Realization of a self-enforcing international environmental agreement by matching schemes

Fujita, Toshiyuki
Faculty of Economics, Kyushu University : Assosiate Professor

出版情報：経済学研究院ディスカッション・ペーパー, 2011. 九州大学大学院経済学研究院
バージョン：
権利関係：

Realization of a self-enforcing international environmental agreement by matching schemes

FUJITA Toshiyuki*

Faculty of Economics, Kyushu University

Abstract

We incorporate matching schemes into a model of transboundary environmental agreements and investigate their effectiveness using three-stage game model. In the first stage, each country decides whether or not to accede to the agreement. In the second stage, the signatories collectively choose a common matching rate. Finally, in the third stage, each signatory and non-signatory determines its unconditional flat abatement noncooperatively, taking the value of the matching rate as given. Each signatory is imposed an additional abatement obtained by multiplying the total of all the other countries' flat abatements by the matching rate. The analysis of a matching agreement game with symmetric countries as players suggests the existence of a self-enforcing agreement leading to an efficient and equitable outcome, and thus shows that matching schemes are effective.

Keywords: Transboundary pollution, International environmental agreements, Matching agreement, Game theory, Self-enforcement.

1 Introduction

One remarkable feature of transboundary environmental problems is that no organization has the supranational power to control anthropogenic pollutants. Hence, to develop measures to protect the environment, it is essential to conclude an international agreement. An agreement is said to be *self-enforcing* if no country has an incentive to change decisions on the accedence. The design of the agreement should prevent any free riding and thus realize the large self-enforcing agreement.

Many theoretical analyses of international environmental agreements have been conducted to date, but earlier studies by Carraro and Siniscalco (1993), Barrett

*6-19-1 Hakozaki, Higashiku, Fukuoka 812-8581, JAPAN, E-mail: tfujita@en.kyushu-u.ac.jp

(1994), and others have shown that the size of the self-enforcing agreement is generally small.[1] They typically describe the agreements as two-stage games with sovereign countries as players. In the first stage, each country decides simultaneously whether to accede to the agreement or not. In the second stage, the signatories collectively choose the abatements to maximize the total payoff to the signatories, whereas each non-signatory behaves noncooperatively. The payoff is defined as a function of the abatements of all countries. Each country considers its final payoff and decides whether to be a signatory.

In this paper, we apply the *matching* concept proposed by Guttman (1978) and developed by Guttman (1987) and Guttman and Schnytzer (1992) to international environmental agreements, and investigate the effectiveness of *matching agreements* on environmental protection. Rübbelke (2006) is the first study that applies matching schemes to environmental problems. Matching agreements do not fix the pollution abatement but the so-called "*matching rate*" for each country. Subsequently, each country fixes an unconditional *flat abatement* noncooperatively. Consequently, both its own flat abatement and the flat abatement of other countries multiplied by its matching rate are imposed on each country. Rübbelke (2006) extends Guttman (1987) and focuses on the ancillary benefits of the global environmental policy. He emphasizes that matching agreements are effective because generally, there is no need for renegotiations regarding the matching rate as each country can update its flat abatement rate whenever a new ancillary benefit is discovered.

Rübbelke (2006) assumes that all countries accept the matching rules unconditionally. However, we cannot deny the possibility of the emergence of countries that refuse the matching. We analyze matching schemes within the framework of self-enforcing agreements based on the simple pollution abatement models by Barrett (1994) and Na and Shin (1998), and show that there exists an equilibrium in which all countries participate in the agreement and that an efficient and equitable outcome is realized.

Carraro and Siniscalco (1993), Barrett (1994), and other earlier studies of international environmental agreements describe that the size of self-enforcing agreements does not increase unless considerable restrictions are imposed on the payoff structure and behavioral patterns of each country in the setting or assumption of the model. The concept of a matching agreement in our model, in contrast, is noteworthy because it does not force the countries to undertake any commitment in addition to following the simple rules of matching. The contributions of our study are as follows: it defines the matching agreement through the application of the matching theory to a framework of an international environmental agreement by Barrett (1994) and Na and Shin (1998); and it clearly shows the self-enforcing agreement scheme by matching that has not been achieved before. In addition, we

[1] More recent studies with various factors include Petrakis and Xepapadeas (1996), Hoel and Schneider (1997), Na and Shin (1998), Barrett (2001), Lange and Vogt (2003), and Boadway *et al.* (2007).

have extended previous studies by allowing the existence of countries that do not commit to the matching, and analyzed the interdependencies between the actions of signatories and those of non-signatories.

Below, in Section 2, after formulating a model of international pollution abatement and obtaining the first-best solution, we define a game that incorporates matching agreements into the basic model and explain the game's solution concepts. In Section 3, we prove the existence of efficient and self-enforcing agreements and show the effectiveness of matching agreements. We also introduce the results of a simple numerical example. Finally, we summarize the concluding remarks in Section 4.

2 Model of matching agreements

2.1 International environmental agreements

Let us start from the description of the game of standard international environmental agreements. Players of the game are the governments of n symmetric countries sharing the environment. Let $N = \{1, \cdots, n\}$ denote the set of countries. We focus on certain transboundary pollutants. The benefit due to a single country's pollution abatement is proportional to the amount of abatement and affects all countries. The abatement cost is incurred entirely by the abating country. Let $\boldsymbol{x} = (x_1, \cdots, x_n)$ denote the vector of abatement of countries, and the payoff to country $i (\in N)$ is

$$\pi_i = \sum_{j \in N} x_j - C(x_i), \tag{1}$$

where $C(\cdot)$ is a three-times differentiable cost function that satisfies $C(0) = C'(0) = 0$, $\lim_{x \to \infty} C(x) = \lim_{x \to \infty} C'(x) = \infty$ and $C' > 0, C'' > 0, C''' \geq 0$ for all $x > 0$. As can be seen from (1), the technological structure and the damage due to pollution for each country are identical.

The first-best abatement that maximizes the total payoff to all countries or, in other words, satisfies the Samuelson condition is $x_i = \Omega(n) \equiv C'^{-1}(n)$ for all $i \in N$, obtained as a solution of $\dfrac{\partial}{\partial x_i} \sum_{j \in N} \pi_j = 0$.[2] Properties $\Omega' > 0$ and $\Omega'' \leq 0$ are derived by the assumptions on cost function C.[3] Let us assume that for each country, the emissions exceed $\Omega(n)$ in the current state. When each country decides an abatement noncooperatively, country i selects $x_i = \Omega(1)$ $(< \Omega(n))$, which maximizes its own payoff by solving $\dfrac{\partial \pi_i}{\partial x_i} = 0$; thus, letting each country act independently does not lead to an efficient outcome.

Let us examine the possibility of solving the problem by an agreement on abatement x_i by using the model of international environmental agreements. The game

[2] Since $C'(\cdot)$ is monotone, it has an inverse function.

[3] We use the property of cost function $C''' \geq 0$ to show that $\Omega'' \leq 0$.

is divided into two stages. In the first stage, countries individually decide whether or not to participate in the agreement. In the second stage, all signatories and each non-signatory determine the amount of abatement noncooperatively. In this framework, when the number of countries is large and the cost function satisfies the properties described above, the agreement is shown to be not self-enforcing.

Proposition 1 *For the pollution-abatement agreement game with* n *symmetric countries, an efficient agreement among all countries is not self-enforcing when the number of countries is four or greater.*

Proof. The final payoff to a country is determined by the number of signatories and whether the country accedes to the agreement. Let s $(= 0, 1, \cdots, n)$ denote the number of signatories. If the payoffs to signatories and non-signatories are written as $\pi^S(s)$ and $\pi^F(s)$, respectively, the condition for an agreement with all countries to be self-enforcing is

$$\pi^S(n) \geq \pi^F(n-1). \tag{2}$$

Equation (2) indicates that participating in an agreement would be rational for a country provided that all other countries participate in the agreement. When the number of signatories is s, the amount of abatement of the signatories is $\Omega(s)$. A simple calculation yields $\pi^S(n) = n\Omega(n) - C(\Omega(n))$ and $\pi^F(n-1) = (n-1)\Omega(n-1) + \Omega(1) - C(\Omega(1))$. Appendix 1 shows that $\pi^S(n) < \pi^F(n-1)$ when $n \geq 4$, and (2) does not hold. □

Using some examples, Barrett (1994) shows that when n is sufficiently large, an efficient agreement by all the countries is not self-enforcing and that some countries cannot be prevented from refusing the agreement and acting noncooperatively. Proposition 1 is considered a generalized version of Proposition 2 (p. 888) of Barrett (1994). As presented in the proof shown above, country i will not join the agreement if all other countries join the agreement because noncooperative behavior outside the agreement would be more beneficial than entering the agreement would. Since this kind of incentive works for some countries, an agreement by all the countries is not self-enforcing when the number of countries is large.

2.2 Matching agreement game and the solution concept

We introduce matching rules in the framework of Section 2.1 and define the matching agreement. The game is divided into three stages. In the first stage, the rules of matching are announced and each country individually decides whether to accede to the agreement or not. Unlike Guttman (1978) and Rübbelke (2006), we do not assume that all the countries accede to the agreement and conduct matching. If there exist countries that accede to the agreement, the signatories collectively determine a common matching rate by negotiation in the second stage. In the third stage, each country determines its flat abatement noncooperatively, taking the matching rate as given. If no countries accede to the agreement in the first stage, the agreement does not exist. Each country determines the abatement in the second stage and the game ends.

Under the matching agreement, each signatory is imposed an additional abatement, which is the amount calculated by multiplying the matching rate fixed in the second stage with the total flat abatement determined by all the other countries, including the non-signatories. We assume that the matching agreement has a certain binding authority and that the signatories are committed to the matching rate and comply with the abatement settled in the agreement after the end of the third stage.[4] As non-signatories do not commit to the matching rules, the actual abatements are the same as their flat abatements. After the actual abatement, the final payoff to each country is determined.

Let a_i (≥ 0) denote the flat abatement of country $i \in N$ and b (≥ 0), the matching rate of the agreement. In addition, let S denote the set of signatories, and let $s \equiv |S|$. If country i is a signatory, that is, $i \in S$, the abatement imposed on i is $x_i = a_i + b\sum_{j\neq i} a_j$ by the matching rules. If it is a non-signatory, that is, $i \notin S$, $x_i = a_i$. Therefore, from (1),

$$\pi_i = \begin{cases} \sum_{j\in S} a_j(1+b(s-1)) + \sum_{j\notin S} a_j(1+bs) - C\left(a_i + b\sum_{j\neq i} a_j\right) & (i \in S), \\ \sum_{j\in S} a_j(1+b(s-1)) + \sum_{j\notin S} a_j(1+bs) - C(a_i) & (i \notin S). \end{cases} \tag{3}$$

As the payoff depends on the flat abatement vector $\boldsymbol{a} \equiv (a_1, \cdots, a_n)$, matching rate b, and the size of agreement (number of signatories) s, we subsequently denote the right-hand side of the first and the second lines of (3) by $\pi_i^S(\boldsymbol{a}, b, s)$ and $\pi_i^F(\boldsymbol{a}, b, s)$, respectively.

Let us define the solution of the matching agreement game. We solve the game backwards. In the third stage, given the size of agreement and the matching rate determined in the first and second stages, country i determines its abatement a_i as the optimal response with respect to the set of flat abatement of other countries $\boldsymbol{a}_{-i} = (a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_n)$. The Kuhn-Tucker condition for maximizing country i's benefit is

$$\frac{\partial \pi_i}{\partial a_i} \leq 0, \quad \frac{\partial \pi_i}{\partial a_i} \cdot a_i = 0. \tag{4}$$

[4]This assumption means that no signatory refuses the additional abatement derived from the matching agreement.

From (3), we obtain

$$\frac{\partial \pi_i}{\partial a_i} = \begin{cases} 1 + b(s-1) - C'\left(a_i + b\sum_{j \neq i} a_j\right) & (i \in S), \\ 1 + bs - C'(a_i) & (i \notin S). \end{cases}$$

When the optimal response $a_i^*(b, s, \boldsymbol{a}_{-i})$ takes a positive value as an interior solution, $\frac{\partial \pi_i}{\partial a_i} = 0$ leads to

$$a_i^*(b, s, \boldsymbol{a}_{-i}) = \begin{cases} \Omega(1 + b(s-1)) - b\sum_{j \neq i} a_j & (i \in S), \\ \Omega(1 + bs) & (i \notin S). \end{cases} \tag{5}$$

If the right-hand side of (5) is nonpositive for some i, then $a_i^*(b, s, \boldsymbol{a}_{-i}) = 0$ from (4). When all countries take optimal responses, that is, when simultaneous equations $a_i^* = a_i^*(b, s, \boldsymbol{a}_{-i}^*)$ $\forall i \in N$ are satisfied, the combination of flat abatement $(a_1^*, \cdots, a_n^*)$ is the equilibrium of the third stage. Below, we express it as $\boldsymbol{a}^*(b, s)$.

The solution of the second stage is b^* which satisfies

$$\sum_{j \in S} \pi_j^S(\boldsymbol{a}^*(b^*, s), b^*, s) \geq \sum_{j \in S} \pi_j^S(\boldsymbol{a}^*(b, s), b, s) \ \ \forall b \geq 0. \tag{6}$$

Inequality (6) shows that the matching rate determined in the second stage is b^*, which maximizes the total payoff to the signatories. As the signatories make decisions taking the size of the agreement as given, b^* is dependent on s. Therefore, this solution is expressed as $b^*(s)$.

Finally, the solution of the first stage is S^*, which satisfies the following equations simultaneously:

$$\forall i \in S^*,\ \pi_i^S(\boldsymbol{a}^*(b^*(s^*), s^*), b^*(s^*), s^*) \geq \pi_i^F(\boldsymbol{a}^*(b^*(s^*-1), s^*-1), b^*(s^*-1), s^*-1), \tag{7}$$

$$\forall j \notin S^*,\ \pi_j^F(\boldsymbol{a}^*(b^*(s^*), s^*), b^*(s^*), s^*) > \pi_j^S(\boldsymbol{a}^*(b^*(s^*+1), s^*+1), b^*(s^*+1), s^*+1), \tag{8}$$

where $s^* \equiv |S^*|$. Note that when $S^* = \emptyset$ ($s^* = 0$), the right-hand side of (7) is not defined, and therefore, (8) is the only condition for $S^* = \emptyset$ to be the solution. Similarly, (7) is the only condition for $S^* = N$ ($s^* = n$) to be the solution. Inequality (7) shows that even if one country withdraws the agreement of size s^*, the final payoff to that country does not increase. Inequality (8) shows that if one country outside the agreement enters into the agreement of size s^*, the final payoff to that country decreases. Inequalities (7) and (8) are the characteristics known as the internal stability and the external stability of the agreement, and an agreement that satisfies these characteristics is self-enforcing. If a certain s^* satisfies (7) and (8) in the first stage, no country has any incentive to change the decision pertaining to entering into the agreement and an agreement of size s^* is realized as a result of equilibrium.

3 Solution of the matching agreement game

3.1 Efficiency and self-enforcement of full agreement

In this section, we examine the solution of matching games. We refer to the agreement in which all countries participate as the full agreement, and Lemma 1 shows that the matching rules realize an efficient outcome when the full agreement is formed.

Lemma 1 *For the pollution-abatement matching agreement game with n symmetric countries, the first-best abatement is led by the solutions of the second and the third stages when the full agreement is formed in the first stage.*

Proof. Let b denote the matching rate determined in the second stage. We assume symmetric equilibrium as the solution of the third stage, that is, $a_i^* = a_j^*$ for all $i, j \in S(= N)$. Subsequently, from (5), we have $a_i^* = \Omega(1+b(n-1)) - b(n-1)a_i^*$ for all $i \in N$, and it follows that $a_i^* = \dfrac{\Omega(1+b(n-1))}{1+b(n-1)}$. The actual abatement of country i becomes $x_i^* = a_i^* + b(n-1)a_i^* = \Omega(1+b(n-1))$. As $\Omega(\cdot)$ is a monotonically increasing function, x_i^* is equal to the first-best level $\Omega(n)$ if and only if $b = 1$. Therefore, the matching rate $b^* = 1$ is determined in the second stage, and an efficient outcome is realized. □

Next, Lemma 2 shows that the full agreement is self-enforcing. This means that all countries accede to the agreement as a result of noncooperative decisions and that no single country has an incentive not to accede.

Lemma 2 *For the first stage of the pollution-abatement matching agreement game with n symmetric countries, the full agreement is self-enforcing.*

Proof. As there are no countries outside the agreement, we need to check only the internal stability. We show that in the case where country i alone does not accede to the agreement ($S = N\backslash\{i\}$), it cannot increase its payoff compared to the case where it accedes to the full agreement.

When $S = N\backslash\{i\}$ and the matching rate is b, from (5), the flat abatement determined by each country is calculated as follows:

$$\begin{cases} a_i^* = \Omega(1+b(n-1)), \\ a_j^* = \max\left(\Omega(1+b(n-2)) - b\displaystyle\sum_{k\neq j} a_k^*,\ 0\right) \quad (j \in S). \end{cases}$$

Again, we assume symmetric equilibrium in the agreement, that is, $a_j^* = a_k^*$ for any $j, k \in S$. Assuming $a_j^* > 0$ for all $j \in S$, we obtain $a_j^* = \Omega(1+b(n-2)) - b(a_i^* + (n-2)a_j^*)$, and it follows that $a_j^* = \dfrac{\Omega(1+b(n-2)) - b\Omega(1+b(n-1))}{1+b(n-2)}$. Consequently, the flat abatement determined by each signatory in the third stage is

$$a_j^* = \begin{cases} \dfrac{\Omega(1+b(n-2)) - b\Omega(1+b(n-1))}{1+b(n-2)} & (b \in B_1), \\ 0 & (b \in B_2), \end{cases}$$

where $B_1 \equiv \{b \geq 0 \mid \Omega(1+b(n-2)) \geq b\Omega(1+b(n-1))\}$ and $B_2 \equiv \{b \geq 0 \mid \Omega(1+b(n-2)) < b\Omega(1+b(n-1))\}$. We have $B_1 \subset [0,\ 1)$ because $\Omega(1+b(n-2)) < b\Omega(1+b(n-1))$ holds when $b \geq 1$.

Let us derive the matching rate that maximizes the total payoff to the signatories. When $b \in B_1$, the actual abatement of each country in equilibrium is $x_i^* = \Omega(1+b(n-1))$ and $x_j^* = \Omega(1+b(n-2))$ $(j \in S)$. Therefore, for all $j \in S$, the payoff in equilibrium is

$$\pi_j^S = \Omega(1+b(n-1)) + (n-1)\Omega(1+b(n-2)) - C(\Omega(1+b(n-2))).$$

Considering this as a function of b and differentiating with respect to b, we have

$$\frac{d\pi_j^S(b)}{db} = (n-1)\Omega'(1+b(n-1)) + (n-2)^2(1-b)\Omega'(1+b(n-2)) > 0;$$

thus, π_j^S is a monotonically increasing function in the range of $b \in B_1$. Thus, if we denote by $\bar{b}$ the maximum value of b that satisfies $\Omega(1+b(n-2)) = b\Omega(1+b(n-1))$, it follows that $\bar{b} \in (0,\ 1)$ and

$$\pi_j^S(\bar{b}) \geq \pi_j^S(b) \ \ \forall b \in B_1. \tag{9}$$

Next, when $b \in B_2$, from $x_i^* = \Omega(1+b(n-1))$ and $x_j^* = b\Omega(1+b(n-1))$ $(j \in S)$, the payoff to country j in equilibrium is

$$\pi_j^S = \{1+b(n-1)\}\Omega(1+b(n-1)) - C(b\Omega(1+b(n-1))).$$

Differentiating this with respect to b again, we obtain

$$\begin{aligned}\frac{d\pi_j^S(b)}{db} = &(n-1)\Omega(1+b(n-1)) + \{1+b(n-1)\}(n-1)\Omega'(1+b(n-1)) \\ &-\{\Omega(1+b(n-1)) + b(n-1)\Omega'(1+b(n-1))\}C'(b\Omega(1+b(n-1))).\end{aligned}$$

Using Appendices 2 and 3, we can show the following inequalities:

$$\lim_{b \to \bar{b}+0} \frac{d\pi_j^S(b)}{db} > 0, \tag{10}$$

$$\frac{d\pi_j^S(b)}{db} < 0 \ \ \forall b \geq 1. \tag{11}$$

From (9), (10), and (11), we have

$$\exists b' \in B_2 \cap (0,\, 1) \text{ s.t. } \pi_j^S(b') \geq \pi_j^S(b) \ \ \forall b \geq 0,$$

which means b', which maximizes π_j^S, exists in the range of (0, 1).

The matching rate determined by the agreement is less than unity; thus we obtain $x_i^* > x_j^*$, that is, payoff to country i is smaller than the average payoff to all countries. Moreover, the total payoff is maximized when the full agreement is formed. Therefore, we have

$$\pi_i^F(n-1) < \frac{1}{n}\left\{\pi_i^F(n-1) + \sum_{j\in S}\pi_j^S(n-1)\right\} \leq \frac{1}{n}\sum_{j\in N}\pi_j^S(n) = \pi_i^S(n).$$

From above, the payoff to country i is smaller than that when it accedes to the full agreement. Thus, there is no incentive for country i to refuse the accedence to the agreement. □

Now, we obtain the following proposition directly from Lemmas 1 and 2.

Proposition 2 *For the first stage of the pollution-abatement matching agreement game with n symmetric countries, there exists a self-enforcing solution that leads to an efficient outcome.*

Proof. Lemmas 1 and 2 show that the full agreement, which leads to an efficient outcome, is self-enforcing. □

When the full agreement is formed, the matching rate is unity; thus, the actual abatement of each country equals to the sum of flat abatements of all countries. Even if an individual country deviates, the matching rate determined by the agreement among the remaining countries becomes less than unity. Consequently, payoff to the deviating country decreases and the full agreement is thus self-enforcing. For the outcome led by the equilibrium, the equality of all the countries' abatements means that payoffs are also equal. This is a desirable outcome from the perspective of equity.

3.2 Examples

Here, we explain how the matching agreement works by using a simple numerical example. Let us assume $n = 4$ and $C(x_i) = x_i^2/2$. The payoff to country i is

$$\pi_i = \sum_{i=1}^{4} x_i - \frac{x_i^2}{2}.$$

The first-best abatements are $x_1 = \cdots = x_4 = 4$, and the corresponding payoffs are $\pi_1 = \cdots = \pi_4 = 8$, but when each country decides its abatement noncooperatively, $x_1 = \cdots = x_4 = 1$ and each country's payoff decreases to 7/2.

Let us suppose that a full matching agreement has been concluded. The matching rate should be determined as unity in the second stage. Equation (5) becomes $a_i^* = 4 - 3a_i^*$ for all $i \in N$, and equilibrium of the third stage is the flat abatement vector $\boldsymbol{a}^* = (1, 1, 1, 1)$. In this case, $x_i^* = 4$ for all i is satisfied and efficiency is achieved. Each country's payoff is 8.

Suppose that only one country, for example, country 1, does not accede to the agreement. If the matching rate determined by the signatories (countries 2, 3, and 4) is b, from (5), the flat abatements become

$$a_1^* = 1 + 3b,\ a_2^* = a_3^* = a_4^* = \begin{cases} 1 - \dfrac{b(1+3b)}{1+2b} & \left(0 \leq b \leq \dfrac{1+\sqrt{13}}{6}\right), \\ 0 & \left(b > \dfrac{1+\sqrt{13}}{6}\right), \end{cases}$$

and the actual abatements are

$$x_1^* = 1 + 3b,\ x_2^* = x_3^* = x_4^* = \begin{cases} 1 + 2b & \left(0 \leq b \leq \dfrac{1+\sqrt{13}}{6}\right), \\ b(1+3b) & \left(b > \dfrac{1+\sqrt{13}}{6}\right). \end{cases}$$

Hence, the payoff to country $j(\neq 1)$ is

$$\pi_j^S = \begin{cases} 1 + 3b + 3(1+2b) - \dfrac{(1+2b)^2}{2} & \left(0 \leq b \leq \dfrac{1+\sqrt{13}}{6}\right), \\ 1 + 3b + 3b(1+3b) - \dfrac{b^2(1+3b)^2}{2} & \left(b > \dfrac{1+\sqrt{13}}{6}\right). \end{cases}$$

Since π_j^S takes a maximum value when $b = \dfrac{-1+\sqrt{145}}{12} \approx 0.92$, the matching rate determined in the second stage is around 0.92. Subsequently, the payoffs to the signatories and the non-signatory (country 1) are, respectively, calculated as 8.15 and 7.07 approximately. Payoff to country 1 is smaller than that when it accedes to the full agreement; thus, there is no incentive for country 1 to deviate.

4 Conclusions

We have considered matching agreements proposed by Rübbelke (2006) as a model of transboundary environmental agreements and investigated their effectiveness.

In Section 2, we first present Proposition 1, stating that efficient agreements on the amount of abatement are not normally self-enforcing in the game of international environmental agreement on pollution abatement. This proposition is regarded as the generalization of the proposition provided by Barrett (1994). Subsequently, we present a model that includes matching rules. The players in the model are symmetric countries, whose benefits are specified as linear functions. In the first stage, the rules of matching are announced and each country individually decides whether to accede to the agreement or not. In the second stage, the signatories collectively determine a common matching rate through negotiations. In the third stage, each country determines its flat abatement noncooperatively, taking the matching rate as given. Under the matching agreement, each signatory is imposed an additional abatement, which is the amount calculated by multiplying the matching rate fixed in the second stage with the total flat abatement determined by all the other countries including the non-signatories.

In Section 3, we examine the solution of the matching agreement game formulated in Section 2. First, we show that the full agreement concluded by all the countries achieves an efficient outcome through equilibrium in stages 2 and 3 (Lemma 1). Next, this agreement turns out to be self-enforcing, that is, it has internal stability (Lemma 2). From these lemmas, we show the existence of a self-enforcing agreement that achieves efficient outcome (Proposition 2). All countries obtain the same payoff. Each country enters into the agreement without being forced and the outcome achieved is efficient and equitable. Hence, we have shown the effectiveness of the matching agreement.

Under a situation where it is difficult to prevent free riders in the international environmental agreement, the concept of matching could provide a major clue. Unlike an agreement on the amount of pollution abatement, even if an individual country refuses to enter into the matching agreement, its payoff will not increase if other countries remain in the agreement. Therefore, there is no incentive to leave the agreement. This is because the decisions of the matching rate and the flat abatement in the second and third stages function as a punishment for the deviator.

Future issues to be studied include more general scenarios such as the analysis of cases where the benefit functions of the countries are more general, analysis of cases where the players are asymmetric, and an examination of the problem of multiple countries departing from the agreement (problem of coalitional rationality).

Appendix

1 Proof that (2) does not hold

Based on the discussions in the proof of Proposition 1, we have

$$\pi^F(n-1)-\pi^S(n) = C(\Omega(n))-C(\Omega(1))-(n-1)\{\Omega(n)-\Omega(n-1)\}-\{\Omega(n)-\Omega(1)\}. \quad \text{(A1)}$$

Convexity of the cost function $C(\cdot)$ and the property $\Omega' > 0$ yield

$$\begin{aligned} C(\Omega(n)) &> C(\Omega(n-1)) + C'(\Omega(n-1))\{\Omega(n)-\Omega(n-1)\} \\ &= C(\Omega(n-1)) + (n-1)\{\Omega(n)-\Omega(n-1)\}. \end{aligned} \quad \text{(A2)}$$

Similarly we have

$$\begin{cases} C(\Omega(n-1)) & > & C(\Omega(n-2)) + (n-2)\{\Omega(n-1)-\Omega(n-2)\}, \\ & \vdots & \\ C(\Omega(2)) & > & C(\Omega(1)) + \Omega(2) - \Omega(1). \end{cases} \quad \text{(A3)}$$

Substituting (A3) in (A2), we obtain

$$C(\Omega(n))-C(\Omega(1)) > (n-1)\{\Omega(n)-\Omega(n-1)\}+(n-2)\{\Omega(n-1)-\Omega(n-2)\}+\cdots+\Omega(2)-\Omega(1). \quad \text{(A4)}$$

From (A1) and (A4), we have when $n \geq 4$

$$\begin{aligned} \pi^F(n-1) - \pi^S(n) &> (n-2)\{\Omega(n-1)-\Omega(n-2)\} + \cdots + \Omega(2) - \Omega(1) - \{\Omega(n)-\Omega(1)\} \\ &= (n-2)\{\Omega(n-1)-\Omega(n-2)\} + \cdots + \Omega(2) - \Omega(1) - \{\Omega(n)-\Omega(n-1)\} \\ &\qquad -\{(\Omega(n-1)-\Omega(n-2)\} - \cdots - \{\Omega(2)-\Omega(1)\} \\ &> \{\Omega(3)-\Omega(2)\} - \{\Omega(n)-\Omega(n-1)\}. \end{aligned} \quad \text{(A5)}$$

Using (A5) and $\Omega'' \leq 0$, we can show $\pi^F(n-1) - \pi^S(n) > 0$. □

2 Proof of (10)

Definition of $\overline{b}$ leads to $\Omega(1+\overline{b}(n-2)) = \overline{b}\Omega(1+\overline{b}(n-1))$. Therefore,

$$\begin{aligned} \left.\frac{d\pi_j^S(b)}{db}\right|_{b=\overline{b}} &= (n-1)\Omega(1+\overline{b}(n-1)) + \{1+\overline{b}(n-1)\}(n-1)\Omega'(1+\overline{b}(n-1)) \\ &\qquad -\{\Omega(1+\overline{b}(n-1)) + \overline{b}(n-1)\Omega'(1+\overline{b}(n-1))\}C'(\overline{b}\Omega(1+\overline{b}(n-1))) \\ &= (n-1)\Omega(1+\overline{b}(n-1)) + \{1+\overline{b}(n-1)\}(n-1)\Omega'(1+\overline{b}(n-1)) \\ &\qquad -\{\Omega(1+\overline{b}(n-1)) + \overline{b}(n-1)\Omega'(1+\overline{b}(n-1))\}\{1+\overline{b}(n-2)\} \\ &= (1-\overline{b})(n-2)\Omega(1+\overline{b}(n-1)) + (n-1)\{1+\overline{b}(1-\overline{b})(n-2)\}\Omega'(1+\overline{b}(n-1)). \end{aligned}$$

It is clear that (10) holds when $n \geq 2$. □

3 Proof of (11)

Since $b\Omega(1+b(n-1))$ is monotonically increasing with b, if $b \geq 1$, then $b\Omega(1+b(n-1)) \geq \Omega(n)$. Monotonicity of $C'(\cdot)$ leads to $C'(b\Omega(1+b(n-1))) \geq C'(\Omega(n)) = n$. Hence, if $b \geq 1$,

$$\begin{aligned} \frac{d\pi_j^S(b)}{db} &\leq (n-1)\Omega(1+b(n-1)) + \{1+b(n-1)\}(n-1)\Omega'(1+b(n-1)) \\ &\qquad -n\{\Omega(1+b(n-1)) + b(n-1)\Omega'(1+b(n-1))\} \\ &= -\Omega(1+b(n-1)) + (1-b)(n-1)\Omega'(1+b(n-1)) < 0, \end{aligned}$$

and (11) holds. □

References

Barrett, S. (1994), "Self-Enforcing International Environmental Agreements," *Oxford Economic Papers*, vol. 46, supplementary issue, pp. 878-894.

Boadway, R., Z. Song and J. -F. Tremblay (2007), "Commitment and Matching Contributions to Public Goods," *Journal of Public Economics*, vol. 91, pp. 1664-1683.

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

Danziger, L. and A. Schnytzer (1991), "Implementing the Lindahl Voluntary-Exchange Mechanism," *European Journal of Political Economy*, vol. 7, pp. 55-64.

Guttman, J. M. (1978), "Understanding Collective Action: Matching Behavior," *American Economic Review*, vol. 68, pp. 251-255.

Guttman, J. M. (1987), "A Non-Cournot Model of Voluntary Collective Action," *Economica*, vol. 54, pp. 1-19.

Guttman, J. M. and A. Schnytzer (1992), "A Solution of the Externality Problem Using Strategic Matching," *Social Choice and Welfare*, vol. 9, pp. 73-88.

Na, S. -l. and H. S. Shin (1998), "International Environmental Agreements under Uncertainty," *Oxford Economic Papers*, vol. 50, pp. 173-185.

Rübbelke, D. T. G. (2006), "Analysis of an International Environmental Matching Agreement," *Environmental Economics and Policy Studies*, vol. 8, pp. 1-31.