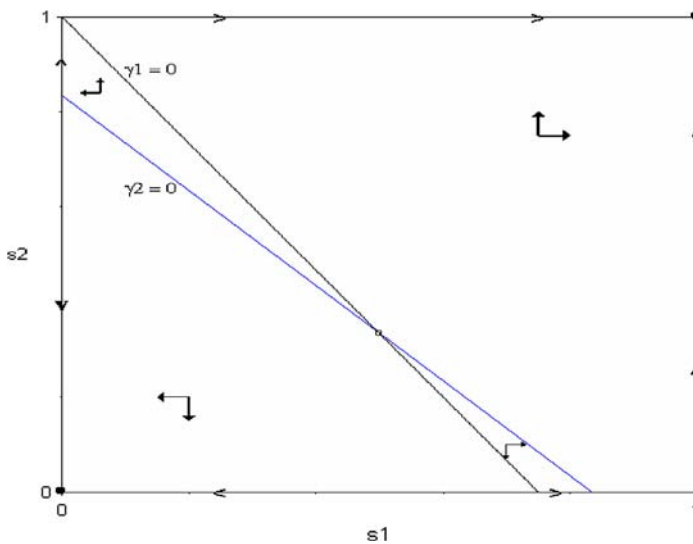
Fig. 6 Coordination game

3.2 Corner Solutions

For all other permissible parameter constellations the EE is along an edge of the unit-simplex. Parameters that satisfy the fundamental characteristics of positive externalities, aggregate payoff monotonically increasing in participation and gains to cooperation, but do not have an interior Nash equilibrium will result in EE along an edge. As an example, consider the Coordination game in Fig. 6: $c_1 = 60, c_2 = 100, b_2 = 1, b_1 = 2, d_2 = 3, d_1 = 4$. The interior Nash equilibrium {0.5,0.33} is unstable. There are two EE at: {0,0} and {1,1}. The full participation IEA can be sustained without transfers as long as the threshold level of participation is above the lower locus.

4 Transfers

Previous work has shown that transfers among signatories have the potential to dramatically increase IEA participation (Barrett 2001, 2003). At an interior EE a transfer to a non-signatory within population is not credible since they have equal payoffs. No signatory could credibly give a transfer to a free-rider since the signatory would earn less post-transfer than a non-signatory. However, transfers may be credible across populations. Credible transfers may decrease participation in one population, increase participation in the other population and increase all nations' payoff. The optimal transfer is that which generates the highest global payoff from the set of post-transfer EE. Optimal transfers eliminate the interior EE when it exists, and eliminate the basin of attraction for the Pareto inferior EE.

4.1 Interior EE

An interior EE is obtained only in a Hawk-Dove game with small asymmetry. Proposition 5 states payoffs are only equated within, not across, populations. This raises the question

if a zero-sum system of transfers among signatories can increase participation and global welfare. With small asymmetry ($d_1 > d_2$) a type 1 signatory adds more to the global payoff, but they are harder to get to join since they have a greater free-rider incentive to overcome $b_1 > b_2$.

The γ_i^τ function is the post-transfer payoff advantage to cooperating: $\gamma_i^\tau \equiv \gamma_i + \tau_i$.

$$\gamma_i^\tau = d_1 s_1 n_1 + d_2 s_2 n_2 - b_i [s_1 n_1 + s_2 n_2] - c_i + \tau_i, i = 1, 2 \tag{16}$$

The $\gamma_i^\tau = 0$ loci, again solved for s_2 (s_1) are:

$$s_2(s_1) = \frac{s_1 n_1 (d_1 - b_i) - c_i + \tau_i}{n_2 (b_i - d_2)} \tag{17}$$

So a positive (negative) transfer shifts up (down) the locus, but does not change the slope of the loci, or effect the stability properties of the EE (sign of the eigenvalues). The interior EE is:

$$
\begin{aligned}
s_1^{*\tau} &= \frac{(c_1 - \tau_1)(d_2 - b_2) + (c_2 - \tau_2)(b_1 - d_2)}{n_1 (b_1 - b_2)(d_1 - d_2)} \\
s_2^{*\tau} &= \frac{(c_1 - \tau_1)(b_2 - d_1) + (c_2 - \tau_2)(d_1 - b_1)}{n_2 (b_1 - b_2)(d_1 - d_2)}
\end{aligned}
\tag{18}
$$

Payoffs at an interior EE, post-transfer are:

$$
\begin{aligned}
\pi_1^a \left(s_1^{*\tau}, s_2^{*\tau}\right) &= \pi_1^P \left(s_1^{*\tau}, s_2^{*\tau}\right) = \frac{b_1(-c_1 + \tau_1 + c_2 - \tau_2)}{b_1 - b_2} \\
\pi_2^a \left(s_1^{*\tau}, s_2^{*\tau}\right) &= \pi_2^P (s_1^{*\tau}, s_2^{*\tau}) = \frac{b_2(-c_1 + \tau_1 + c_2 - \tau_2)}{b_1 - b_2}
\end{aligned}
\tag{19}
$$

Of course, the transfers are zero-sum and satisfy:

$$s_1^{*\tau} n_1 \tau_1 + s_2^{*\tau} n_2 \tau_2 = 0 \tag{20}$$

The optimal transfer results in the highest global payoff that can be sustained as an EE.

Proposition 6 *With an interior EE (Hawk-Dove, small asymmetry) the optimal transfer shifts the loci so their intersection is an EE at the edge of the unit-simplex, exacerbating inequality to achieve a higher payoff for all nations.*

Proof A Hawk-Dove game with small asymmetry is defined as: $c_1 < 0, c_2 < 0, d_2 < d_1 < b_2 < b_1$ and the interior Nash equilibrium restriction in Table 3 requires $|c_1| > |c_2|$. At an interior EE global payoff is: $\Pi(s_1^{*\tau}, s_2^{*\tau}) = \frac{(n_1 b_1 + n_2 b_2)(-c_1 + \tau_1 + c_2 - \tau_2)}{b_1 - b_2}$. Given $b_2 < b_1$, $\frac{d\Pi}{d\tau_1} > 0$ and $\frac{d\Pi}{d\tau_2} < 0$, thus global payoff is increased by a transfer from population 2 to population 1: $\tau_1 > 0, \tau_2 < 0$. Payoff inequality without and with transfers is: $\pi_1(s_1^*, s_2^*) - \pi_2(s_1^*, s_2^*) = -c_1 + c_2$, and $\pi_1(s_1^{*\tau}, s_2^{*\tau}) - \pi_2(s_1^{*\tau}, s_2^{*\tau}) = -c_1 + \tau_1 + c_2 - \tau_2$, respectively. Clearly, $\tau_1 > 0, \tau_2 < 0$ increases inequality. □

Consider the Hawk-Dove small asymmetry example in Fig. 4: $-c_1 = 160, -c_2 = 100, d_2 = 1, d_1 = 2, b_2 = 3, b_1 = 4$. The NE is: $s_1^* = 0.50, s_2^* = 0.67$. The payoffs at the EE are: $\pi_1^a (s_1^*, s_2^*) = \pi_1^P(s_1^*, s_2^*) = 240, \pi_2^a (s_1^*, s_2^*) = \pi_2^P (s_1^*, s_2^*) = 180$. So, it would seem that a transfer from population 1 (the rich) to population 2 (the poor) would be in order since population 1 has a higher payoff. However, each type 1 that joins the IEA adds more to the agreement since $-c_1 > -c_2$ and $d_1 > d_2$. So, we are in the difficult situation where the lower payoff group makes a positive transfer. Given the linearity of the payoff

functions, the optimal transfers are such that the loci are shifted until they just intersect at the $s_1 = 1$ edge of the unit-simplex in Fig. 4.

The optimal transfer pair for this example is: $\tau_1 = 3.16$, $\tau_2 = -4.56$, generating an EE of: $s_1 = 1$, $s_2^* = 0.46$. The optimal transfer pair shifts the γ_1^τ locus up and the γ_2^τ locus down in Fig. 4 until the loci intersect at the $s_1 = 1$ edge of the unit-simplex. Any increase in the transfer from this point will only reduce participation in population 2, and is therefore not optimal. With the transfers payoffs are: $\pi_1^a = 271$, $\pi_2^a = \pi_2^P = 203$, so all nations are better off, but transfers have increased the difference in payoffs across populations. Greater efficiency requires greater inequality in this case. The transfers are effective in closing the potential payoff gap. Global payoff is 23,025, or 47% of the difference between the no-transfer EE and the full participation payoff level. There are now $40 + 27.6 = 67.6$ nations cooperating compared to $20 + 40.2 = 60.2$ without transfers.

4.2 Corner EE

Transfers may also serve a purpose when there are two EE as in the Hawk-Dove large asymmetry in Fig. 5. The payoff dominant EE is along the $s_2 = 1$ edge since $d_2 > d_1$. The optimal transfer pair can eliminate the basin of attraction for the inferior EE (at the $s_1 = 1$ edge) by shifting the $\gamma_1^\tau = 0$ locus down and the $\gamma_2^\tau = 0$ locus up, until the loci intersect at the $s_1 = 1$ edge of the unit-simplex. The reduced form solution for this transfer pair involves solving a system of three Eqs. (17) and (20), evaluated at $s_1 = 1$. The three endogenous variables are s_2, τ_1 and τ_2. For example, using the $\gamma_1^\tau = 0$ locus and the zero-sum constraint to eliminate τ_1 results in an equation in s_2 and τ_2. Then using the $\gamma_2^\tau = 0$ locus to eliminate s_2 results in the reduced form transfer τ_2.

$$\tau_2 = \frac{1}{2}\left[n_1\left(b_2 - b_1 + d_2 - d_1\right) + c_2 + \sqrt{\rho}\right] > 0 \tag{21}$$

where $\rho \equiv n_1^2[(b_1 - b_2)^2 + (d_1 - d_2)^2 + 2(b_1 - b_2)(d_2 - d_1)] + 2n_1[-2c_1(b_2 - d_2) - c_2(b_1 + b_2 - d_1 - d_2)] + c_2^2$. Substituting (21) into the $\gamma_2^\tau = 0$ locus results in the reduced form s_2, and τ_1 follows from the zero-sum constraint.

Proposition 7 *For a Hawk-Dove game with large asymmetry the basin of attraction for the Pareto inferior EE can be eliminated by a transfer that reduces inequality. The transfer is credible since all players earn a higher payoff.*

Proof The Jacobian formed by (16) shows the stability properties are independent of the transfers. The reduced form transfers follow from direct substitution of Eq. (21) into the $\gamma_i^\tau = 0$ loci (17) and the zero-sum constraint (20). As in Proposition 6, the payoff difference across populations post-transfer is $\pi_1(s_1^{*\tau}, s_2^{*\tau}) - \pi_2(s_1^{*\tau}, s_2^{*\tau}) = -c_1 + \tau_1 + c_2 - \tau_2$ and $\tau_1 < 0$, $\tau_2 > 0$ decreases inequality. □

Consider the Fig. 5 Hawk-Dove large asymmetry example with two EE: {0.42,1} and {1,0.33}. Transfers from population 1 to population 2 can eliminate the basin of attraction of the inferior EE {1,0.33}. At {1,0.33} the optimal transfer is such that the loci intersect at the $s = 1$ edge of the unit-simplex. The transfer pair that accomplishes this is: $\tau_1 = -2.30$, $\tau_2 = 3.85$. However, the key difference is that the transfer need only be specified at the Pareto inferior EE to eliminate the basin of attraction. Once the basin of attraction is eliminated then the Pareto superior EE is obtained and no transfers are required to maintain the EE {0.42,1}. For a Hawk-Dove with large asymmetry the IEA only specifies a transfer in the event an inferior EE is obtained. Here the transfer serves as a coordination device and again it can be agreed on with unanimity.

5 Conclusion

Evolutionary game theory provides a natural framework to analyze IEAs. Governments are more likely to respond to the current relative payoffs from their action, with some inertia, than to lay out a complete contingency plan to be followed by subsequent administrations. Furthermore, it is difficult to justify being in an IEA when a nation can earn a higher payoff outside the IEA. This issue associated with a stage game analysis from previous work is eliminated in an evolutionary model.

This paper investigates IEAs as evolutionary games and shows that transfers cannot increase abatement in a single population framework. With two populations there is a unique interior EE in a Hawk-Dove game (decreasing returns to abatement) with small asymmetry in the externalities across populations. The population with the greater free-rider externality earns a higher payoff at an interior EE. For a Hawk-Dove with large asymmetry or a Coordination game the EE is along an edge of the unit-simplex.

Transfers can increase payoffs for all nations from an interior EE, but come at the expense of lower participation in one of the populations. Transfers can also eliminate the Pareto inferior EE in a Hawk-Dove with large asymmetry. Transfers have a limited role in Coordination games and can not be credibly made in a Prisoner's Dilemma. If there are decreasing returns to abatement (Hawk-Dove) then the post-transfer EE may have substantially more abatement than a single population model. This is a potential rationale for requiring only a subset of nations (Annex I) to abate in the Kyoto Treaty, effectively creating a two-population model.

A surprising policy implication emerges from the model. Both the "polluter-pays" and "ability-to-pay" principles dictate that rich nations are required to abate. This is one of the central forces behind the Kyoto Treaty where reducing global inequality is clearly a goal. If Annex I nations were allowed to meet their abatement requirements by purchasing tradable pollution permits from the non-Annex I nations, then actual abatement relative to the requirements under the treaty would implicitly define a transfer from the rich to the poor. However, with decreasing returns to abatement and small asymmetry all nations may be better off if the poor make a transfer to the rich. The rich bring more to the agreement, but have a greater incentive to remain outside. Overcoming their greater free-rider incentive is the reason that the rich nations receive a positive transfer. An effective IEA may be Pareto improving and increase global inequality. One possible way to overcome this increase in inequality is issue-linkage, discussed in the introduction. A multi-issue agreement has the potential to increase all payoffs and reduce inequality.

Appendix: Parameter Restrictions

Positive externalities requires: $\frac{\partial \pi_i^p}{\partial s_1} = b_i n_1 > 0$, $\frac{\partial \pi_i^p}{\partial s_2} = b_i n_2 > 0$, $\frac{\partial \pi_i^a}{\partial s_1} = d_1 n_1 > 0$ and $\frac{\partial \pi_i^a}{\partial s_2} = d_2 n_2 > 0$. Thus, $b_i > 0$ and $d_i > 0$ for $i = 1, 2$. Global payoff is:

$$\Pi(s_1, s_2) \equiv s_1 n_1 \pi_1^a + s_2 n_2 \pi_2^a + (1 - s_1) n_1 \pi_1^p + (1 - s_2) n_2 \pi_2^p \tag{22}$$

The derivatives are:

$$\frac{\partial \Pi(s_1, s_2)}{\partial s_1} = n_1 \left[-c_1 + 2 s_1 n_1 (d_1 - b_1) + s_2 n_2 (d_1 + d_2 - b_1 - b_2) + n_1 b_1 + n_2 b_2 \right]$$

$$\frac{\partial \Pi(s_1, s_2)}{\partial s_2} = n_2 \left[-c_2 + s_1 n_1 (d_1 + d_2 - b_1 - b_2) + 2 s_2 n_2 (d_2 - b_2) + n_1 b_1 + n_2 b_2 \right] \tag{23}$$

The derivatives in (23) are required to be positive for all points in the unit-simplex. (i) For a Prisoner's Dilemma or Coordination game ($d_i \geq b_i$) the minimum value of the partials occurs at $s_1 = s_2 = 0$.

$$c_i < n_1 b_1 + n_2 b_2 \text{ for } i = 1, 2 \tag{24}$$

(ii) For a Hawk-Dove (or a Prisoner's Dilemma) with decreasing returns ($d_i < b_i$) the minimum value occurs at $s_1 = s_2 = 1$.

$$c_i < n_i(2d_i - b_j) + n_j(d_1 + d_2 - b_i) \text{ for } i \neq j = 1, 2 \tag{25}$$

The global payoff with no cooperation is: $\Pi(s_1 = 0, s_2 = 0) = 0$, and with full cooperation is:

$$\Pi(s_1 = 1, s_2 = 1) = -n_1 c_1 - n_2 c_2 + d_1 (n_1)^2 + d_2 (n_2)^2 + n_1 n_2 (d_1 + d_2) \tag{26}$$

The parameter restriction for a positive global payoff difference is: $n_1 c_1 + n_2 c_2 < n_1 n_2 (d_1 + d_2) + d_1 (n_1)^2 + d_2 (n_2)^2$.

The second set of restrictions is given by the levels of γ_i at the diagonal corners ($s_1 = 0, s_2 = 0$) and ($s_1 = 1, s_2 = 1$). A Prisoner's Dilemma requires $\gamma_i(s_i, s_j)$ for all $s_i \in [0, 1]$. With decreasing returns all three terms in γ_i are negative so $\gamma_i(s_i, s_j) < 0 \; \forall s_i, s_j \in [0, 1]$ and there is no additional restriction required. With increasing returns the maximum value occurs at $\gamma_i(1, 1)$, yielding the restriction:

$$c_i > n_1(d_1 - b_i) + n_2(d_2 - b_i) \tag{27}$$

Together with (24), the increasing returns Prisoner's Dilemma restriction is:

$$n_1(d_1 - b_i) + n_2(d_2 - b_i) < c_i < n_1 b_1 + n_2 b_2 \tag{28}$$

The Coordination game has increasing returns and requires $\gamma_i(0, 0) < 0$ (or $c > 0$) and $\gamma_i(1, 1) > 0$. The latter requires:

$$c_i < n_1(d_1 - b_i) + n_2(d_2 - b_i) \text{ for } i = 1, 2 \tag{29}$$

Together with (24) the restriction is:

$$c_i < n_1(d_1 - b_i) + n_2(d_2 - b_i) \text{ and } c_i < n_1 b_1 + n_2 b_2 \text{ for } i = 1, 2 \tag{30}$$

The Hawk-Dove has decreasing returns and requires $\gamma_i(0, 0) > 0$, or $c_i < 0$ and $\gamma_i(1, 1) > 0$. The latter requires:

$$-c_i < n_1(b_i - d_1) + n_2(b_i - d_2) \tag{31}$$

The Hawk-Dove restriction is:

$$n_i(b_i - 2d_i) + n_j(b_i - d_1 - d_2) < -c_i < n_1(b_i - d_1) + n_2(b_i - d_2) \tag{32}$$

An interior Nash equilibrium occurs when both s_1^* and s_2^* are contained in $(0, 1)$. If $c_1 = c_2 = c$ then the EE values become: $s_1^* = \frac{c}{n_1(d_1 - d_2)}$ and $s_2^* = \frac{c}{n_2(d_2 - d_1)}$, which indicates that there is no participation in at least one population (either $s_1^* = 0, s_2^* = 0$). If $c_1 = c_2 = c$ and $d_2 > d_1$ then $s_1^* = 0$, and if $d_1 > d_2$ then $s_2^* = 0$. If $b_1 = b_2$, or $d_1 = d_2$ then the EE values in (12) are undefined. Asymmetry across populations is needed to generate an interior Nash equilibrium. There is an interior Nash equilibrium for the following parameter restrictions (Table 3).

Table 3 Interior Nash equilibrium parameter restrictions

	$s_1^* \in (0, 1)$	$s_2^* \in (0, 1)$	$s_1^*, s_2^* \in (0, 1)$
(1) Coordination small asymmetry			
$\quad b_2 < b_1 < d_2 < d_1$	iff $\frac{c_1}{c_2} > \frac{d_2-b_1}{d_2-b_2}$	iff $\frac{c_1}{c_2} < \frac{d_1-b_1}{d_1-b_2}$	iff $\frac{d_2-b_1}{d_2-b_2} < \frac{c_1}{c_2} < \frac{d_1-b_1}{d_1-b_2}$
(2) Coordination large asymmetry			
$\quad b_1 < b_2 < d_2 < d_1$	iff $\frac{c_1}{c_2} < \frac{d_2-b_1}{d_2-b_2}$	iff $\frac{c_1}{c_2} > \frac{d_1-b_1}{d_1-b_2}$	iff $\frac{d_1-b_1}{d_1-b_2} < \frac{c_1}{c_2} < \frac{d_2-b_1}{d_2-b_2}$
(3) Hawk-Dove small asymmetry			
$\quad d_2 < d_1 < b_2 < b_1$	iff $\frac{c_1}{c_2} > \frac{d_2-b_1}{d_2-b_2}$	iff $\frac{c_1}{c_2} < \frac{d_1-b_1}{d_1-b_2}$	iff $\frac{d_2-b_1}{d_2-b_2} < \frac{c_1}{c_2} < \frac{d_1-b_1}{d_1-b_2}$
(4) Hawk-Dove large asymmetry			
$\quad d_1 < d_2 < b_2 < b_1$	iff $\frac{c_1}{c_2} < \frac{d_2-b_1}{d_2-b_2}$	iff $\frac{c_1}{c_2} > \frac{d_1-b_1}{d_1-b_2}$	iff $\frac{d_1-b_1}{d_1-b_2} < \frac{c_1}{c_2} < \frac{d_2-b_1}{d_2-b_2}$

The two populations must also be of sufficient size for an interior equilibrium to exist: $n_1 > \frac{c_1(d_2-b_2)+c_2(b_1-d_2)}{(b_1-b_2)(d_1-d_2)}$ and $n_2 > \frac{c_1(d_1-b_2)+c_2(b_1-d_1)}{(b_1-b_2)(d_2-d_1)}$

References

Barrett S (1994) Self-enforcing international environmental agreements. Oxf Econ Pap 46:878–894

Barrett S (1997) Heterogeneous international environmental agreements. In: Carraro C (ed) International environmental negotiations. Strategic policy issues. Edward Elgar, Cheltenham, UK

Barrett S (1999) A theory of full international cooperation. J Theor Politics 11:519–541

Barrett S (2001) International cooperation for sale. Eur Econ Rev 45:1835–1850

Barrett S (2003) Environment and statecraft: the strategy of environmental treaty-making. Oxford Press, Oxford University Press, UK

Carraro C, Siniscalco D (1993) Strategies for the international protection of the environment. J Public Econ 52:309–328

Chander P, Tulkens H (1995) A core-theoretic solution for the design of cooperative agreements on transfrontier pollution. Int Tax Public Financ 2:279–293

d'Aspremont C, Jacquemin A, Gabszewicz JJ, Weymark J (1983) On the stability of collusive price leadership. Can J Econ 16:17–25

Dijkstra B, de Vries F (2006) Location choice by households and polluting firms: an evolutionary approach. Eur Econ Rev 50:425–446

Finus M (2003) Stability and design of international environmental agreements: the case of transboundary pollution. In: Folmer H, Tietenberg T (eds) International yearbook of environmental and resource economics, 2003/4. Edward Elgar, Cheltenham, UK, pp 82–158

Fisher E, Kakkar V (2004) On the evolution of comparative advantage in matching models. J Int Econ 64:169–193

Folmer H, van Mouche P, Ragland J (1993) Interconnected games and international environmental problems. Environ Resour Econ 3:313–335

Folmer H, van Mouche P (1994) Interconnected games and international environmental problems, II. Ann Oper Res 54:97–117

Friedman D (1991) Evolutionary games in economics. Econometrica 59(3):637–666

Friedman D (1996) Equilibrium in evolutionary games: some experimental results. Econ J 106(434):1–25

Friedman D (1998) On economic applications of evolutionary game theory. J Evol Econ 8(1):15–43

Friedman D, Fung KC (1996) International trade and the internal organization of firms: an evolutionary approach. J Int Econ 41:113–137

Mas-Colell A, Whinston M, Green J (1995) Microeconomic theory. Oxford University Press, New York

Maynard Smith J (1974) The theory of games and the evolution of animal conflicts. J Theor Biol 47:209–221

McGinty M (2007) International environmental agreements among asymmetric nations. Oxf Econ Pap 59(1):45–62

Nordhaus W, Yang Z (1996) A regional dynamic general-equilibrium model of alternative climate-change strategies. Am Econ Rev 86(4):741–765

Weibull J (1995) Evolutionary game theory. MIT Press, Cambridge, MA

Weikard H (2009) Cartel stability under an optimal sharing rule. Manchester School 77(5): 575–593