KIER DISCUSSION PAPER SERIES

KYOTO INSTITUTE
OF
ECONOMIC RESEARCH

Discussion Paper No.1026

"Free Trade Agreements with Environmental Standards"

Hideo Konishi Minoru Nakada Akihisa Shibata

March 2020 (Revised: April 2020)

KYOTO UNIVERSITY

KYOTO, JAPAN

Free Trade Agreements with Environmental Standards*

Hideo Konishi† Minoru Nakada‡ Akihisa Shibata§

March 13, 2020

Abstract

In this paper, we investigate the effects of a free trade agreement (FTA) with environmental standards between Northern and Southern countries, with explicit considerations for transferring clean technology and enforcing reduced emissions. Southern producers benefit greatly from having unimpeded access to a Northern market, but they are reluctant to use new high-cost, clean technology provided by the North. Thus, environmentally conscious Northern countries should design an FTA where Southern countries are provided with sufficient membership benefits but must follow tighter enforcement requirements. Since including too many Southern countries dilutes the benefits of FTA membership, it is in the best interest of the North to limit the number of Southern memberships while strictly enforcing emissions reduction. This may result in unequal treatment among the Southern countries. We provide a quantitative evaluation of FTA policies using a numerical example.

Keywords: Free trade agreements; Environmental standards.

JEL Classification: Q56, F53

---

*This research was financially supported by JSPS Grant-in-Aid for Scientific Research (No.17H02516) and the joint research programs of KIER (Kyoto University) and ISER (Osaka University).

†Hideo Konishi: Department of Economics, Boston College, USA. Email: hideo.konishi@bc.edu

‡Minoru Nakada: Corresponding author, Graduate School of Environmental Studies, Nagoya University, B4-1(780) Furo Chikusa Nagoya 464-8601, Japan. Email: mnakada@cc.nagoya-u.ac.jp, Tel: +81-(0)52-789-4733, Fax: +81-(0)52-789-4820

§Akihisa Shibata: Institute of Economic Research, Kyoto University, Japan. Email: shibata@kier.kyoto-u.ac.jp

1 Introduction

In the current era of globalization, the natural environment of a country is significantly affected by other countries' economic activities and abatement efforts. Therefore, when it comes to fixing international environmental problems such as acid rain, ozone layer depletion, and climate change, multinational negotiation among nations has become an essential way to reach agreements on how to handle these issues. In the negotiation processes, countries that share common interests may organize themselves into groups to strengthen their bargaining power. In the case of climate change, during the negotiation process under the UN Framework Convention on Climate Change (UNFCCC), nations have been organized into seven groups.[1] However, there are many difficulties countries could face in organizing themselves this way. They may differ on the degree of potential damages and on the levels of costs and benefits in pollution abatement activities, even if they share general common interests in the region. Income levels may be the fundamental factor, as countries can have different priorities in terms of their GDP, living standards, and the quality of their environment. Thus, negotiations between groups with different income levels, such as between developed and developing countries, take longer to establish multinational environmental agreements (MEAs).[2]

In contrast, there is an increasing number of free-trade agreements (FTAs) between developed and developing countries.[3] According to the World Trade Organization Database (WTO 2019), as of June 26, 2019, 294 trade agreements are in effect. Around 10% of these

---

[1] These are the African Group, the Arab States, the Environmental Integrity Group, the European Union, Least Developed Countries, the Small Island Developing States, and the Umbrella Group (UNFCCC 2019).

[2] According to Bodansky (2016), in the case of climate change, negotiations to seek a second commitment period on the Kyoto Protocol first began in 2005 and, next, discussions for the Post-Kyoto Protocol framework began at the Bali conference in 2007; both of these were intended to be concluded at the Copenhagen conference in 2009. In fact, it took more than ten years from the start of negotiations to reach the Paris Agreement in 2016.

[3] Article XXIV of GATT provides an exception for regional trade agreements (RTAs), including FTAs and customs unions (CUs). This treatment appears contrary to the most-favoured-nation (MFN) clause and multilateralism, which require a unanimous agreement among the WTO members. However, the WTO seems to think that having RTAs helps accelerate negotiations between regional groups and eliminate the tariffs and non-tariff barriers that the WTO has long been trying to remove or at least reduce. Similarly, FTAs might help in achieving some of the objectives of MEAs if they are associated with environmental provisions - this is because there is empirical evidence that RTAs with environmental provisions contribute to emission reductions (Baghdadi et al., 2013).

agreements are between developed countries, while close to 30% are between developed and developing countries (Behar and Cirera-i-Criville 2013). Forming an FTA expands the pie for its member countries—for developing countries, access to markets in developed countries is a lucrative reward. As long as the number of countries participating in FTAs is limited, this reward may be attractive to developing countries. If these rewards can be used as negotiation tools for MEAs among countries with different income levels, it may make sense to utilize an FTA with environmental provisions as an alternative framework to address transboundary environmental problems.[4] In the case of the North American Free Trade Agreement (NAFTA), the North American Agreement on Environmental Cooperation (NAAEC) was ratified by Canada, Mexico, and the United States. Although it allows each country to establish its own level of domestic environmental protection, the agreement requires member countries to conserve the environment; moreover, it is recommended that each country provides high levels of environmental protection (CEC 1993).

In FTAs with environmental provisions, the transfer of clean technologies from developed to developing countries is essential for effective agreements. Gutierrez and Teshima (2018) pointed out the importance of the technology upgrades induced by NAFTA for pollution reduction in Mexico. Such evidence highlights the importance of providing developing countries access to markets as the motivation to adopt cleaner technologies. The diffusion of such technologies via trade might be essential for developing countries to not only expand their markets, but also reduce pollution (Taylor 2005). However, these rewards do not necessarily compensate for the costs of adopting clean technologies (Otsuki et al. 2001). Therefore, there exist cases of developed countries providing financial support for developing economies to adopt such technologies by means of technology transfer or capacity building.[5]

To make matters more complicated, there is the "standards divide" problem, in which there is a gap between an environmental standard that a provision requires and a standard

---

[4] In the case of international negotiations on MEAs, side payments are helpful in reaching an agreement (Barrett 2001).

[5] The US had several programs for supporting Mexico's compliance with environmental laws and increasing enforcement capacity along their border (EPA 1991).

already present in developing countries (Wilson and Abiola 2003).[6] During the negotiation process, these standards are likely to converge to the more stringent (and more costly) ones often employed in developed countries (Disdier 2014). Consequently, clean technologies tend to be too costly to be implemented, especially for developing countries. The World Bank suggests that governments in developing countries do not necessarily have sufficient capacity for policy implementation and hence their enforcement levels should be improved (WB 2007). Even if developed countries successfully transfer clean technologies to developing ones, it does not necessarily follow that developing countries would employ them. Using dirty technologies is usually less costly; thus, firms in a developing country may employ low-cost dirty technologies if the government's enforcement level is low. In such a case, developed countries also have to help the government monitor and enforce its policies.

In this paper, we develop a model for a free trade agreement (FTA) with environmental standards.[7] Unlike most existing papers that deal with stable MEAs among symmetric countries, we assume that there is one North country and multiple Southern countries; the Northern country can sign an FTA with any number of Southern countries.[8] We consider costly clean and cheap dirty technologies that produce a manufacturing good to be traded; the Northern country has clean technology[9] and the Southern countries have only dirty technology without free trade agreements with the North. If a Southern country

---

[6] Fischer and Serra (2000) investigate the implication of the minimum standards on a product that is produced by domestic firms and foreign competitors and show that the domestic country may choose the smallest standard to force out foreign competitors.

[7] In the literature on international agreements on climate change or international pollution, many researchers study self-enforcing environmental agreements (SIEAs) among symmetric countries - see Barrett (1994), Carraro and Siniscalco (1993), Eichner and Pethig (2013a, 2013b and 2015), and Kuhn et al. (2015). Barrett (2001) mentioned that it is difficult to reach an agreement if there is high asymmetry between countries (for instance, significant income differences).

[8] Our analysis on international agreements among asymmetric countries may be restrictive in the sense that it is limited to static games, as McGinty (2007) and Pavlova and de Zeeuw (2013). Zagonari (1998) extends the dynamic model of international pollution control developed by Long (1992) and Dockner and Long (1993) to the case of asymmetric countries and shows that the unilateral actions of one country produce less pollution than the cooperative solution when both countries use linear feedback strategies and the other country cares less about the environment. See Calvo and Rudio (2012) and Long (2010) for surveys of a dynamic game approach to international environmental control.

[9] In a dynamic game model of transboundary pollution à la Long (1992) and Dockner and Long (1993), Benchekroun and Chaudhuri (2014) show that adopting a cleaner (low-emission) technology may increase the long-run pollution stock if the initial stock is high and the natural rate of decay of the pollution is low. Because our model is static, we will not consider such a possibility.

establishes the agreement with the North, it can adopt the clean technology. However, the Southern country has an incentive to cheat and use the cheaper dirty technology. In order for a Southern country to monitor whether its firms are using the clean technology, it must spend an enforcement cost, which can differ from country to country. In order to ensure that Southern countries use the clean technology, the Northern country might need to provide monetary support to allow them to join the FTA with a certain enforcement level. That is, there are tradeoffs between the number of Southern participants and the enforcement level that the FTA imposes on its members. With this model, we ask the following questions: Is there a stable FTA? How many and which kinds of Southern countries would be invited to the FTA by the Northern country? How could Southern countries be incentivized to adopt costly clean technology? What would happen if a potential trade partner has a government that lacks the capacity to enforce policy? Is monetary compensation the way to encourage Southern countries to engage in enforcement effort? Which enforcement level should be chosen if the incentive program becomes more expensive for stricter enforcement?

The main incentive method used by Northern countries to encourage Southern countries to use the high-cost clean technology with strict enforcement is to limit the number of Southern participants.[10] If the number of Southern countries in the FTA is small, these countries receive great benefits from being included in the FTA (i.e., by having exclusive accesses to a lucrative Northern market), and thus they are willing to enforce the high-cost clean technology while demanding less transfers. Obviously, the Northern country's consumers may want more Southern competitors for lower prices, but by including more Southern countries, the enforcement level goes down and they demand more transfers. Note that this arrangement necessarily involves inequality among Southern countries. The ones in the FTA get access to the Northern market and prosper. In contrast, the ones excluded from the FTA lose business with the Northern country and become even poorer than before. Thus, FTA with environmental standards may increase inequality

---

[10] In a repeated game with two symmetric countries, Benchekroun and Yildiz (2011) analyze the sustainability of an international environmental agreement with or without free trade. They show that under free trade, the enforcement of the agreement becomes harder and determine the emission standard endogenously with the logic of repeated game.

among Southern countries.

We first show that for any given level of enforcement and monetary support, there is a stable free trade agreement in the sense that (i) no member country wants to quit the FTA unilaterally, and (ii) no outsider wants to participate in the FTA unilaterally (Proposition 1). This stability notion was first introduced by d'Aspremont et al. (1983) for analyzing cartels and is widely used by environmental economists (see Barrett 1994). Then, using linear technologies and demand, we show that if a Northern country is setting the rule by maximizing its social welfare, then the enforcement level of the clean technology usage (the fraction of production that uses the clean technology) goes down as the number of Southern participants increases (Proposition 2).

With Proposition 2, it is easy to see that there is a tradeoff between having more Southern countries in the FTA and the level of enforcement, but there are other tradeoffs as well. With more Southern memberships, a Northern country's consumer surplus increases while its domestic firm's profits and its tariff revenue decrease. We also do not know how the total level of emissions would be affected by an increase in the number of Southern countries in the FTA, since the enforcement level for the FTA members goes down while the number of Southern countries goes up. Moreover, as the Southern membership goes up, the total transfers become more and more costly for the Northern countries. Since all of these factors are important and it is hard to get qualitative results, we will present an example with reasonable parameter values and observe the optimal FTA policy for the Northern country and its environmental implications.

In the numerical example, we confirm that these considerations play important roles in evaluating the FTA policies. Limiting Southern memberships is desirable for Northern countries, but it results in sizable inequality between the FTA members and nonmembers among Southern countries. Comparative static analyses of the numerical example demonstrate that if the number of member states is kept constant, an increase in emissions from Southern countries (as their dirty technology worsens) raises the aggregate emissions. However, it also shows that once the number of member states is endogenized, its overall effect on the aggregate emissions is negative, due to the subsequent increase in the number of Southern participants that adopt clean technologies.

2 The Model

2.1 The basic structure of the model

There is one Northern country and m Southern countries in the world, and all Southern countries are identical ex ante. The set of Southern countries is denoted by $S = \{1, ..., m\}$. The Northern country (denoted by 0) has an inverse demand function for an industrial good $P(\bar{Q})$, while Southern countries have identical inverse demand functions for the industrial good $p(q_j)$, where $\bar{Q}$ and q_j are aggregated quantities in the Northern and Southern country j's markets, respectively. We assume that P and p are twice continuously differentiable. The Northern and Southern countries' wage rates (opportunity cost of labor) are exogenously fixed at w_N and w_S, respectively ($w_N > w_S > 0$).[11] There are two technologies that produce industrial goods, one clean and one dirty. In order to produce one unit of an industrial good, the clean and dirty technologies (C and D, respectively) require α_C and α_D units of labor, respectively ($\alpha_C > \alpha_D > 0$). That is, the clean technology requires more labor input to produce one unit of output than the dirty technology. Initially, the Northern country has the clean technology C, while all Southern countries have the same dirty technology D. The amount of emissions from producing one unit with the clean and dirty technologies are denoted by e_C and e_D, respectively, with $e_D > e_C \geq 0$.

The Northern country applies a common specific tariff rate $\tau > 0$ on imports from Southern countries. Unless Southern country j has a free trade agreement with the Northern country, tariff rate τ applies. We fix τ throughout this paper (τ is not a policy variable). This is because the WTO prohibits increasing tariffs when countries form an FTA.[12]

---

[11] This assumption implies that w_N units of the numeraire good can be produced from one unit of labor in the Northern country, while w_S units of the numeraire good can be produced from one unit of labor in Southern countries. Country 0 produces one unit of an industrial good by using α_C units of labor, which means that it gives up $w_N \alpha_C$ units of the numeraire good (opportunity cost) by producing one unit of the industrial good.

[12] One of the key principles of the WTO is nondiscrimination (Obviously, an FTA is itself discriminatory, but GATT's Article 24 allows for FTAs and custom unions as long as they do not provide negative

A free trade agreement does not allow a country to indirectly export goods via a third country. Each country $j \in \{0\} \cup S$ has a single firm (only Northern countries consume industrial goods). Country j's export quantity to Northern country 0 is denoted by Q_j and country 0's domestic supply is denoted by Q_0. Thus, the total supply in country 0 is $\bar{Q} = \sum_{j \in S} Q_j + Q_0$.

We will also assume that Southern countries do not import industrial goods. This assumption is imposed for simplicity of analysis and deincentivizes Southern countries from participating in the FTA.

2.2 Free trade agreement, environmental standard, and law enforcement

The WTO allows for countries to form FTAs, but requires that countries in an FTA mutually abolish tariff rates on imports from all member countries (although in our model Southern countries do not import industrial goods). Since our interest is in how international trade affects total world emissions, we assume that for a Southern country to form an FTA with Northern country 0, it must accept an environmental standard set by the North with a required enforcement level. We denote FTA partners with Northern country 0 by set $A \subseteq S$. This means that when Northern country 0 and country $j \in S$ form an FTA, country j must adopt clean technology C that requires α_C units of labor and must enforce its usage to at least some extent by spending a fixed cost to establish law enforcement. This is because the dirty technology has a lower marginal cost than the clean technology: $\alpha_D < \alpha_C$. Without an enforcement mechanism, producers are tempted to use the dirty technology, so law enforcement needs to randomly audit to check if the clean technology is being used. We will denote the level of enforcement of the clean technology implicitly by $\xi_j \in [0, 1]$: country j's firm produces only a fraction ξ_j of its output with the clean technology and the rest of its output $(1 - \xi_j)$ is produced with the dirty technology to save money. Enforcing the usage of the clean technology can be costly,

---

externalities to outsiders.). Increasing τ appears to discriminate outsiders from FTA members, even though it is motivated by a Northern country's intention to encourage Southern countries to join.

since it requires strong infrastructure, such as an audit system and well-disciplined police, which in turn requires a fixed cost. Let $F_j(\xi)$ be country j's cost of introducing the clean technology together with the cost to establish law enforcement that achieve enforcement level $\xi \in [0,1]$. We assume $F_j(\xi) = F + f_j(\xi)$ with $F \geq 0$, $f_j(0) = 0$, $f_j'(\cdot) > 0$, and $f_j''(\cdot) > 0$. We assume that F_js are ordered by the efficiency of enforcement technology: i.e., *for any* $\xi \in [0,1]$, $f_1(\xi) \leq f_2(\xi) \leq ... \leq f_m(\xi)$ *and* $f_1'(\xi) \leq f_2'(\xi) \leq ... \leq f_m'(\xi)$ *holds.*

Let the total amount of pollutive emissions in the world be described by

$$E = e_C Q_0 + \sum_{j \in S} \left(\xi_j e_C + (1-\xi_j) e_D\right)(Q_j + q_j),$$

where $\xi_j e_C + (1-\xi_j)e_D$ is country j's emission rate for $j \in A$, and $Q \equiv (Q_0, ..., Q_m)$ and $q \equiv (q_1, ..., q_m)$ denote supply vectors in the Northern and Southern countries, respectively. Northern and Southern countries receive negative externalities from pollutive emissions in an additive manner (global pollutive emissions) by $d_N E$ and $d_S E$, respectively, where $0 \leq d_S < d_N$.

3 Analysis

3.1 Northern market equilibrium allocation

We will analyze Northern country 0's market equilibrium. Firms in different countries have different effective marginal costs. The firm in country 0 has marginal cost $c_0 = w_N \alpha_C$, the one in Southern country $j \in A$ has marginal cost $c_j = w_S \alpha_C$ if $j \in A$, and the one in country $j \in S \backslash A$ has marginal cost $c_j = w_S \alpha_D + \tau$ if $j \in S \backslash A$. When there are m countries that supply the product to country i, and they have heterogeneous costs $(c_0, c_1 ..., c_m)$, the standard Cournot equilibrium solution can be obtained in the following manner: Country j's best response to q_{-j}^i is a solution of

$$\max_{Q_j} P\left(\sum_{i=0}^{m} Q_i\right) Q_j - c_j Q_j,$$

i.e., the first order condition

$$P\left(\sum_{i=0}^{m} Q_i\right) - c_j + P'\left(\sum_{i=0}^{m} Q_i\right) Q_j = 0.$$

Summing them up, we have

$$(m+1)\, P\left(\bar{Q}\right) - \sum_{i=0}^{m} c_i + P'\left(\bar{Q}\right)\bar{Q} = 0. \tag{1}$$

This equation determines $\bar{Q} = \sum_{j=0}^{m} Q_j$ and $P(\bar{Q})$ uniquely as long as the **strategic substitute condition** ($P'(\bar{Q}) + P''(\bar{Q})Q_j \leq 0$ for all $\bar{Q}$ and $Q_j < \bar{Q}$) is satisfied. The equilibrium allocation is described only by $\bar{Q}$: for all $j = 0, ..., m$

$$Q_j(\bar{Q}) = \frac{P(\bar{Q}) - c_j}{-P'(\bar{Q})}$$

and

$$\Pi_j(\bar{Q}) = \frac{(P(\bar{Q}) - c_j)^2}{-P'(\bar{Q})}, \tag{2}$$

as long as $P(\bar{Q}) \geq c_j$ is satisfied (otherwise, $Q_j = 0$ holds and firm j becomes an inactive firm: i.e., the number of firms in the market shrinks, but all nice properties still hold, even after some firms become inactive). We can show that under the strategic substitute condition, we have

$$\frac{d\bar{Q}}{d\left(\sum_{i=0}^{m} c_i\right)} = (m+2)\, P'\left(\bar{Q}\right) + P''\left(\bar{Q}\right)\bar{Q} < 0 \tag{3}$$

and $\bar{Q}$ is uniquely determined by $\sum_{i=0}^{m} c_i$ (monotonic decreasing function). This in turn determines firm j's profit, which is a decreasing function of Q:

$$\begin{aligned}\frac{\partial \Pi_j}{\partial \bar{Q}} &= \frac{2P'\left(P - c_j\right)\left(-P'\right) - \left(P - c_j\right)^2\left(-P''\right)}{\left(-P'\right)^2} \\ &= \frac{\left(P - c_j\right)}{\left(-P'\right)}\left[2P' + P''q_j\right] < 0.\end{aligned}$$

Thus, keeping c_j constant, if $\sum_{i=0}^{m} c_i$ decreases, Π_j goes down.

3.2 Southern market equilibrium allocation

In contrast, we greatly simplify each Southern country's market equilibrium. Let country j's domestic inverse demand function be $p(q_j)$.

If country j is not participating in an FTA with Northern country 0, then the firm in country j uses the dirty technology D:

$$\pi_j(q_j) = p(q_j)q_j - w_S\alpha_D q_j. \tag{4}$$

Clearly, when operating in a nonmember country, firm j chooses a monopoly output level given marginal cost $w_S\alpha_D$: $p - w_S\alpha_D + p'q_D = 0$. Let us denote the Southern countries' monopoly output and profit with the dirty technology by q_D and $\pi_D = \frac{(p(q_D)-w_S\alpha_D)^2}{-P'(q_D)}$.

There are several different possible scenarios for the marginal cost of production of a Southern FTA member country with enforcement level ξ. One reasonable assumption is that the marginal cost of production in deciding how much to produce is based on the clean technology's marginal cost $c_j = \alpha_C w_S$, even though the average cost is $(\xi_j\alpha_C + (1-\xi_j)\alpha_D)\, w_S$. This case is justified if the firm itself has the good intention to use the clean technology, while workers shirk by producing a fraction $1-\xi_j$ of its output to earn the difference in marginal costs. *Throughout the paper, we will assume that country j's firm operates using its marginal cost $c_j = \alpha_C w_S$.*[13]

Under this assumption, an FTA member country's monopoly output q_C is determined by $p - w_S\alpha_C + p'q_j = 0$. Its profit is denoted by $\pi_C = \frac{(p(q_C)-w_S\alpha_C)^2}{-p'(q_C)}$. Since $\alpha_D < \alpha_C$, $q_C < q_D$ and $\pi_C < \pi_D$ hold. The firm earns the exporting and domestic profits with the clean technology, and cheating workers get $(1-\xi_j)\,(\alpha_C - \alpha_D)\, w_S\,(Q_j + q_j)$.

---

[13] Practically, if a firm in a Southern country determines its output with a marginal cost lower than $w_S\alpha_C$, it becomes obvious that its firm is using the dirty technology. Thus, our assumption makes sense. However, it is also easy to assume that the marginal cost is the same as the average cost $c_j = (\xi_j\alpha_C + (1-\xi_j)\alpha_D)\, w_S$, which can be justified if the firm is choosing its output level based on knowledge of the usage of dirty technology (so the firm's output decision is affected by ξ). In the former case, $\frac{dc_j}{d\xi} = 0$, while in the latter case, $\frac{dc_j}{d\xi} > 0$ holds. Despite the difference in the underlying assumption, the quantitative results are the same.

3.3 Global equilibrium allocation with an FTA

Suppose that k Southern countries are in the FTA ($|A| = k$) and agree to use the clean technology, i.e., countries in $A \cup \{0\}$ adopt the technology. Since Southern countries' marginal costs depend only on the (official) technologies they use, the equilibrium output allocation vector is solely determined by A (or k). The agreed upon enforcement level ξ affects social welfare through the worldwide emission of pollutive substances E and Southern member countries' policy enforcement only.

Let $\bar{Q}(k)$ be the solution of equation (??) for $c_0 = w_N \alpha_C$, $c_j = w_S \alpha_C$ for all $j \in A$, and $c_j = w_S \alpha_D$ for all $j \notin A$. The Northern country's consumer surplus is described by $CS(k) = \int_0^{\bar{Q}(k)} \left(P(\tilde{Q}) - P(\bar{Q}(k)) \right) d\tilde{Q}$. Let $Q(k) \equiv (Q_0(k), Q_1(k), ..., Q_m(k))$ and $\Pi(k) \equiv (\Pi_0(k), \Pi_1(k), ..., \Pi_m(k))$ be such that $Q_j(k) \equiv Q_j(\bar{Q}(k))$ and $\Pi_j(k) \equiv \Pi_j(\bar{Q}(k))$ for the above $c = (c_0, c_1, ..., c_m)$. Countries' supply and profit vectors in the Northern market are dependent on their technologies: $Q_j(k) = Q_C(k)$ and $\Pi_j(k) = \Pi_C(k)$ for $j \in A$, and $Q_j(k) = Q_D(k)$ and $\Pi_j(k) = \Pi_D(k)$ for $j \notin A$. The Southern countries' domestic supply vector is simply determined as $q_j = q_C$ if $j \in A$, and $q_j = q_D$ otherwise.

The Northern country sets a clean-technology enforcement level $\xi \in [0, 1]$ and a sign-up subsidy $\sigma \geq 0$ for its FTA member (Southern) countries, and the Northern country agrees to form a free trade agreement with Southern country j as long as country j is willing to adopt the clean technology by spending enforcement cost $F_j(\xi) \geq 0$ (open membership, or non-discrimination). The worldwide emission of pollutive substance under this free trade agreement is described by

$$\begin{aligned} E(k, \xi) &= e_C Q_0(k) + \sum_{j \in A} \left(\xi e_C + (1 - \xi) e_D \right) \left(Q_j(k) + q_C \right) + \sum_{j \in S \setminus A} e_D (Q_j(k) + q_D) \\ &= e_C Q_0 + k \left(\xi e_C + (1 - \xi) e_D \right) \left(Q_C + q_C \right) + (m - k) e_D (Q_D + q_D). \end{aligned}$$

The Northern country's social welfare can be written as

$$SW(k, \xi, \sigma) = CS(k) + \Pi_0(k) - k\sigma - d_N E(k, \xi).$$

The Southern countries' consumer surplus is described by $cs_j = cs_D \equiv \int_0^{q_D} (p(q) - p(q_D))\, dq$ if $j \notin A$, and $cs_j = cs_C \equiv \int_0^{q_C} (p(q) - p(q_C))\, dq$ if $j \in A$. Their social welfare can be written as

$$sw^{OUT}(k, \xi) = sw(k, \xi) \equiv cs_D + \Pi_D(k) + \pi_D - d_S E(k, \xi) \tag{5}$$

if $j \notin A$, and

$$\begin{aligned} sw^{IN}(k, \xi) = sw(k, \xi) &\equiv cs_C + \Pi_C(k) + \pi_C + \sigma - F(\xi) \\ &\quad + (1 - \xi)(\alpha_C - \alpha_D)\, w_S (Q_C + q_C) - d_S E(k, \xi) \end{aligned} \tag{6}$$

if $j \in A$.

3.4 Participation decision in an FTA

Here, we consider an FTA between Northern country 0 and some Southern countries and analyze the set of equilibrium participants in the free trade agreements with Northern country 0. Let $A \subset S$ be the set of Southern countries that participate in free trade agreements, and let its cardinality be $a = |A|$. Note that all countries j in A have marginal costs $c_j = w_S \alpha_C$ and countries j in $S \backslash A$ have marginal costs $c_j = w_S \alpha_D + \tau$. The equilibrium set A of the Southern FTA member countries k is described by the following two inequalities:

$$sw^{IN}(k, \xi) - F - f_j(\xi) + \sigma \geq sw^{OUT}(k - 1, \xi) \text{ for all } j \in A \text{ (internal stability)}$$

and

$$sw^{IN}(k + 1, \xi) - F - f_j(\xi) + \sigma \leq sw^{OUT}(k, \xi) \text{ for all } j \notin A \text{ (external stability).}$$

If a set of Southern country members satisfies both internal and external stability conditions, then it is called a stable FTA. Extending the proof by d'Aspremont et al. (1983, Theorem), we can show that there always exists a stable FTA.

Proposition 1. For all $\xi \in [0,1]$ and all $\sigma \geq 0$, there exists a stable FTA.

Proof. First note $f_1(\xi) \leq f_2(\xi) \leq ... \leq f_m(\xi)$ for all $\xi \in [0,1]$ by assumption. Thus, if $sw^{IN}(k,\xi) - F - f_k(\xi) + \sigma \geq sw^{OUT}(k-1,\xi)$ holds, then $sw^{IN}(k,\xi) - F - f_{k'}(\xi) + \sigma \geq sw^{OUT}(k-1,\xi)$ holds for all $k' \leq k$. And if $sw^{IN}(k+1,\xi) - F - f_k(\xi) + \sigma \leq sw^{OUT}(k,\xi)$ holds, then $sw^{IN}(k+1,\xi) - F - f_{k'}(\xi) + \sigma \leq sw^{OUT}(k,\xi)$ for all $k' \geq k$.

We will prove the statement by contradiction. Suppose that there is no stable FTA. We will use an induction argument.

1. Start with $k = 0$. If $sw^{IN}(1,\xi) - F - f_1(\xi) + \sigma \leq sw^{OUT}(0,\xi)$, then $k = 0$ is a stable FTA. Since there is no stable FTA, we have $sw^{IN}(1,\xi) - F - f_1(\xi) + \sigma \leq sw^{OUT}(0,\xi)$.

2. For $k \geq 1$, suppose that $sw^{IN}(k',\xi) - F - f_{k'}(\xi) + \sigma > sw^{OUT}(k'-1,\xi)$ holds for all $k' \leq k$. This implies $sw^{IN}(k,\xi) - F - f_k(\xi) + \sigma > sw^{OUT}(k-1,\xi)$. If $sw^{IN}(k+1,\xi) - F - f_{k+1}(\xi) + \sigma \leq sw^{OUT}(k,\xi)$, then $A = \{1,...,k\}$ is a stable FTA. Thus, we have $sw^{IN}(k+1,\xi) - F - f_{k+1}(\xi) + \sigma > sw^{OUT}(k,\xi)$. By induction, $sw^{IN}(k',\xi) - F - f_{k'}(\xi) + \sigma > sw^{OUT}(k'-1,\xi)$ holds for all $k' \leq k$. This implies that $sw^{IN}(k',\xi) - F - f_{k'}(\xi) + \sigma > sw^{OUT}(k'-1,\xi)$ holds for all $k' \leq k+1$.

3. By induction, $sw^{IN}(k',\xi) - F - f_{k'}(\xi) + \sigma > sw^{OUT}(k'-1,\xi)$ holds for all $k' \leq m$. This implies that $A = S$ is internally stable. Since there are no more Southern countries, we conclude that $A = S$ is a stable FTA.

This is a contradiction.□

With general functional forms, it is hard to make general statements besides the existence of equilibrium, so we will adopt linear demand functions to describe the optimal FTA participation rule for the Northern country in the next section.

4 Optimal FTA Rules

Here, we allow the Northern country to set the FTA rule, and Southern countries can passively decide whether or not they will participate. We will assume that the Northern country can choose a policy combination of the enforcement level ξ of the clean technology usage and a sign-up subsidy σ for FTA participation. We will use linear demand functions so that we can discuss the optimal policy mix.

4.1 Linear Demand Functions

Here, we assume that the Northern country has the inverse demand function $P(Q) = 1 - Q$, and each Southern country has $p(q) = a - bq$. We have the following basic results (the proof is in Appendix A).

Lemma 1. Suppose that there are k Southern countries in the FTA. The equilibrium total output in the Northern market, the Northern country's output, the Southern FTA countries' and the non-FTA country's export to the Northern market, and the Northern country's equilibrium consumer surplus CS are

$$\bar{Q}(k) = \sum_{i=0}^{m} Q_i(k) = \frac{(m+1) - (c_0 + kc_C + (m-k)(c_D + \tau))}{m+2},$$

$$Q_0(k) = \frac{1}{m+2}\left\{1 + (kc_C + (m-k)(c_D+\tau)) - (m+1)c_0\right\},$$

$$Q_C(k) = \frac{1 + c_0 - (m-k+2)c_C + (m-k)(c_D+\tau)}{m+2},$$

$$Q_D(k) = \frac{1 + c_0 + kc_C - (k+2)(c_D+\tau)}{m+2},$$

$$CS(k) = \frac{\left[(m+1) - (c_0 + kc_C + (m-k)(c_D+\tau))\right]^2}{2(m+2)^2},$$

respectively. Profits from the Northern market earned by firms in the Northern country, the Southern FTA country (with the clean technology), and the Southern non-FTA country (with the dirty technology) are

$$\Pi_0(k) = \left(\frac{1}{m+2}\right)^2 \left[1 - (m+1)c_0 + kc_C + (m-k)(c_D+\tau)\right]^2,$$

$$\Pi_C(k) = \left(\frac{1}{m+2}\right)^2 \left[1 + c_0 - (m-k+2)c_C + (m-k)(c_D+\tau)\right]^2,$$

$$\Pi_D(k) = \left(\frac{1}{m+2}\right)^2 \left[1 + c_0 + kc_C - (k+2)\left(c_D + \tau\right)\right]^2,$$

respectively. Domestic outputs, profits, and consumer surpluses in FTA and non-FTA Southern countries are $q_C = \frac{a-c_C}{2b}$, $\pi_C = \frac{(a-c_C)^2}{4b}$, $cs_C = \frac{(a-c_C)^2}{8b}$, and $q_D = \frac{a-c_D}{2b}$, $\pi_D = \frac{(a-c_D)^2}{4b}$, $cs_D = \frac{(a-c_D)^2}{8b}$, respectively. Finally, the amount of equilibrium total emissions is

$$\begin{aligned}
E(k,\xi) = {} & (2e_D - e_C)\left(\frac{m+1}{m+2} - \frac{c_0 + kc_C + (m-k)\left(c_D + \tau\right)}{m+2}\right) \\
& - \left(e_D - e_C\right)\left(1 - c_C\right) + e_D\left\{k\frac{a-c_C}{2b} + (m-k)\frac{a-c_D}{2b}\right\} \\
& - \left(e_D - e_C\right)\left[\frac{1 + c_0 + kc_C + (m-k)\left(c_D+\tau\right) - (m+2)\,c_0}{m+2}\right] \\
& - \left(e_D - e_C\right)k\xi\left\{\frac{1 + c_0 + kc_C + (m-k)c_D - (m+2)\,c_C}{m+2} + \frac{a-c_C}{2b}\right\}.
\end{aligned}$$

With these basic results, we can analyze the optimal FTA rule for the Northern country. The Northern country can choose a policy combination, the enforcement level $\xi \in [0,1]$, and a sign-up subsidy $\sigma \geq 0$ to the participants of the FTA from Southern countries in order to maximize its social welfare.

$$SW(k,\xi,\sigma) = CS(k) + \Pi_0(k) + \tau\,(m-k)\,Q_D(k) - k\sigma - d_N E(k,\xi). \tag{7}$$

In order to find the optimal FTA policy for the Northern country, we can use the following two-step procedure. First, for each $k = 1, ..., m$, find an optimal combination of policies (ξ^k, σ^k) by solving the following problem:

$$(\xi^k, \sigma^k) \in \arg\max_{\xi,\sigma} SW(k,\xi,\sigma) \quad s.t. \quad sw^{IN}(k,\xi) - F - f_k(\xi) + \sigma \geq sw^{OUT}(k-1,\xi). \tag{8}$$

Second, choose the optimal size of an FTA k:

$$k^* = \arg\max_k SW(k, \xi^k, \sigma^k).$$

Then, $\left(\xi^{k^*}, \sigma^{k^*}\right)$ is the optimal policy that implements a size k^* FTA. *Recall that τ is an uncontrollable variable (see footnote 5).* It is easy to see that a prohibitive tariff is optimal as long as there is at least one Southern FTA member. The tariff also minimizes non-FTA countries' emissions, since it prohibits their access to the Northern market.

In the first step of the analysis, we rewrite the welfare maximization problem (??).

Lemma 2. The constraint of (??) with equality can be written as

$$\begin{aligned}
\sigma(k,\xi) = & -\frac{3\left(a-c_C\right)^2}{8b}+\frac{3\left(a-c_D\right)^2}{8b}+F+f_k(\xi) \\
& -\left(\frac{1}{m+2}\right)^2(m-1)(-c_C+(c_D+\tau)) \\
& \times\left\{2\left(1+c_0\right)-(m-2k+3)\,c_C+(m-2k-1)\left(c_D+\tau\right)\right\} \\
& +d_S\left[\left(3e_D-2e_C\right)\left(\frac{-c_C+\left(c_D+\tau\right)}{m+2}\right)-e_D\left\{-\frac{a-c_C}{2b}+\frac{a-c_D}{2b}\right\}\right. \\
& -\left(e_D-e_C\right)\xi\left\{\frac{1+c_0+kc_C+(m-k)\left(c_D+\tau\right)-(m+2)\,c_C}{m+2}+\frac{a-c_C}{2b}\right\} \\
& \left.+\left(e_D-e_C\right)(k-1)\,\xi\left\{\frac{-c_C+\left(c_D+\tau\right)}{m+2}\right\}\right].
\end{aligned}$$

This implies $\frac{\partial\sigma}{\partial k}>0$ and the constraint gets tighter as k increases. We can convert (??) into an unconstrained maximization problem by substituting this formula into (??).

Proposition 2. Under linear demand, we have $1 \geq \xi_1^* \geq \xi_2^* \geq ... \geq \xi_m^* \geq 0$ with strict inequalities $\xi_{k-1}^* > \xi_k^* > \xi_{k+1}^*$ for all ks with an interior solution $1 > \xi_k^* > 0$.

Proof. Problem (??) can be written as

$$SW(k,\xi,\sigma(k,\xi)) = CS(k)+\Pi_0(k)+\tau\,(m-k)\,Q_D(k)-k\sigma(k,\xi)-d_N E(k,\xi).$$

Thus, given k, the social optimum ξ_k^* is characterized by

$$k\frac{\partial\sigma}{\partial\xi}+d_N\frac{\partial E}{\partial\xi}=0.$$

Rewriting this, we obtain

$$\begin{aligned}
f_k'(\xi_k^*) = \left(e_D-e_C\right) & \left[\left(d_N+d_S\right)\left\{\frac{1+c_0+kc_C+(m-k)\left(c_D+\tau\right)-(m+2)\,c_C}{m+2}+\frac{a-c_C}{2b}\right\}\right. \\
& \left.-(k-1)\,d_S\left(\frac{-c_C+\left(c_D+\tau\right)}{m+2}\right)\right].
\end{aligned}$$

Since $(c_D + \tau) > c_C$, the RHS is decreasing in k. Since $f_k''(\xi) > 0$ and $f_k'(\xi) \geq f_{k-1}'(\xi)$ for all ξ, we conclude that $\xi_k^* < \xi_{k-1}^*$ holds for all k as long as they are interior solutions.□

This proposition shows that there is a tradeoff between the number of Southern participants and the level of enforcement. Although it is hard to analyze whether or not equilibrium σ increases monotonically without specifying f_k functions, it is quite natural to assume that the total subsidy payment $k\sigma_k^* < (k+1)\,\sigma_{k+1}^*$ holds for all k as long as solutions are interior. Thus, the Northern country cannot expand the membership of the FTA too much, since such an expansion means that the program becomes more costly and the level of enforcement goes down.

5 A Numerical Example

In this section, we provide a numerical example to illustrate the quantitative properties of our model. We specify the f_k function in the following manner:

$$f_k(\xi) = \frac{1}{2}\beta_k \xi^2,$$

where $\beta_1 \leq \beta_2 \leq ... \leq \beta_m$. This formulation satisfies $f_k'(0) = 0$ while $f_k(1) = \beta_k < \infty$. Then, ξ_k^* is written as

$$\xi_k^* = \frac{(e_D - e_C)}{\beta_k}\left[(d_N + d_S)\left\{\frac{1 + c_0 + kc_C + (m-k)\,(c_D+\tau) - (m+2)\,c_C}{m+2} + \frac{a - c_C}{2b}\right\}\right.$$
$$\left. - (k-1)\,d_S\left(\frac{-c_C + (c_D+\tau)}{m+2}\right)\right]$$

if the RHS is less than 1, and $\xi_k^* = 1$ otherwise.

We set the parameter values as $m = 10$, $c_0 = 0.25$, $c_C = 0.2$, $c_D = 0.15$, $\tau = 0.1$, $a = 0.3$, $b = 1$, $d_N = 0.5$, $d_S = 0$, $e_D = 0.3$, and $e_C = 0.1$. We also assume that $\beta_k = \beta = 0.017$ for all k and $F = 0$.

This numerical example is not the most realistic one, but it provides a good understanding of the model. Our main findings are as follows.

Table 1: A Numerical Example

k	0	1	2	3	4	5	6	7	8	9	10
$\bar{Q}$	0.6875	0.69167	0.69583	0.7	0.70417	0.70833	0.7125	0.71667	0.72083	0.725	0.79217
P	0.3125	0.30833	0.30417	0.3	0.29583	0.29167	0.2875	0.28333	0.27917	0.275	0.27083
Q_0	0.0625	0.05833	0.05417	0.05	0.04583	0.04167	0.0375	0.03333	0.02917	0.025	0.02083
Q_C	-	0.10833	0.10417	0.1	0.09583	0.09167	0.0875	0.08333	0.07917	0.075	0.07083
Q_D	0.0625	0.05833	0.05417	0.05	0.04583	0.04167	0.0375	0.03333	0.02917	0.025	0.02083
Π_0	0.00391	0.0034	0.00293	0.0025	0.0021	0.00174	0.00141	0.00111	0.00085	0.00063	0.00043
Π_C	-	0.01174	0.01085	0.01	0.00918	0.0084	0.00766	0.00694	0.00627	0.00563	0.00502
Π_D	0.00391	0.0034	0.00293	0.0025	0.0021	0.00174	0.00141	0.00111	0.00085	0.00063	0.00043
CS	0.23633	0.2392	0.24209	0.245	0.24793	0.25087	0.25383	0.25681	0.2598	0.26281	0.26584
ξ	-	0.93137	0.90686	0.88235	0.85784	0.83333	0.80882	0.78431	0.7598	0.73529	0.71078
E	0.24438	0.22218	0.20263	0.18562	0.17102	0.15872	0.14857	0.14047	0.13429	0.12991	0.12721
σ	-	0.00024	0.00055	0.00078	0.00107	0.00137	0.00168	0.002	0.00233	0.00267	0.00302
TR	0.0625	0.0525	0.04333	0.035	0.0275	0.02083	0.015	0.01	0.00583	0.0025	0
SW	0.18057	0.18378	0.18603	0.18734	*0.18772*	0.18722	0.18586	0.18368	0.18069	0.17695	0.17246

(1) Starting from no free trade agreement, if one Southern country joins the FTA, it gets a high market share of the Northern market. Thus, if only one country joins the agreement, a high enforcement rate can be imposed with only a small sign-up subsidy. (Depending on the parameter values, $\xi = 1$ and $\sigma = 0$ can occur very easily).

(2) In this set of parameter values, tariff revenue plays a strong role in the Northern country's social welfare; as a result, it cares less about FTA.

(3) With this set of parameter values, E_k^* is monotonically decreasing in k but the magnitude of marginal reduction in k is decreasing. Note that the level of enforcement ξ_k^* is monotonically decreasing. Thus, depending on parameter values, the movement of total emissions E_k^* can be non monotonic in k. As ξ_k^* decreases, E_k^* can turn back upward. This is because the Northern market is much larger than the Southern market.

(4) The Northern country needs to evaluate the benefits and costs of changing its policies (ξ and σ) to increase Southern countries' membership by evaluating CS, Π_0, and TR (tariff revenues), in addition to emissions E. Here, $k = 4$ is the optimal number of Southern countries in the FTA.

(5) Under some parameter values, nonmember Southern countries can be effectively excluded from the Northern market (if $P(k) < c_C + \tau$).

Moreover, we can easily see how changes in the enforcement cost β, the tariff rate τ, the cost of the clean technology c_C, and the emission from the dirty technology e_D, affect the optimal number of Southern countries participating in the FTA. In Appendix B, we show the results of the changes in these values ($\beta_k = \beta$ from 0.017 to 0.02, τ from 0.1 to 0.15, c_C from 0.2 to 0.18, and e_D from 0.3 to 0.5), from which we can observe the following.

(1) If the enforcement efficiency is lower (higher β), the enforcement of clean-tech implementation is more difficult and FTA membership declines. Therefore, it would be better to exclude a state with a high probability of cheating.

(2) The higher tariff rate (τ) increases the number of member states. Whereas the Northern country tries to decrease the number, the Southern countries have more incentive to become a member to avoid the considerably high tariff rate.

(3) If the clean technology is less costly (lower c_C), more states will join the FTA. Additionally, emissions decline because such reduction will be easier.

(4) An increase in the emission rate (higher e_D) in Southern countries raises the aggregate emissions as long as the number of member states is kept constant. However, these higher emissions induce the Northern country to persuade Southern countries to become members. Thus, the number of member countries adopting the clean technology increases and eventually the aggregate level of emissions declines.

6 Conclusion

In this paper, we analyzed the optimal free trade agreement (FTA) between Northern and Southern countries by explicitly considering the environmental aspects of trade. We first proved the existence of a stable free trade agreement. We then showed that there exists an interior solution to the optimal number of Southern member countries. Although the firms in Southern member countries take advantage of unimpeded access to the Northern market, they are unwilling to employ clean but costly technology provided by Northern country. Then, the Northern country has to propose a sufficiently beneficial FTA to Southern countries in order to enforce the implementation of tighter environmental regulation. Since an excessive number of Southern participants discourages others from joining the FTA, it is essential for the North to restrict Southern memberships when strict enforcement of emission reduction is required. We have also provided quantitative evaluation of FTA policies using a numerical example. We demonstrated that, on the one hand, an increase in the emission rate in Southern countries (which may be due to economic growth) raises the aggregate emissions if the number of member states is kept constant. On the other hand, its overall effect on the aggregate emissions is negative due to the corresponding increase in Southern member countries, all of which adopt the clean technology.

Apart from the modeling, there might be another political reason to tie a high environmental standard to free trade agreements. Imposing a high environmental standard (enforcement of the clean technology) makes it politically easier for a Northern country to form an FTA with Southern countries (Bill Clinton forced Mexico to satisfy higher environmental standards, for instance.). The number of Southern countries will be reduced as a byproduct, which also helps to pass the bill in Congress/Parliament. In such a case, it might also be interesting to analyze whether political turnover would affect the number of Southern participants or global emissions. These factors may require further investigation.

Appendix A: Linear Demand

Here, we assume that the Northern country has the following demand function: $P(Q) = 1 - Q$. Firm j's profit maximization problem is

$$\max_{q_j^0} \left(1 - \sum_{i=0}^{m} Q_i\right) Q_j - c_j Q_j.$$

The first order condition is

$$1 - \sum_{i=0}^{m} Q_i - Q_j - c_j = 0.$$

Summing them up, we obtain

$$(m+1) - \left((m+2)\sum_{i=0}^{m} Q_i\right) - \sum_{i=0}^{m} c_i = 0$$

and

$$\bar{Q} = \sum_{i=0}^{m} Q_i = \frac{m+1}{m+2} - \frac{1}{m+2}\sum_{i=0}^{m} c_i.$$

Let $\alpha_C w_N = c_0$, $\alpha_C w_S = c_C$, and $\alpha_D w_S = c_D$. We assume that in the presence of a tariff charged by the Northern country, the marginal cost of using the clean technology in the FTA is lower than the one using the dirty technology outside of the FTA if they export $c^{OUT} = c_D + \tau > c^{IN} = c_C$ naturally (although $c_C > c_D$ holds). The equilibrium output by country j when k Southern countries participate in the FTA is

$$\begin{aligned} Q_j &= \frac{1}{m+2} + \frac{1}{m+2}\sum_{i=0}^{m} c_i - c_j \\ &= \frac{1}{m+2}\left\{1 + \left(c_0 + k c_C + (m-k)(c_D + \tau)\right) - (m+2)\, c_j\right\}. \end{aligned}$$

Thus, the Northern country's output and FTA and non-FTA Southern countries' exports are written as

$$Q_0(k) = \frac{1}{m+2}\left\{1 + \left(k c_C + (m-k)(c_D + \tau)\right) - (m+1)\, c_0\right\},$$

$$Q_C(k) = \frac{1}{m+2}\left[1 + c_0 - (m-k+2)\, c_C + (m-k)(c_D + \tau)\right],$$

$$Q_D(k) = \frac{1}{m+2}\left[1 + c_0 + kc_C - (k+2)(c_D+\tau)\right],$$

respectively. Since $\Pi_j = Q_j^2$, we have the following

$$\Pi_C(k) = \left(\frac{1}{m+2}\right)^2 \left[1 + c_0 - (m-k+2)\,c_C + (m-k)(c_D+\tau)\right]^2,$$

$$\Pi_D(k) = \left(\frac{1}{m+2}\right)^2 \left[1 + c_0 + kc_C - (k+2)(c_D+\tau)\right]^2.$$

Substituting Q_js and q_js into $E(k,\xi)$, we obtain

$$\begin{aligned}
E(k,\xi) &= e_D\bar{Q}(k) - (e_D - e_C)\,Q_0(k) + e_D\left\{kq_C + (m-k)\,q_D\right\} - (e_D - e_C)\,k\xi\left\{Q_C(k) + q_C\right\}\\
&= (2e_D - e_C)\left(\frac{m+1}{m+2} - \frac{c_0 + kc_C + (m-k)(c_D+\tau)}{m+2}\right)\\
&- (e_D - e_C)(1 - c_C) + e_D\left\{k\frac{a - c_C}{2b} + (m-k)\frac{a - c_D}{2b}\right\}\\
&- (e_D - e_C)\left[\frac{1 + c_0 + kc_C + (m-k)(c_D+\tau) - (m+2)\,c_0}{m+2}\right]\\
&- (e_D - e_C)\,k\xi\left\{\frac{1 + c_0 + kc_C + (m-k)(c_D+\tau) - (m+2)\,c_C}{m+2} + \frac{a - c_C}{2b}\right\}.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&E(k,\xi) - E(k-1,\xi)\\
&= (3e_D - 2e_C)\left(\frac{-c_C + (c_D+\tau)}{m+2}\right) + e_D\left\{\frac{a - c_C}{2b} - \frac{a - c_D}{2b}\right\}\\
&- (e_D - e_C)\,k\xi\left\{\frac{1 + c_0 + kc_C + (m-k)(c_D+\tau) - (m+2)\,c_C}{m+2} + \frac{a - c_C}{2b}\right\}\\
&+ (e_D - e_C)(k-1)\,\xi\left\{\frac{1 + c_0 + (k-1)\,c_C + (m-k+1)(c_D+\tau) - (m+2)\,c_C}{m+2} + \frac{a - c_C}{2b}\right\}\\
&= (3e_D - 2e_C)\left(\frac{-c_C + (c_D+\tau)}{m+2}\right) - e_D\left\{-\frac{a - c_C}{2b} + \frac{a - c_D}{2b}\right\}\\
&- (e_D - e_C)\,\xi\left\{\frac{1 + c_0 + kc_C + (m-k)(c_D+\tau) - (m+2)\,c_C}{m+2} + \frac{a - c_C}{2b}\right\}\\
&+ (e_D - e_C)(k-1)\,\xi\left\{\frac{-c_C + (c_D+\tau)}{m+2}\right\}.
\end{aligned}$$

We can interpret the above formula as follows. The first term is an indirect effect of equilibrium output that increases in the Northern market by giving another Southern country access to the Northern market. The second term is an output reduction effect in a new Southern entrant country. The third term is the direct effect of reducing emissions by having another country with clean technology in the Northern market. The fourth term represents an indirect effect of reduction in clean technology production in the existing $k-1$ Southern member countries crowded out by the kth Southern country's participation.

Southern country j's social welfare is provided for two different cases: being a member or a nonmember of the FTA. Southern countries' social welfare can be written as

$$\begin{aligned} sw^{OUT}(k,\xi) &= cs_D + \Pi_D(k) + \pi_D - d_S E(k,\xi) \\ &= \frac{(a-c_D)^2}{8b} + \left(\frac{1}{m+2}\right)^2 \left[1 + c_0 + kc_C - (k+2)(c_D+\tau)\right]^2 + \frac{(a-c_D)^2}{4b} \\ &- d_S E(k,\xi) \end{aligned}$$

if $j \notin A$, and

$$\begin{aligned} sw^{IN}(k,\xi) &= cs_C + \Pi_C(k) + \pi_C + (1-\xi)(\alpha_C - \alpha_D)\, w_S (Q_C + q_C) - d_S E(k,\xi) \\ &= \frac{(a-c_C)^2}{8b} + \left(\frac{1}{m+2}\right)^2 \left[1 + c_0 - (m-k+2)\, c_C + (m-k)(c_D+\tau)\right]^2 \\ &+ \frac{(a-c_C)^2}{4b} - d_S E(k,\xi) \end{aligned}$$

if $j \in A$.

This implies

$$\begin{aligned}
& sw^{IN}(k,\xi) - sw^{OUT}(k-1,\xi) \\
& = \frac{(a-c_C)^2}{8b} + \left(\frac{1}{m+2}\right)^2 \left[1 + c_0 - (m-k+2)\,c_C + (m-k)\,(c_D+\tau)\right]^2 + \frac{(a-c_C)^2}{4b} \\
& - d_S E(k,\xi) - \frac{(a-c_D)^2}{8b} - \left(\frac{1}{m+2}\right)^2 \left[1 + c_0 + (k-1)\,c_C - (k+1)\,(c_D+\tau)\right]^2 \\
& - \frac{(a-c_D)^2}{4b} + d_S E(k-1,\xi) \\
& = \frac{3\,(a-c_C)^2}{8b} - \frac{3\,(a-c_D)^2}{8b} \\
& + \left(\frac{1}{m+2}\right)^2 (m-1)(c_D+\tau-c_C) \times \left\{2\,(1+c_0) - (m-2k+3)\,c_C + (m-2k-1)\,(c_D+\tau)\right\} \\
& + d_S(E(k-1,\xi) - E(k,\xi)).
\end{aligned}$$

Northern country 0 can choose a policy combination: the enforcement level $\xi \in [0,1]$ and a sign-up subsidy $\sigma \geq 0$ to the FTA participants from Southern countries. In order to find the optimal FTA policy for the Northern country, we can use the following procedure. First, for each $k = 1, ..., m$, find an optimal combination of policies (ξ^k, σ^k) by solving the following problem:

$$(\xi^k, \sigma^k) \in \arg\max_{\xi,\sigma} SW(k,\xi,\sigma) \quad s.t. \quad sw^{IN}(k,\xi) - F - f_k(\xi) + \sigma \geq sw^{OUT}(k-1,\xi).$$

When describing the binding constraint of the above problem, we express the subsidy amount as a function of ξ and k:

$$\begin{aligned}
\sigma &= s(\xi,k) \\
&= -sw^{IN}(k,\xi)+F(0)+f(\xi)+sw^{OUT}(k-1,\xi) \\
&= \frac{3\,(2a-c_C-c_D)\,(c_C-c_D)}{8b} \\
&- (1-\xi)\,(c_C-c_D)\left[\frac{1+c_0-(m-k+2)\,c_C+(m-k)\,(c_D+\tau)}{m+2}+\frac{a-c_C}{2b}\right] \\
&- d_S\left(\frac{1}{m+2}\right)^2(m-1)(c_D+\tau-c_C)\times\{2\,(1+c_0)+(m-2k+1)\,(c_D+\tau-c_C)\} \\
&+ F+f_k(\xi)+d_S E(k,\xi).
\end{aligned}$$

Problem (??) can be written as

$$SW(k,\xi,\sigma(k,\xi)) = CS(k)+\Pi_0(k)+\tau\,(m-k)\,Q_D(k)-k\sigma(k,\xi)-d_N E(k,\xi).$$

Thus, given k, the social optimum ξ_k^* is characterized by

$$k\frac{\partial\sigma}{\partial\xi}+d_N\frac{\partial E}{\partial\xi}=0.$$

Thus, we have

$$\begin{aligned}
&k f_k'(\xi_k^*)-d_S\,(e_D-e_C)\,k\left\{\frac{1+c_0+kc_C+(m-k)\,(c_D+\tau)-(m+2)\,c_C}{m+2}+\frac{a-c_C}{2b}\right\} \\
&+ d_S\,(e_D-e_C)\,k\,(k-1)\left\{\frac{-c_C+(c_D+\tau)}{m+2}\right\} \\
&- d_N\,(e_D-e_C)\,k\left\{\frac{1+c_0+kc_C+(m-k)\,(c_D+\tau)-(m+2)\,c_C}{m+2}+\frac{a-c_C}{2b}\right\} \\
&= 0.
\end{aligned}$$

Rewriting this, we obtain

$$\begin{aligned}
f_k'(\xi_k^*) = (e_D-e_C)\Bigg[&(d_N+d_S)\left\{\frac{1+c_0+kc_C+(m-k)\,(c_D+\tau)-(m+2)\,c_C}{m+2}+\frac{a-c_C}{2b}\right\} \\
&-(k-1)\,d_S\left(\frac{-c_C+(c_D+\tau)}{m+2}\right)\Bigg].
\end{aligned}$$

Since $(c_D + \tau) > c_C$, the RHS is decreasing in k. Since $f_k''(\xi) > 0$ and $f_k'(\xi) \geq f_{k-1}'(\xi)$ for all ξ, we conclude that $\xi_k^* < \xi_{k-1}^*$ holds for all k as long as they are interior solutions.

Appendix B: More numerical examples

Table A1: Lower Efficiency of Enforcement: $\beta = \beta_k = 0.02$

k	0	1	2	3	4	5	6	7	8	9	10
$\bar{Q}$	0.6875	0.69167	0.69583	0.7	0.70417	0.70833	0.7125	0.71667	0.72083	0.725	0.72917
P	0.3125	0.30833	0.30417	0.3	0.29583	0.29167	0.2875	0.28333	0.27917	0.275	0.27083
Q_0	0.0625	0.05833	0.05417	0.05	0.04583	0.04167	0.0375	0.03333	0.02917	0.025	0.02083
Q_C	-	0.10833	0.10417	0.1	0.09583	0.09167	0.0875	0.08333	0.07917	0.075	0.07083
Q_D	0.0625	0.05833	0.05417	0.05	0.04583	0.04167	0.0375	0.03333	0.02917	0.025	0.02083
Π_0	0.00391	0.0034	0.00293	0.0025	0.0021	0.00174	0.00141	0.00111	0.00085	0.00063	0.00043
Π_C	-	0.01174	0.01085	0.01	0.00918	0.0084	0.00766	0.00694	0.00627	0.00563	0.00502
Π_D	0.00391	0.0034	0.00293	0.0025	0.0021	0.00174	0.00141	0.00111	0.00085	0.00063	0.00043
CS	0.23633	0.2392	0.24209	0.245	0.24793	0.25087	0.25383	0.25681	0.2598	0.26281	0.26584
ξ	1	0.79167	0.77083	0.75	0.72917	0.70833	0.6875	0.66667	0.64583	0.625	0.60417
E	0.24438	0.22544	0.20876	0.19423	0.18176	0.17123	0.16254	0.15559	0.15027	0.14648	0.14411
σ	-	-	-	0.00012	0.00046	0.0008	0.00115	0.00151	0.00188	0.00226	0.00264
TR	0.0625	0.0525	0.04333	0.035	0.0275	0.02083	0.015	0.01	0.00583	0.0025	-
SW	0.18055	0.18238	0.18398	*0.18502*	0.18482	0.18383	0.18206	0.17955	0.17632	0.17239	0.16779

Table A2: Higher Tariff Rate: $\tau = 0.15$

k	0	1	2	3	4	5	6	7	8	9	10
$\bar{Q}$	0.64583	0.65417	0.6625	0.67083	0.67917	0.6875	0.69583	0.70417	0.7125	0.72083	0.72917
P	0.35417	0.34583	0.3375	0.32917	0.32083	0.3125	0.30417	0.29583	0.2875	0.27917	0.27083
Q_0	0.10417	0.09583	0.0875	0.07917	0.07083	0.0625	0.05417	0.04583	0.0375	0.02917	0.02083
Q_C	-	0.14583	0.1375	0.12917	0.12083	0.1125	0.10417	0.09583	0.0875	0.07917	0.07083
Q_D	0.05417	0.04583	0.0375	0.02917	0.02083	0.0125	0.00417	-	-	-	-
Π_0	0.01085	0.00918	0.00766	0.00627	0.00502	0.00391	0.00293	0.0021	0.00141	0.00085	0.00043
Π_C	-	0.02127	0.01891	0.01668	0.0146	0.01266	0.01085	0.00918	0.00766	0.00627	0.00502
Π_D	0.00293	0.0021	0.00141	0.00085	0.00043	0.00016	0.00002	-	-	-	-
CS	0.20855	0.21397	0.21945	0.22501	0.23063	0.23633	0.24209	0.24793	0.25383	0.2598	0.26584
ξ	1	1	1	1	1	0.95588	0.90686	0.85784	0.80882	0.7598	0.71078
E	0.22354	0.19432	0.16844	0.14589	0.12667	0.11612	0.11081	0.10976	0.11247	0.11845	0.12721
σ	-	-	-	-	-	-	-	-	-	0.00122	0.0024
TR	0.08125	0.06188	0.045	0.03063	0.01875	0.00938	0.0025	-	-	-	-
SW	0.18888	0.18786	0.18789	0.18896	0.19107	0.19155	0.19212	0.19514	*0.19900*	0.19045	0.17871

Table A3: Cheaper Clean Technology: $c_C = 0.18$

k	0	1	2	3	4	5	6	7	8	9	10
$\bar{Q}$	0.6875	0.69333	0.69917	0.705	0.71083	0.71667	0.7225	0.72833	0.73417	0.74	0.74583
P	0.3125	0.30667	0.30083	0.295	0.28917	0.28333	0.2775	0.27167	0.26583	0.26	0.25417
Q_0	0.0625	0.05667	0.05083	0.045	0.03917	0.03333	0.0275	0.02167	0.01583	0.01	0.00417
Q_C	-	0.12667	0.12083	0.115	0.10917	0.10333	0.0975	0.09167	0.08583	0.08	0.07417
Q_D	0.0625	0.05667	0.05083	0.045	0.03917	0.03333	0.0275	0.02167	0.01583	0.01	0.00417
Π_0	0.00391	0.00321	0.00258	0.00203	0.00153	0.00111	0.00076	0.00047	0.00025	0.0001	0.00002
Π_C	-	0.01604	0.01460	0.01323	0.01192	0.01068	0.00951	0.00840	0.00737	0.0064	0.0055
Π_D	0.00391	0.00321	0.00258	0.00203	0.00153	0.00111	0.00076	0.00047	0.00025	0.0001	0.00002
CS	0.23633	0.24036	0.24442	0.24851	0.25264	0.25681	0.261	0.26523	0.26950	0.2738	0.27813
ξ	1	1	1	1	0.9951	0.96078	0.92647	0.89216	0.85784	0.82353	0.78922
E	0.24438	0.21774	0.19344	0.17147	0.15229	0.13891	0.12891	0.12204	0.11807	0.11675	0.11784
σ	-	-	-	-	-	-	-	0.00049	0.00117	0.00186	0.00257
TR	0.0625	0.051	0.04067	0.0315	0.0235	0.01667	0.011	0.0065	0.00317	0.001	0
SW	0.18055	0.1857	0.19095	0.1963	0.20153	0.20513	*0.2083*	0.20772	0.2045	0.19974	0.19355

Table A4: Higher Emission Rate: $e_D = 0.5$

k	0	1	2	3	4	5	6	7	8	9	10
$\bar{Q}$	0.6875	0.69333	0.69917	0.705	0.71083	0.71667	0.7225	0.72833	0.73417	0.74	0.74583
P	0.31250	0.30667	0.30083	0.295	0.28917	0.28333	0.2775	0.27167	0.26583	0.26	0.25417
Q_0	0.0625	0.05667	0.05083	0.045	0.03917	0.03333	0.0275	0.02167	0.01583	0.01	0.00417
Q_C	-	0.12667	0.12083	0.115	0.10917	0.10333	0.0975	0.09167	0.08583	0.08	0.07417
Q_D	0.0625	0.05667	0.05083	0.045	0.03917	0.03333	0.0275	0.02167	0.01583	0.01	0.00417
Π_0	0.00391	0.00321	0.00258	0.00203	0.00153	0.00111	0.00076	0.00047	0.00025	0.0001	0.00002
Π_C	-	0.01604	0.0146	0.01323	0.01192	0.01068	0.00951	0.0084	0.00737	0.0064	0.0055
Π_D	0.00391	0.00321	0.00258	0.00203	0.00153	0.00111	0.00076	0.00047	0.00025	0.0001	0.00002
CS	0.23633	0.24036	0.24442	0.24851	0.25264	0.25681	0.261	0.26523	0.26950	0.27380	0.27813
ξ	1	1	1	1	1	1	1	1	1	1	1
E	0.40313	0.35011	0.30177	0.25809	0.21908	0.18474	0.15506	0.13005	0.10971	0.09403	0.08302
σ	-	-	-	-	-	-	0.00004	0.00079	0.00154	0.00229	0.00304
TR	0.0625	0.051	0.04067	0.0315	0.0235	0.01667	0.011	0.0065	0.00317	0.001	-
SW	0.10117	0.11951	0.13678	0.15299	0.16813	0.18221	0.19497	0.20164	0.20575	*0.20729*	0.20627

References

[1] Baghdadi, Leila, Inmaculada Martinez-Zarzoso, and Habib Zitouna (2013). Are RTA agreements with environmental provisions reducing emissions? Journal of International Economics 90(2), 378-390.

[2] Barrett, Scott (1994). Self-enforcing international environmental agreements. Oxford Economic Papers 46(Supplement 1), 878-894.

[3] Barrett, Scott (2001). International cooperation for sale. European Economic Review 45(10), 1835-1850.

[4] Behar, Alberto, and Laia Cirera-i-Crivillé (2013). Does it matter who you sign with? Comparing the impacts of North-South and South-South trade agreements on bilateral trade. Review of International Economics 21(4), 765-782.

[5] Benchekroun, Hassan, and Amrita Ray Chaudhuri (2014). Transboundary pollution and clean technologies. Resource and Energy Economics, 36(2), 601-619.

[6] Benchekroun, Hassan, and Halis Murat Yildiz (2011). Free trade, autarky and the sustainability of an international environmental agreement. The BE Journal of Economic Analysis & Policy 11(1).

[7] Bodansky, Daniel (2016). The Paris climate change agreement: a new hope? American Journal of International Law 110(2), 288-319.

[8] Calvo, Emilio, and Santiago J. Rubio (2012). Dynamic models of international environmental agreements: a differential game approach. International Review of Environmental and Resource Economics 6(4), 289-339.

[9] Carraro, Carlo, and Domenico Siniscalco (1993). Strategies for the international protection of the environment. Journal of Public Economics 52(3), 309-328.

[10] CEC (1993). North American Agreement on Environmental Cooperation, Secretariat of the Commission for Environmental Cooperation (`http://www.cec.org/about-us/NAAEC`).

[11] Disdier, Anne-Célia, Lionel Fontagné, and Olivier Cadot (2014). North-South standards harmonization and international trade. The World Bank Economic Review 29(2), 327-352.

[12] Dockner, Engelbert J., and Ngo Van Long (1993). International pollution control: Cooperative versus noncooperative strategies. Journal of Environmental Economics and Management 25(1), 13-29.

[13] Eichner, Thomas and Rüdiger Pethig (2013a). Self-enforcing environmental agreements and international trade. Journal of Public Economics 102, 37-50.

[14] Eichner, Thomas and Rüdiger Pethig (2013b). Trade tariffs and self-enforcing environmental agreements. CESifo Working Paper No. 4464.

[15] Eichner, Thomas and Rüdiger Pethig (2015). Is trade liberalization conducive to the formation of climate coalitions? International Tax and Public Finance 22, 932-955.

[16] EPA (1994). Environmental Protection Along the U.S.-Mexico Border.

[17] Fischer, Ronald, and Pablo Serra (2000). Standards and protection. Journal of International Economics 52(2), 377-400.

[18] Gutiérrez, Emilio, and Kensuke Teshima (2018). Abatement expenditures, technology choice, and environmental performance: Evidence from firm responses to import competition in Mexico. Journal of Development Economics 133, 264-274.

[19] Kuhn, Thomas, Radomir Pestow, and Anja Zenker (2015). Self-enforcing climate coalitions and preferential free trade areas. WWDP 123/2015.

[20] Long, Ngo Van (1992). Pollution control: A differential game approach. Annals of Operations Research 37(1), 283-296.

[21] Long, Ngo, Van (2010). A Survey of Dynamic Games in Economics. Vol. 1. World Scientific.

[22] McGinty, Matthew (2007). International environmental agreements among asymmetric nations. Oxford Economic Papers 59(1), 45-62.

[23] Otsuki, Tsunehiro, John S. Wilson, and Mirvat Sewadeh (2001). Saving two in a billion: quantifying the trade effect of European food safety standards on African exports. Food Policy 26(5), 495-514.

[24] Pavlova, Yulia, and Aart De Zeeuw (2013) Asymmetries in international environmental agreements. Environment and Development Economics 18(1), 51-68.

[25] Taylor, M. Scott (2005). Unbundling the pollution haven hypothesis. The BE Journal of Economic Analysis & Policy 3(2), 1-28.

[26] UNFCCC (2019). Party Groupings (`https://unfccc.int/process-and-meetings/parties-non-party-stakeholders/parties/party-groupings`).

[27] WB (2007). Journey to a Cleaner Future, The World Bank (`http://documents.worldbank.org/curated/en/863081468023342359/pdf/422020WP0Box031CleanerFutureRoadmap.pdf`).

[28] Wilson, John S., and Victor O. Abiola (2003). Standards and Global Trade: A Voice for Africa. The World Bank.

[29] WTO (2019). Regional Trade Agreements (`https://www.wto.org/english/tratop_e/region_e/region_e.htm`).

[30] Zagonari, Fabio (1998). International pollution problems: Unilateral initiatives by environmental groups in one country. Journal of Environmental Economics and Management 36(1), 46-69.