Strategic Behavior and the Environment, 2024, 10: 15–38

Supporting Environmental Agreements under Asymmetry and Minimum Participation Rule

Qian Li[1], Toshiyuki Fujita[2,*] and Yutian Zhang[3]

[1]*Faculty of Economics, Nagoya University, Furocho, Chikusaku, Nagoya 464-8601, Japan*
[2]*Faculty of Economics, Kyushu University, 744 Motooka, Nishiku, Fukuoka 819-0395, Japan; tfujita@econ.kyushu-u.ac.jp*
[3]*Graduate School of Engineering, Kyushu University, 744 Motooka, Nishiku, Fukuoka 819-0395, Japan*

ABSTRACT

This study analyzes the efficiency of external transfers in international environmental agreements (IEAs) with two types of heterogeneous countries differing in abatement benefit or abatement cost parameters. We introduce a minimum participation rule concerning the number of supporters, who commit to transfer welfare to induce all the other countries to form a self-enforcing IEA. The analytical result shows that an equilibrium exists where all countries except supporters become members of the agreement under a certain condition. The simulation results suggest that the higher the heterogeneity in the abatement benefit, the larger the IEA size and the higher the relative gains of IEAs; the degree of

---

*Corresponding author.

This research was supported by the JSPS KAKENHI, Grant Number JP21K01485. We are grateful to the editor and two anonymous reviewers for their valuable comments and suggestions. We are also grateful to Hsiao-chi Chen and the participants in the 8th Congress of the East Asian Association of Environmental and Resource Economics.

MS received on 6 September 2022; revised 19 September 2023; accepted 25 September 2023
ISSN 1944-012X; DOI 10.1561/102.00000106

heterogeneity in the abatement cost has little impact on the IEA size but contributes to the relative gains of IEAs.

Keywords: International environmental agreements; coalition formation; minimum participation rule; transfers; heterogeneous countries

JEL Codes: C72, Q50

Introduction

Transboundary pollutant problems, such as global warming, are caused by human activity in each country and have a negative impact on the other countries. Since there is no international agency that can force countries to reduce their emissions, solving these problems relies on countries' voluntary actions by forming international environmental agreements (IEAs). In recent decades, the efficiency of IEAs has been investigated from different aspects, and free riding is regarded as the main obstacle to their success. Barrett (1994) shows that self-enforcing IEAs typically include few countries when a large coalition can greatly improve the non-cooperative result. The concept of self-enforcing IEAs originates from D'Aspremont *et al.* (1983) and describes that neither a member nor a free rider of IEAs has any incentive to change its membership status. Therefore, the design of IEAs should overcome free riding at the participation stage.

Theoretical studies show that using external transfers is an effective way to increase participants in IEAs in different contexts. Carraro and Siniscalco (1993) show that if some countries outside the agreement commit to transfer welfare to incentivize the remaining countries' participation, a stable coalition including all the remaining countries can be formed. Ansink *et al.* (2019) extend this scheme into a four-stage sequential game played by symmetric countries, where countries paying transfers are called supporters. In the first stage, countries decide whether to become supporters and in the second stage, supporters decide whether to accept the transfer proposal given by the international climate agency. In the third stage, the remaining countries determine their IEA membership and in the last stage, each country chooses its abatement level. The supporter scheme leads to the same result as that of Carraro and Siniscalco (1993).

In reality, countries are asymmetric in terms of abatement benefits and abatement costs of pollution, taking the technology level and monetary valuation of abatement into account. Some IEA literature have considered the implications of asymmetric countries by allowing for unlimited types of countries (Bakalova and Eyckmans, 2019; Finus and McGinty, 2019; McGinty, 2007, 2020) while others allow for two types of countries (Diamantoudi *et al.*, 2018a,b; Fuentes-Albero and Rubio, 2010; Pavlova and de Zeeuw, 2013). Since a member country with a low-cost parameter abates a large amount at a cost-effective solution, transfers within the coalition (internal transfers) to compensate these countries are always necessary to stabilize a coalition that includes asymmetric countries (Diamantoudi *et al.*, 2018a; Fuentes-Albero and Rubio, 2010). Besides, these studies investigate the relationship among the degree of asymmetry, the number of members (the size), and the efficiency of IEAs with internal transfers. As for external transfers, Li and Fujita (2021) incorporate asymmetry in the model of Ansink *et al.* (2019) to explore the efficiency of transfers under strong asymmetry among countries. The analytical result confirms their efficiency and finds them more effective than internal transfers.

This study reinforces earlier analyses on external transfers with asymmetric countries in Li and Fujita (2021) by focusing on the impact of heterogeneity on the efficiency of IEAs. We consider two types of countries and use simulation method to analyze how asymmetry affects the stability and efficiency of IEAs. Each country has a linear-quadratic benefit–cost function for abatement levels while type 2 countries (developed countries) are assumed to have higher benefit and cost parameters than type 1 countries (developing countries). This study is the first to make such an analysis under the framework of external transfers.

Regarding the implementation of external transfers, this study combines commitment and minimum participation rule (MPR) in the coalition formation game. In the first stage, type 2 countries vote for an MPR regarding the number of supporters. The supporter scheme is effective only if the number of countries that become supporters exceeds a certain value called a minimum participation number.

In the second stage, these countries will decide whether to become supporters. Once the MPR is satisfied, supporters will commit to transfer welfare to guarantee a self-enforcing coalition consisting of all other countries. This study does not consider an international climate agency that proposes a grim transfer scheme, as in Ansink *et al.* (2019) and Li and Fujita (2021), since it serves the same purpose as the MPR of supporters and their commitment. International agency prescribes the amount of transfers depending on the number of supporters. Therefore, both the international agency and MPR function as a threat that the supporter scheme will stop working if the number of supporters ends up small.

Note that we only require commitment from supporters about paying transfers. Certainly, all countries' commitment to reduce emissions would lead to a first-best outcome in IEAs and external transfers would be pointless. However, achieving such a commitment in IEAs is unreasonable. Rather, supporters' commitment is reasonable if developed countries' reputation is considered. Choosing to become supporters means agreeing to contribute to the enlargement of IEAs when minimum requirement is satisfied, thus, they will not ruin their reputation by denying commitment on transfers. The remaining stages are the same as in Ansink *et al.* (2019).

Although many papers have suggested that MPR can increase participation in IEAs, none have linked it to non-member countries. Rutz (2001) examines exogenous MPR which is characterized by the number of member countries and Carraro *et al.* (2009) endogenizes it, both assuming symmetric countries. Weikard *et al.* (2015) consider asymmetric countries with internal transfers, and the MPR is specified as the sum of the member countries' non-cooperative abatement levels. Our study is the first to implement an endogenous MPR characterized by the number of supporters and study how it enlarges the size of IEAs.

The analytical result shows that an equilibrium of the game exists where countries except for supporters, of which the number equals the minimum participation number, will get transfers and become IEA members. The simulation results suggest that the higher the heterogeneity in the abatement benefits, the smaller the number of supporters, the larger the size of the IEA, and the higher the relative gains of the IEA. The relative gains measure the extent to which the

IEA closes the gap between all countries' payoffs in the no cooperation and full cooperation scenarios (Bakalova and Eyckmans, 2019; Fuentes-Albero and Rubio, 2010; McGinty, 2007). Thus, with external transfers, the efficiency of IEAs increases with the heterogeneity in abatement benefits. However, the degree of heterogeneity in abatement costs has no impact on the size of the IEA in most cases, while it increases the relative gains. Additionally, external transfers can encourage participation in the IEA compared to the no-transfer case, which internal transfers cannot do as suggested by Fuentes-Albero and Rubio (2010).[1]

The remainder of the paper is structured as follows. In the next section, the model setting and the coalition formation game are explained. In the section "Existence of equilibria with supporters", we show the condition for the existence of subgame perfect Nash equilibrium (SPNE). The section "The impact of heterogeneity on equilibria" describes the simulation results for how heterogeneity only in abatement benefits, and that in abatement costs, affect the outcome of IEA. The last section summarizes the conclusions.

IEAs with Supporters

The Model

The formation of IEAs can be modeled using the non-cooperative game. Players of the game are the governments of n_1 type 1 countries and n_2 type 2 countries. Countries of the same type are identical. Let N_i denote the set of type i countries. Every country j of type $i, i \in \{1, 2\}$ selects the continuous abatement level q_j^i of its pollutant emissions, and the aggregate abatement taken by all countries is $Q = \sum_i \sum_{j \in N_i} q_j^i$. Then, each country derives benefits from aggregate abatement, expressed as function $B_j^i(Q) = \beta_i Q$. It also suffers loss from its abatement action, represented by the cost function $C_j^i(q_j^i) = \frac{1}{2}\gamma_i(q_j^i)^2$.

The payoff of each country j of type i, π_j^i, is defined as the difference between its benefit from aggregate abatement and the cost from its

[1] Diamantoudi *et al.* (2018a) use quadratic-quadratic benefit-cost function and find that the higher the heterogeneity in the abatement benefit, the larger the size of IEAs while the improvement in both global welfare and abatement level are very small relative to the no-transfers case.

abatement:

$$\pi_j^i = \beta_i Q - \frac{1}{2}\gamma_i (q_j^i)^2 \tag{1}$$

Moreover, following Fuentes-Albero and Rubio (2010) and Li and Fujita (2021), we assume that $\beta_2 > \beta_1$ and $\gamma_2 > \gamma_1$.

In the context of climate change, we can think of type 2 countries as being developed and type 1 countries as being developing. Developed countries value the environment more and they are willing to pay more to induce global abatement, which means the aggregate abatement level will bring them higher benefits than developing countries ($\beta_2 > \beta_1$). Developed countries have already exploited advanced abatement technology like hydropower energy, therefore their marginal abatement costs for the same abatement level will be higher ($\gamma_2 > \gamma_1$).

The Game Procedures

We only allow type 2 countries to be supporters, considering their strong incentives to induce abatements by others. High benefits incentivize them to increase abatements, but high costs incentivize them to buy others' abatements rather than reduce by themselves. The MPR, characterized by the number of supporters, is voted by all type 2 countries, as no type 2 country can be forced to accept a participation constraint that they do not agree upon, and as this participation constraint is not binding for type 1 countries. The setting of the MPR follows the same procedure as in Carraro *et al.* (2009) and Weikard *et al.* (2015). Based on the above considerations, the formation of IEAs with supporters can be modeled as the following four-stage game.

Stage 1: An arbitrary type 2 country proposes a positive minimum participation number of supporters and the decision is made via unanimous voting by all type 2 countries. Let us denote this number as c. If the proposal is vetoed by any country, the supporter scheme is not applied.

Stage 2: Type 2 countries independently and simultaneously choose whether or not to become supporters, who commit to transfer welfare to stimulate all other countries to enter into the agreement. A supporter coalition will go into effect only when the number of type 2 countries which become supporters exceeds the minimum participation number. In such case, the coalition is denoted as S and the number of supporters is $s = |S|$. If the number of supporters is smaller than the minimum participation number, no transfers will be paid and the supporters consequently become free riders.

Stage 3: The remaining countries decide whether to become members of the IEA or not. An agreement will be formed and m_1 (m_2) is the number of type 1 (type 2) countries being part of the agreement. This leads to a partition of the set of countries into three subsets: supporters, members, and free riders.

Stage 4: Members choose the abatement levels to maximize coalition payoff while other countries (supporters and free riders) act as singletons to maximize individual payoffs. Each member receives a transfer from supporters.

Solving the Game

Next, we derive the SPNE by solving the game backwards. The solution for stage 4 is the abatement levels chosen by coalition members, free riders, and supporters, given that a coalition $M = M_1 \cup M_2$ formed in stage 3, where M_i denotes a set of type i members. The first-order condition for a member $j \in M_i$ is $\frac{\partial \sum_i \sum_{j \in M_i} \pi_j^i}{\partial q_j^i} = 0$, which yields $q_j^i = \frac{m_1\beta_1 + m_2\beta_2}{\gamma_i}$. Let us denote it as $q_M^i(m_1, m_2)$. Correspondingly, the first-order condition for the free rider $j \in N_i \setminus (S \cup M_i)$ is $\frac{\partial \pi_j^i}{\partial q_j^i} = 0$, which yields $q_j^i = \frac{\beta_i}{\gamma_i}$. Let us denote it as q_F^i; and certainly, the supporter's abatement level is $q_S^2 = q_F^2$. Note that a free rider always has a higher payoff than the member of the same type, since the cost part is lower while the benefit part remains the same.

In stage 3, every country except the supporters decides whether to become an IEA member. The solution for this stage is the number of type 1 and type 2 member countries, denoted as (m_1^*, m_2^*), that satisfies internal and external stability conditions. Let $T_i(s)$ denote the total amount of transfer to type i members. Note that it depends only on the number of supporters, not on the number of members. It seems intuitively reasonable to allocate equal transfers to member countries of the same type.

The internal stability requires that neither types of member country can increase its payoff if it withdraws:

$$\pi_M^1(m_1^*, m_2^*) + \frac{T_1(s)}{m_1^*} \geq \pi_F^1(m_1^* - 1, m_2^*) \tag{2}$$

$$\pi_M^2(m_1^*, m_2^*) + \frac{T_2(s)}{m_2^*} \geq \pi_F^2(m_1^*, m_2^* - 1), \tag{3}$$

where π_M^i and π_F^i denote the payoff of type i members (without transfers) and free riders, respectively.

The external stability conditions show that both types of free rider will suffer a loss if it enters into the agreement:

$$\pi_F^1(m_1^*, m_2^*) > \pi_M^1(m_1^* + 1, m_2^*) + \frac{T_1(s)}{m_1^* + 1}, \tag{4}$$

$$\pi_F^2(m_1^*, m_2^*) > \pi_M^2(m_1^*, m_2^* + 1) + \frac{T_2(s)}{m_2^* + 1}. \tag{5}$$

Based on (2) and (3), we can define the minimum transfers to stabilize a coalition (m_1, m_2) as

$$\tau_1(m_1, m_2) = m_1 \left(\pi_F^1(m_1 - 1, m_2) - \pi_M^1(m_1, m_2)\right), \tag{6}$$

$$\tau_2(m_1, m_2) = m_2 \left(\pi_F^2(m_1, m_2 - 1) - \pi_M^2(m_1, m_2)\right). \tag{7}$$

Note that $\frac{\tau_i(m_1,m_2)}{m_i}$ is equal to the net benefit of a type i country from withdrawing the agreement joined by m_1 type 1 countries and m_2 type 2 countries. In what follows, we write m_i^* as a function of transfers: $m_i^*(T_1(s), T_2(s))$.

In stage 2, the number of supporters will be decided and if it is not smaller than the minimum participation number c, the supporter coalition will determine the transfers to each type of member countries. Since no supporter wants to pay more transfers than the rules have required, the amount of transfers should be the minimum transfer to induce an agreement including all countries except for supporters, i.e., for $s \neq 0$,

$$T_i(s) = \begin{cases} \tau_i(n_1, n_2 - s) & (s \geq c) \\ 0 & (s < c) \end{cases} \quad (i = 1, 2),$$

and $T_i(0) = 0$ since no transfers are made without supporters.

Subsequently, all type 2 countries decide whether to become supporters or not. Choosing not to become a supporter means becoming an IEA member or a free rider in the equilibrium of the subgame following stage 2. Henceforth, their decisions depend on the associated payoff of being supporters and on that of not being supporters. The equilibrium number of supporters, denoted as s^*, implies that no supporter deviates to become non-supporter country and no non-supporter country deviates to become supporter, taking solutions after stage 2 into consideration. Therefore, the following internal and external stability conditions for the supporter coalition must hold.

Internal Stability

If a supporter country of type 2 deviates to become a free rider, its payoff with $s^* - 1$ supporters should be no higher than its supporter payoff with s^* supporters:

$$\begin{aligned}\pi_F^2\left(m_1^*(T_1(s^*),T_2(s^*)),m_2^*(T_1(s^*),T_2(s^*))\right) - \frac{T_1(s^*)+T_2(s^*)}{s^*}\\ \geq \pi_F^2\left(m_1^*(T_1(s^*-1),T_2(s^*-1)),m_2^*(T_1(s^*-1),T_2(s^*-1))\right).\end{aligned} \tag{8}$$

If a supporter country of type 2 deviates to become an IEA member, its payoff with $s^* - 1$ supporters should be no higher than its supporter payoff with s^* supporters:

$$\begin{aligned}\pi_F^2\left(m_1^*(T_1(s^*),T_2(s^*)),m_2^*(T_1(s^*),T_2(s^*))\right) - \frac{T_1(s^*)+T_2(s^*)}{s^*}\\ \geq \pi_M^2\left(m_1^*(T_1(s^*-1),T_2(s^*-1)),m_2^*(T_1(s^*-1),T_2(s^*-1))\right)\\ + \frac{T_2(s^*-1)}{m_2^*(T_1(s^*-1),T_2(s^*-1))}.\end{aligned} \tag{9}$$

The internal stability conditions say that a type 2 country will become a supporter if the associated payoff is at least as high as what can be gained in the subgame where it becomes a member or a free rider.

External Stability

For a type 2 non-supporting country expected to become a free rider in the subgame, its supporter payoff with $s^* + 1$ supporters is lower than its free rider payoff with s^* supporters:

$$\begin{aligned}
&\pi_F^2\left(m_1^*(T_1(s^*), T_2(s^*)), m_2^*(T_1(s^*), T_2(s^*))\right) \\
&\quad > \pi_F^2\left(m_1^*(T_1(s^*+1), T_2(s^*+1)), m_2^*(T_1(s^*+1), T_2(s^*+1))\right) \\
&\qquad - \frac{T_1(s^*+1) + T_2(s^*+1)}{s^*+1}. \qquad (10)
\end{aligned}$$

For a type 2 non-supporting country expected to become an IEA member in the subgame, its supporter payoff with s^*+1 supporters is lower than its IEA member payoff with s^* supporters:

$$\begin{aligned}
&\pi_M^2\left(m_1^*(T_1(s^*), T_2(s^*)), m_2^*(T_1(s^*), T_2(s^*))\right) + \frac{T_2(s^*)}{m_2^*} \\
&\quad > \pi_F^2\left(m_1^*(T_1(s^*+1), T_2(s^*+1)), m_2^*(T_1(s^*+1), T_2(s^*+1))\right) \\
&\qquad - \frac{T_1(s^*+1) + T_2(s^*+1)}{s^*+1}. \qquad (11)
\end{aligned}$$

The external stability conditions say that a type 2 country will not become a supporter if the associated payoff is higher than the payoff it can gain when it becomes a supporter in the subgame. Let us denote s^* satisfying (8)–(11) as a function of the minimum participation number: $s^*(c)$.

In stage 1, type 2 countries decide which minimum participation number "c" is optimal via unanimous voting. Two aspects need to be considered. First, each type 2 country should prefer the equilibrium under the MPR to the one without it. This condition defines the set of MPRs that would be unanimously voted by all type 2 countries. To get the rule c approved unanimously, whatever decisions a type 2 country makes in stages 2 and 3, its lowest possible payoff when the MPR is introduced should be no less than its highest possible payoff it would receive in the absence of the concensus on the MPR. In other words, its supporter payoff should be no less than its payoff as a free rider without the supporter scheme, which is implied by (13) below. Second, the MPR in this set should be optimal in the sense that it maximizes the type 2 country's expected payoff in stage 1. We use expected payoff because in this stage each type 2 country does not know whether it will be a supporter, member, or free rider in the following stages. Given symmetry, the final decision can be simplified to a type 2 country's optimal choice by solving the following:

$$\begin{aligned}
&\max_c E\pi_2(c)\\
&\quad = \frac{s^*(c)}{n_2} \times \Big\{\pi_F^2\left(m_1^*(T_1(s^*(c)), T_2(s^*(c))), m_2^*(T_1(s^*(c)), T_2(s^*(c)))\right)\\
&\qquad\qquad - \frac{T_1(s^*(c)) + T_2(s^*(c))}{s^*(c)}\Big\}\\
&\qquad + \frac{m_2^*(T_1(s^*(c)), T_2(s^*(c)))}{n_2}\\
&\qquad \times \Big\{\pi_M^2\left(m_1^*(T_1(s^*(c)), T_2(s^*(c))), m_2^*(T_1(s^*(c)), T_2(s^*(c)))\right)\\
&\qquad\qquad + \frac{T_2(s^*(c))}{m_2^*(T_1(s^*(c)), T_2(s^*(c)))}\Big\}\\
&\qquad + \frac{n_2 - s^*(c) - m_2^*\left(T_1(s^*(c)), T_2(s^*(c))\right)}{n_2}\\
&\qquad \times \pi_F^2\left(m_1^*(T_1(s^*(c)), T_2(s^*(c))), m_2^*(T_1(s^*(c)), T_2(s^*(c)))\right), \qquad (12)\\
&\text{s.t. } \pi_F^2\left(m_1^*(T_1(s^*(c)), T_2(s^*(c))), m_2^*(T_1(s^*(c)), T_2(s^*(c)))\right)\\
&\qquad - \frac{T_1(s^*(c)) + T_2(s^*(c))}{s^*(c)}\\
&\quad \geq \pi_F^2\left(m_1^*(T_1(0), T_2(0)), m_2^*(T_1(0), T_2(0))\right), \qquad (13)
\end{aligned}$$

where $E\pi_2(c)$ in Eq. (12) is a type 2 country's expected payoff in the first stage; $s^*(c)$ is an equilibrium number of supporters given the value of c determined in the first stage; $\frac{s^*(c)}{n_2}$, $\frac{m_2^*(T_1(s^*(c)),T_2(s^*(c)))}{n_2}$, and $\frac{n_2-s^*(c)-m_2^*(T_1(s^*(c)),T_2(s^*(c)))}{n_2}$ are the probability of it being a supporter, member, and free rider in stage 3, respectively. Let us denote the solution to the above maximization problem as c^*.

Existence of Equilibria with Supporters

In the previous section, we have described the details of the game and how to solve it. The following proposition shows the existence of the SPNE.

Proposition 1. *Suppose that there exist some integers c $(1 \leq c \leq n_2)$ satisfying*[2]

$$c \geq \left\lceil \frac{T_1(c) + T_2(c)}{\pi_F^2(n_1, n_2 - c) - \pi_F^2(m_1^*(0,0), m_2^*(0,0))} \right\rceil. \tag{14}$$

Let $C(\neq \emptyset)$ denote the set of c satisfying (14). Subsequently, there exists a unique SPNE in which a minimum participation number is $c^* \in C$, $s^*(c^*) = c^*$ supporters will pay transfers as $T_1(c^*) = \tau_1(n_1, n_2 - c^*)$ and $T_2(c^*) = \tau_2(n_1, n_2 - c^*)$, and the number of member countries is $(m_1^*(T_1(c^*), T_2(c^*)), m_2^*(T_1(c^*), T_2(c^*))) = (n_1, n_2 - c^*)$.

Proof. See Appendix. □

Proposition 1 establishes that using external transfers is effective in increasing participation in IEAs under MPR. The number of supporters equals the minimum participation number while the remaining countries are IEA members. Note that we only focus on equilibria with no free riders, because if the number of supporters is given, the existence of less free riders means more efficiency. If the amount of transfers becomes lower, we will have equilibria with free riders.

Rearrangement of condition (14) in Proposition 1 states that what supporter coalition gains from the enlargement of IEAs should be no less than their transfer payment, which is the minimum transfer required to internally stabilize the IEA that includes all countries except for supporters. It is not too restrictive, as we will present rough analyses below and demonstrate with specific parameter values in the simulation results in the next section.

Let us roughly evaluate the value of the right-hand side of (14) to predict the effect of parameter values on the equilibrium number of supporters. Assume that $\beta_1 = \beta$, $\beta_2 = 1$, $\gamma_1 = \gamma$, and $\gamma_2 = 1 (\beta, \gamma \in (0,1))$. Abatement levels are $q_M^1 = \frac{\beta n_1 + n_2 - c}{\gamma}$, $q_F^1 = \frac{\beta}{\gamma}$, $q_M^2 = \beta n_1 + n_2 - c$, and $q_F^2 = 1$. From the definition of the minimum transfers, we have $T_1(c) = \tau_1(n_1, n_2 - c) = n_1(\pi_F^1(n_1 - 1, n_2 - c) - \pi_M^1(n_1, n_2 - c))$. Considering that the cost reduction is usually much larger than the benefit reduction when a member deviates to become a free rider and that the abatement costs of free riders is nearly

[2] $\lceil x \rceil$ is the minimum integer which equals or is larger than x.

zero, we have $T_1(c) \approx n_1 \cdot \frac{\gamma}{2}(q_M^1)^2 = \frac{n_1(\beta n_1+n_2-c)^2}{2\gamma}$. Similarly, we have $T_2(c) \approx (n_2 - c) \cdot \frac{1}{2}(q_M^2)^2 = \frac{(n_2-c)(\beta n_1+n_2-c)^2}{2}$. Next, $\pi_F^2(n_1, n_2 - c) - \pi_F^2(m_1^*(0,0), m_2^*(0,0))$ is approximately equal to the benefit from the total abatement of n_1+n_2-c members. Thus, we have $\pi_F^2(n_1, n_2-c) - \pi_F^2(m_1^*(0,0), m_2^*(0,0)) \approx n_1\ q_M^1 + (n_2 - c)q_M^2 = \left(\frac{n_1}{\gamma} + n_2 - c\right)(\beta n_1 + n_2 - c)$. From above, $\frac{T_1(c)+T_2(c)}{\pi_F^2(n_1,n_2-c)-\pi_F^2(m_1^*(0,0),m_2^*(0,0))} \approx \frac{\beta n_1+n_2-c}{2}$, thus, we can see that the right-hand side of (14) is affected little by the cost parameters. The solution of $c \geq \frac{\beta n_1+n_2-c}{2}$ is $c \geq \frac{\beta n_1+n_2}{3}$, which means that if the benefit heterogeneity increases (β is smaller), the lower bound of the range of c satisfying (14) is likely to decrease. Let $c' \equiv \left[\frac{\beta n_1+n_2}{3}\right]$.

However, we must note that c' can be different from c^* determined in stage 1. Let us consider the effects of c on the expected payoff of type 2 countries at equilibrium. A supporter's payoff is $\pi_F^2(n_1, n_2 - c) - \frac{T_1(c)+T_2(c)}{c}$ which is nearly equal to the equilibrium payoff in the no-transfer case when $c = c'$. A type 2 member's payoff is $\pi_F^2(n_1, n_2-c-1)$. As c increases, the supporter's payoff will increase[3] because of less transfers, whereas the member's payoff will decrease. Although the optimal $c^* \in [c', n_2]$ which maximizes $E\pi_2(c)$ can be derived only by conducting numerical analyses, we can expect larger c^* when γ is smaller. The reason is that with small γ, the cost burden of supporters due to payment of transfers is larger, and the effect of increase in c on the payoff of the supporter is greater than that on the payoff of the member. Similarly, when β is smaller, the cost burden of supporters is smaller, therefore we expect smaller value of c^* which is close to c'. We will confirm these facts using a simulation in the next section.

The Impact of Heterogeneity on Equilibria

Given the existence of equilibria, we analyze the impact of heterogeneity on the result based on the simulation method (using MATLAB software).[4] To separate the effects of heterogeneity in the abatement costs and the benefits, we first assume that countries have the same abatement

[3]The proof is available upon request.
[4]All programs used in this section are available upon request.

benefit parameter and then the same abatement cost parameter. We consider 20 type 1 countries and varying numbers (integers between 12 and 20) of type 2 countries.[5]

The Heterogeneity in Abatement Costs

Following the analyses in the last section, we assume that $\beta_1 = \beta_2 = 1$, $\gamma_2 = 1$, and γ_1 varies between 0 and 1 by taking values of 10,000 linearly spaced numbers in the interval (0,1). Note that the smaller the value of γ, which equals $\frac{\gamma_1}{\gamma_2}$, the higher the heterogeneity in the abatement costs. Given each set of n_2 and γ, the program finds the equilibrium minimum participation number, c^*, which satisfies (14) and maximizes (12).

Figure 1 presents the relationship between the degree of abatement cost heterogeneity and the minimum participation number.[6] In general, the degree of heterogeneity has no impact on the minimum participation number and consequently on the equilibrium numbers of supporters and member countries. It seems that the value of c^* equals either c' or n_2 in all cases. We can also see that when the number of type 2 countries is 15 and 16, smaller γ or higher heterogeneity results in larger c^*. These results are consistent with the prediction in the last section. As the number of type 2 countries increases, a smaller share of them become supporters and the number of supporters does not change much. This is because the gain an extra supporter can obtain from an enlarged IEA can compensate much of the increased burden, and consequently much more supporters are unnecessary to stabilize the enlarged IEA when there are more type 2 countries. Henceforth, the share of supporters decreases.

A common way to measure the efficiency of IEA is relative gains, calculated as $\frac{\pi(n_1,n_2-c^*)-\pi(0,0)}{\pi(n_1,n_2)-\pi(0,0)}$, where $\pi(m_1, m_2)$ is all countries' total payoff under coalition (m_1, m_2). Figure 2 shows the impact of the

---

[5]We limit the number of type 2 countries to save space.

[6]Even though there are two equilibria in the no-transfer case as proved in Fuentes-Albero and Rubio (2010), i.e., $(m_1^*(0,0), m_2^*(0,0)) = (3,0)$ and $(m_1^*(0,0), m_2^*(0,0)) = (0,3)$, our program suggests the same figures, which can be explained by the sufficiently small values of $\pi_F^2(m_1^*(0,0), m_2^*(0,0))$ compared with $\pi_F^2(n_1, n_2 - c)$. Therefore, which equilibrium is chosen will not affect the final equilibrium result concerning c.

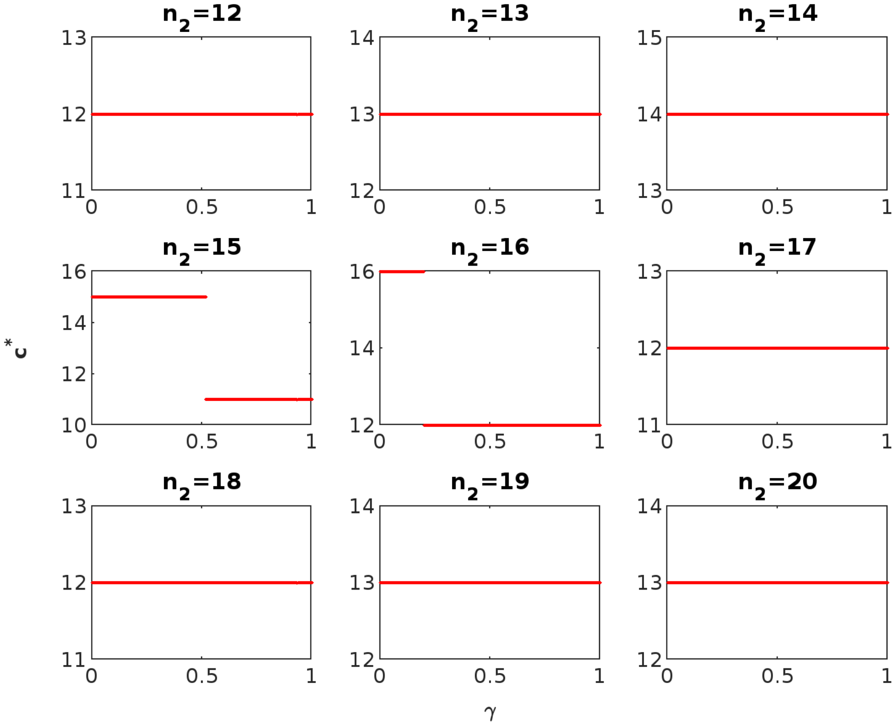

Figure 1: Equilibrium minimum participation number under different degrees of abatement cost heterogeneity.

abatement cost heterogeneity on relative gains. Obviously, the smaller the γ, the higher the relative gains. We summarize the results as follows.

Result 1. *Under the framework of external transfers, the higher the heterogeneity in the abatement costs, the higher the relative gains of a self-enforcing IEA while its composition is generally unchanged although in some cases higher heterogeneity leads to more supporters.*

This partly confirms the result in Fuentes-Albero and Rubio (2010) that if abatement costs represent the only difference among countries, the degree of heterogeneity has no impact on the composition of IEAs. However, we find that external transfers can enlarge the size of IEAs

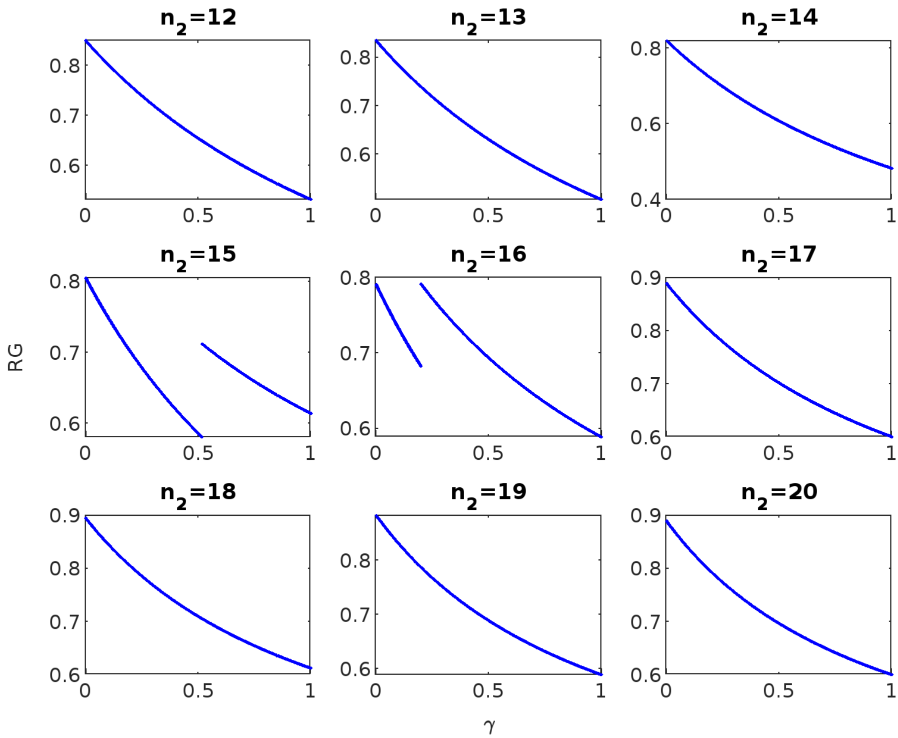

Figure 2: Relative gains of IEAs with supporters under different degrees of abatement cost heterogeneity.

compared to the no-transfer case, which Fuentes-Albero and Rubio (2010) did not find under internal transfers.

Figure 3 illustrates individual country's payoff gain rate, which is the percentage change in each country's payoff from no-transfer case to transfer case. R_M^1, R_M^2 and R_S refer to type 1 member, type 2 member and supporter's payoff gain rate, separately. Obviously, supporters always gain the least from cooperation.[7]

[7] We do not consider a type 2 member's payoff gain rate, when it was a free rider in the no-transfer case, since the result is close to R_M^2 given the reason explained in Footnote 6.

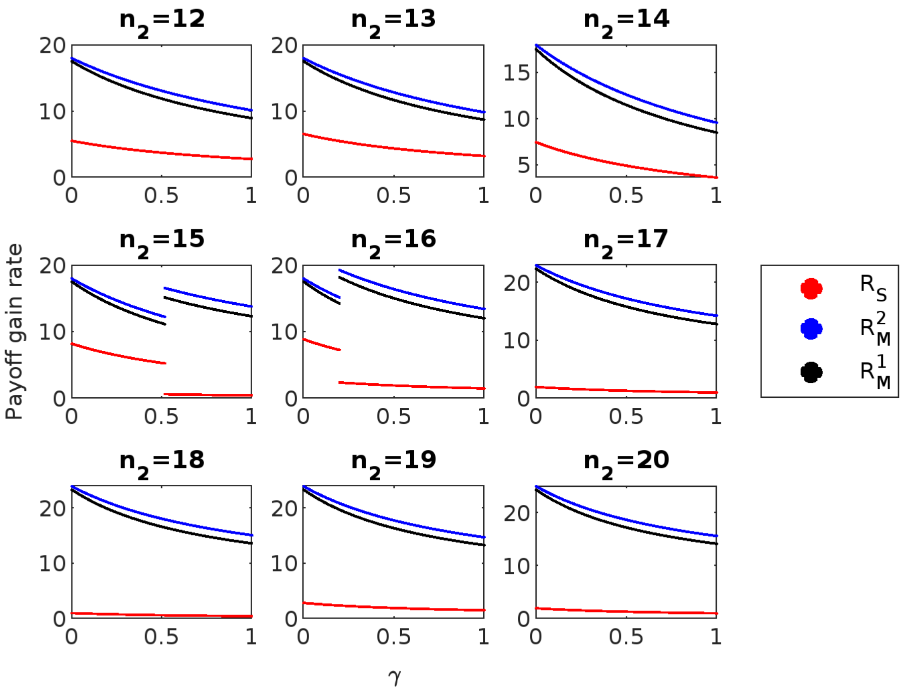

Figure 3: Individual payoff gain rate of three categories of countries under different degrees of abatement cost heterogeneity.

The Heterogeneity in Abatement Benefits

In what follows, we investigate the agreement with transfers when type 1 and type 2 countries differ only in abatement benefits. Other assumptions remain, $\gamma_1 = \gamma_2 = 1$, $\beta_2 = 1$, β_1 varies between 0 and 1 by taking values of 10,000 linearly spaced numbers in the interval $(0, 1)$.

Figure 4 shows the relationship between the abatement benefit heterogeneity and the equilibrium minimum participation number. It is clear that in general, the higher the heterogeneity in abatement benefits, the smaller the equilibrium number of supporters and consequently the larger the size of IEAs. The value of c^* coincides with the lower bound c' under higher heterogeneity as we expect, but there are cases where $c^* = n_2$ holds with lower heterogeneity.

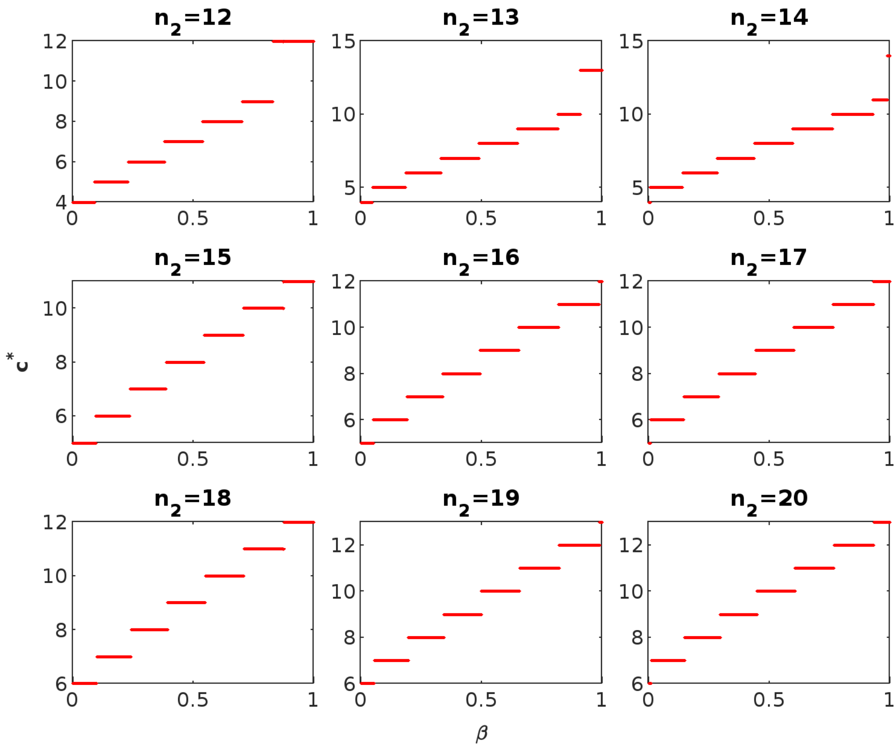

Figure 4: Equilibrium minimum participation number under different degrees of abatement benefit heterogeneity.

Figure 5 shows the relationship between the abatement benefit heterogeneity and the relative gains. We can see that the smaller the β, the higher the relative gains. This result is explained by the fact that greater heterogeneity leads to fewer supporters and more members. The analysis above provides the following result.

Result 2. *Under the framework of external transfers, the higher the heterogeneity in the abatement benefit, the larger the size of a self-enforcing IEA and the higher its relative gains.*

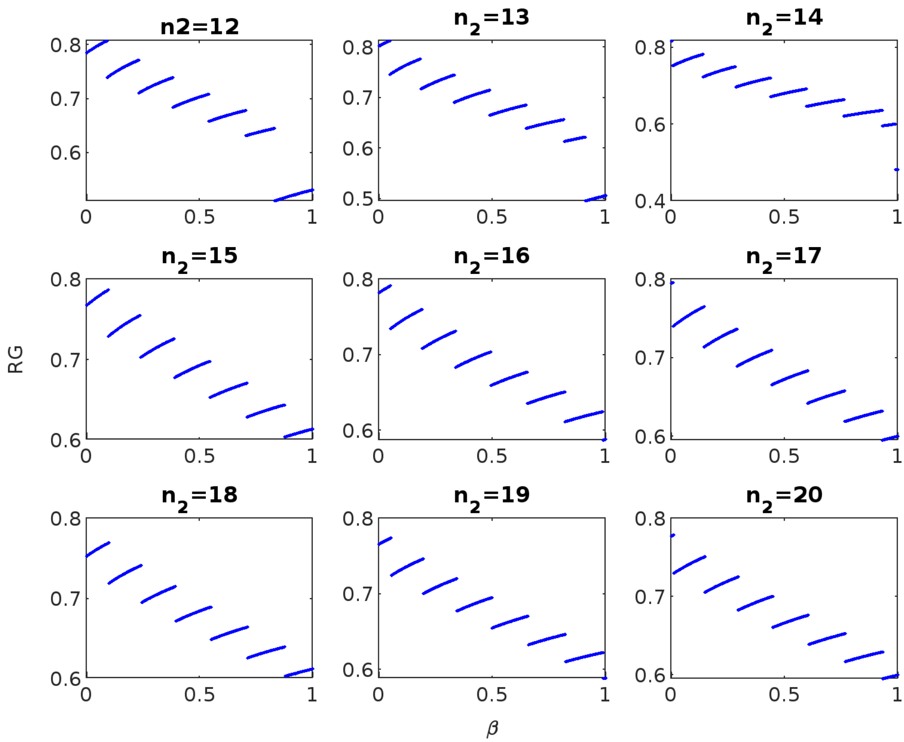

Figure 5: Relative gains of IEAs with supporters under different degrees of abatement benefit heterogeneity.

Result 2 is similar to Fuentes-Albero and Rubio (2010), which shows that the size of the IEA increases with the heterogeneity in the abatement benefit provided that there are only one or two high-benefit member countries. Moreover, Result 2 generalizes such relationship since in the context of external transfers, no limitation is placed on the number of high-benefit member countries.

In addition, Figure 6 indicates that supporter's payoff gain rate is lowest comparing with the member countries, which gives similar result to that in Figure 3.

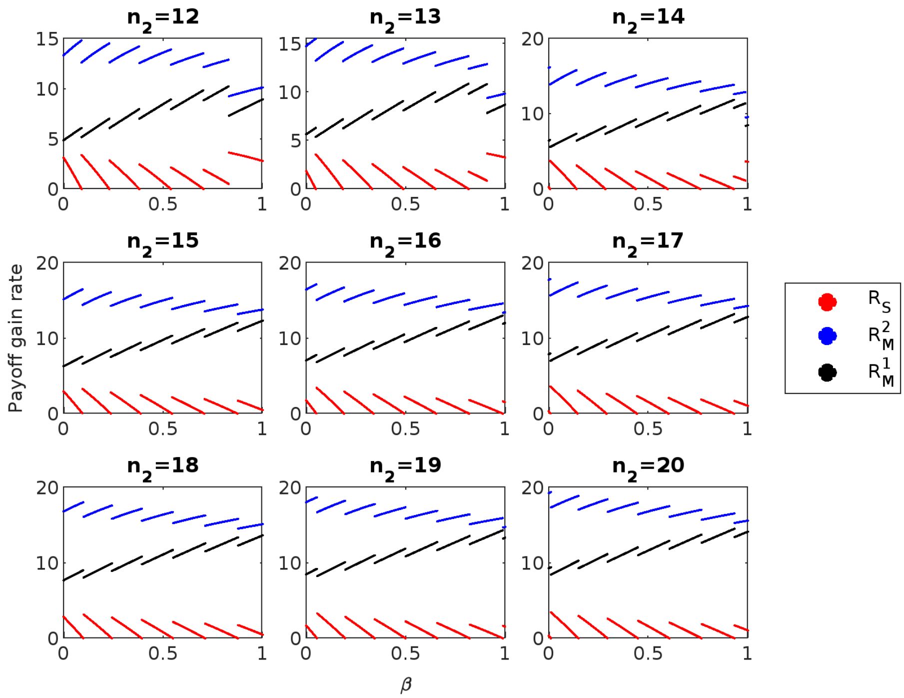

Figure 6: Individual payoff gain rate of three categories of countries under different degrees of abatement benefit heterogeneity.

Discussion and Conclusion

This study explores the efficiency of a supporter scheme (external transfers) in IEAs when countries are asymmetric and investigates the impact of the degree of heterogeneity on the size and on the efficiency of self-enforcing IEAs. Supporters are the countries that commit to transfer welfare to stabilize a coalition including all the remaining countries. The supporter coalition is formed through an MPR that is unanimously voted by all type 2 countries which have larger abatement cost and benefit parameters than type 1 countries.

Our results show that an equilibrium exists and external transfers are effective in encouraging participants in IEAs under asymmetry. When heterogeneity exists only in abatement cost, the degree of heterogeneity has no impact on the number of supporters in most cases, but there are some cases where higher heterogeneity leads to more supporters. As for relative gains, they get higher with higher heterogeneity. In addition, external transfers can encourage participants under such asymmetry assumption, which internal transfers cannot do, as pointed out by Fuentes-Albero and Rubio (2010).

The story changes concerning abatement benefit heterogeneity. We find that, in general, the higher the heterogeneity, the lower the number of supporters that are needed and consequently the larger the size and the higher the efficiency of the IEA. These parallel the results in the internal transfers case (Fuentes-Albero and Rubio, 2010).

When participation in an environmental coalition becomes a huge barrier to the implementation of IEAs, transfers as an incentive to encourage participants seem to be necessary. The support scheme introduced in our work provides a new direction for international cooperation. Assuming one-sided asymmetry, external transfers have a larger potential to encourage cooperation than internal transfers do. In reality, if some industrialized countries become supporters and make transfer commitments, more countries will participate into the IEA.

The future challenges would be to confirm this conclusion in different functional forms and to consider different parameter sets. As pointed out by Ansink *et al.* (2019), under quadratic-quadratic specification where the abatement benefit function is quadratic and concave, the leakage exists, which means that non-member countries will decrease their abatement levels when the size of the IEA is enlarged. However, since this leakage is not complete, the total abatement level increases. This makes being a non-member country more attractive since its benefits increase and costs decrease. Henceforth, it becomes harder for supporters to stabilize large IEAs, and the equilibrium number of supporters is expected to increase. Moreover, it would be interesting to consider incorporating altruism or preferences for equity into the countries' payoffs (Lange and Vogt, 2003; Mason, 2022).

In addition, we must note that the results in this study depend on the assumption that a type 1 country has lower cost and benefit parameters than type 2 ($\gamma < 1$ and $\beta < 1$). Let us consider if these assumptions are relaxed. If γ is much higher than 1, the equilibrium number of supporters c^* will be close to the lower bound c' because the supporters' payoff does not increase much as c increases. If β is sufficiently high, then the supporter scheme will not work because the supporters will gain less than the no-transfer case.

Appendix: Proof of Proposition 1

We prove Proposition 1 by showing that c^*, $s^*(c) = c$ and $m_1^*(T_1(s), T_2(s)) = n_1$, and $m_2^*(T_1(s), T_2(s)) = n_2 - s$ $(s \geq c)$ satisfy conditions (2)–(5) and conditions (8)–(13).

It is obvious that given the transfers, the member countries' payoffs are equal to their payoffs when they deviate to become free riders. Therefore, conditions (2) and (3) are satisfied. Since there is no free rider, (4) and (5) need not be considered. Thus, we have $m_1^*(T_1(s), T_2(s)) = n_1$, and $m_2^*(T_1(s), T_2(s)) = n_2 - s$ $(s \geq c)$ as a solution for stage 3.

Let us turn our attention to stage 2. Suppose that the minimum participation number c is given. From the rule, $T_1(c-1) = T_2(c-1) = 0$. In this case, (9) is implied by (8) given that the left-hand sides of both are the same while on the right-hand side, $\pi_M^2(m_1^*(0,0), m_2^*(0,0)) < \pi_F^2(m_1^*(0,0), m_2^*(0,0))$. According to the solution for stage 3, we do not need to consider (10) because no country becomes a free rider. When $c = n_2$, we do not need to consider (11). When $c \leq n_2 - 1$, (11) is always satisfied because it can be rewritten as $\pi_M^2(n_1, n_2 - c) + \pi_F^2(n_1, n_2 - c - 1) - \pi_M^2(n_1, n_2 - c) = \pi_F^2(n_1, n_2 - c - 1) \geq \pi_F^2(n_1, n_2 - c - 1) - \frac{T_1(c+1)+T_2(c+1)}{c+1}$. Therefore, the only binding condition for $s^*(c) = c$ to be a solution for stage 2 is (8).

Given $c = c^*$, $s^*(c^*) = c^*$, and $(m_1^*(T(c^*), m_2^*(T(c^*)) = (n_1, n_2 - c^*)$, constraint (13) becomes $\pi_F^2(n_1, n_2 - c^*) - \frac{T_1(c^*)+T_2(c^*)}{c^*} \geq \pi_F^2(m_1^*(0,0), m_2^*(0,0))$, which coincides with (8). Therefore, if there exist some integers satisfying (8), i.e., $C \neq \emptyset$, from them we can choose an integer that maximizes (12). This is how the value of c^* is determined. The simple rearrangement of (8) or (13) leads to condition (14) in Proposition 1.

We can show that when c^* satisfies (8) or (13), $s^*(c^*) = c^*$ is a unique solution for stage 2 after the announcement of c^* as follows. If $s^*(c^*) = c \geq c^* + 1$, then (8) is irrelevant, but (9) becomes $\pi_F^2(n_1, n_2 - c) - \frac{T_1(c)+T_2(c)}{c} \geq \pi_F^2(n_1, n_2 - c)$, which does not hold. If $s^*(c^*) = c^* - 1$, then (11) is violated because (11) becomes $\pi_M^2(m_1^*(0,0), m_2^*(0,0)) > \pi_F^2(n_1, n_2 - c^*) - \frac{T_1(c^*)+T_2(c^*)}{c^*}$, which contradicts (8) with $s^* = c^*$. If $s^*(c^*) \leq c^* - 2$, then both (10) and (11) are violated.

References

Ansink, E., H.-P. Weikard, and C. Withagen (2019), "International Environmental Agreements with Support", *Journal of Environmental Economics and Management*, 97, 241–52, DOI: 10.1016/j.jeem.2018.02.001.

Bakalova, I. and J. Eyckmans (2019), "Simulating the Impact of Heterogeneity on Stability and Effectiveness of International Environmental Agreements", *European Journal of Operational Research*, 277(3), 1151–62, DOI: 10.1016/j.ejor.2019.03.028.

Barrett, S. (1994), "Self-Enforcing International Environmental Agreements", *Oxford Economic Papers*, 46, 878–94, DOI: 10.1093/oep/46.Supplement1.878.

Carraro, C., C. Marchiori, and S. Oreffice (2009), "Endogenous Minimum Participation in International Environmental Treaties", *Environmental and Resource Economics*, 42(3), 411–25, DOI: 10.1007/s10640-009-9264-x.

Carraro, C. and D. Siniscalco (1993), "Strategies for the International Protection of the Environment", *Journal of Public Economics*, 52(3), 309–28, DOI: 10.1016/0047-2727(93)90037-T.

D'Aspremont, C., A. Jacquemin, J. J. Gabszewicz, and J. A. Weymark (1983), "On the Stability of Collusive Price Leadership", *Canadian Journal of Economics*, 16(1), 17–25, DOI: 10.2307/134972.

Diamantoudi, E., E. S. Sartzetakis, and S. Strantza (2018a), "International Environmental Agreements — Stability with Transfers among Countries", (May 25, 2018), *Fondazione Eni Enrico Mattei Working Papers*, Paper 1245.

Diamantoudi, E., E. S. Sartzetakis, and S. Strantza (2018b), "International Environmental Agreements — The Impact of Heterogeneity among Countries on Stability", (June 21, 2018), *Fondazione Eni Enrico Mattei Working Papers*, Paper 1247.

Finus, M. and M. McGinty (2019), "The Anti-paradox of Cooperation: Diversity May Pay!", *Journal of Economic Behavior and Organization*, 157, 541–59, DOI: 10.1016/j.jebo.2018.10.015.

Fuentes-Albero, C. and S. J. Rubio (2010), "Can International Environmental Cooperation be Bought?", *European Journal of Operational Research*, 202(1), 255–64, DOI: 10.1016/j.ejor.2009.05.006.

Lange, A. and C. Vogt (2003), "Cooperation in International Environmental Negotiations due to a Preference for Equity", *Journal of Public Economics*, 87(9–10), 2049–67, DOI: 10.1016/S0047-2727(02)00044-0.

Li, Q. and T. Fujita (2021), "Transfers in International Environmental Agreements under Heterogeneity", *Applied Economics*, 53(33), 3899–910, DOI: 10.1080/00036846.2020.1861209.

Mason, C. F. (2022), "Cooperation in Dynamic Games with Asymmetric Players: The Role of Social Preferences", *Dynamic Games and Applications*, 12(3), 977–95, DOI: 10.1007/s13235-022-00435-1.

McGinty, M. (2007), "International Environmental Agreements among Asymmetric Nations", *Oxford Economic Papers*, 59(1), 45–62, DOI: 10.1093/oep/gpl028.

McGinty, M. (2020), "Leadership and Free-riding: Decomposing and Explaining the Paradox of Cooperation in International Environmental Agreements", *Environmental and Resource Economics*, 77(2), 449–74, DOI: 10.1007/s10640-020-00505-1.

Pavlova, Y. and A. de Zeeuw (2013), "Asymmetries in International Environmental Agreements", *Environmental and Development Economics*, 18(1), 51–68, DOI: 10.1017/s1355770x12000289.

Rutz, S. (2001), "Minimum Participation Rules and the Effectiveness of Multilateral Environmental Agreements", Working Paper 01/22, Center for Economic Research, SFIT, Zurich. DOI: 10.3929/ethz-a-004296584.

Weikard, H.-P., L. Wangler, and A. Freytag (2015), "Minimum Participation Rules with Heterogeneous Countries", *Environmental and Resource Economics*, 62(4), 711–27, DOI: 10.1007/s10640-014-9861-1.