INTERNATIONAL AGREEMENTS AND THE ENVIRONMENT

Raed Safadi

Working Paper 9501

Please address correspondence to: Raed Safadi, OECD, 2 Rue Andre Pascal, 75775 PARIS CEDEX 16, France. Fax: 33 1 45 24 15 39.

International Agreements and the Environment

Raed Safadi

OECD*

* The views expressed here reflect the authors' own, and not those of the OECD Secretariat or any of its member governments.

Abstract

This paper focuses upon the links between trade and the environment, and the use of trade policies to address environmental spillovers among countries. The paper empahsizes the existence of alternative forms of international cooperation, ranging from binding agreements to loose coordinating arrangements. A game theoretic model is developed to derive the conditions under which implicit cooperation may be as efficient in terms of environmental outcomes as a binding agreement. The reason for being interested in implicit cooperation is that the establishmment of binding international agreements may prove costly and elusive. The search for international commitments can distract attention from the possibility of taking domestically-based action, and leave the environment worse off.

ملخص

تركز هذه الورقة على العلاقة بين التجارة والبيئة، وعلى استخدام السياسات التجارية في مواجهة التأثيرات البيئية عبر الدول. وتؤكد الورقة على وجود خيارات من صور التعاون الدولى تقع بين اتفاقيات الالتزام الكامل واتفاقيات التعاون والتنسيق الغير مقيدة. و تم تطوير نموذج العاب نظرى لاشتقاق الشروط التى يصبح بموجبها اتفاق التعاون والتنسيق غير المقيد بكفاءة مماثلة للاتفاق الملزم فى اثره على البيئة. والسبب بالاهتمام بالتعاون الضمنى هو ان التوصل الى اتفاق دولى ملزم قد يكون مكلفا وصعبا. كما ان البحث عن التزام دولى قد يحول الانتباه عن امكانية التحرك المحلى، مما يترك البيئة اكثر سوء.

Contents

I- Introduction

Growing demands for environmental quality raise thorny problems for policy makers. Some of the most difficult choices to be made in addressing environmental degradation concern the interaction among sovereign states. Where environmental problems involve transfrontier spillovers (such as acid rain and river pollution) or the global commons (depletion of the ozone layer and the greenhouse effect), options open to governments range from unilateral action to cooperation through internationally binding agreements.

Unilaterally-defined international environment policy, especially on the part of large countries, is likely to entail the application (or at least the threat) of action against other countries perceived to be responsible for a given problem of environmental degradation. Whether such measures are characterized as retaliatory, punitive, or offsetting, the fact that they are unilaterally-defined from an exclusively national welfare perspective will ensure their suboptimality. On the other hand, it is far from obvious that the solution at the other end of the spectrum, that of binding international agreements, can be relied upon to ensure the best outcome. International agreements take time to negotiate and they involve a variety of transactions costs in their formulation and implementation. They also divert attention from domestic initiatives to improve the environment, and tend to promote "lowest common denominator" outcomes in terms of the substance of binding commitments. Another factor to consider is that because environmental policy is in its formative stages in many countries, it relies to a significant degree on relatively inefficient and costly direct regulation.[1] There may be additional costs from tying countries into this kind of approach in international agreements.[2] Moreover, some of the estimates that have been made of the magnitude of the policy changes (and implied international commitments) necessary to have a significant impact upon polluting emissions globally are such as to raise serious doubts about the feasibility of internationally-binding commitments in this sphere.[3]

---

[1] See Eskeland and Jimenez (1991) for a survey of available policy instruments for pollution control.

[2] For an analysis in these terms on the Montreal Protocol on Substances that Deplete the Ozone Layer, see Bohm (1990).

[3] See Whalley (1991) and Winter (1991) on estimates of carbon taxes necessary to address the greenhouse effect.

A limited amount of work has been undertaken on the costs and benefits of binding international agreements aimed at improving environmental quality.[4] This suggests that they do not have a positive payoff. Yet there is growing clamor for such agreements. The objective of this paper is to analyze, from a game-theoretic perspective, the circumstances in which binding international agreements may be expected to produce the best outcome in terms of the sum of national actions taken to address an international pollution problem. It is shown that there is a range of situations in which, given the costs of establishing and maintaining binding agreements, other forms of cooperation (either implicit or weaker than binding) can produce results at least as satisfactory as binding agreements.

The analysis is cast in terms of a two-country, two-commodity model, where each country produces both commodities but only exports one of them (and imports the other), and the production of only one of the commodities (produced in both countries) generates pollution. The policy instrument available to each country for dealing with transfrontier pollution is a tariff.[5] It is important to stress at the outset that a "pollution tariff" is a second-best policy, since it is imposed in the interests of the importing country and does not take account of global welfare. Although, in principle, a tariff can play a somewhat similar role in addressing transfrontier externalities as a Pigouvian tax performs within a country, they are not perfect substitutes. Take, for example, a polluting country that levies an emission tax to close the wedge between social and private costs, while another affected country imposes a tariff equal to its marginal damage. The resulting tax and tariff combination will generally not yield the resource allocation that would have resulted if the polluting country had imposed an internationally optimal Pigouvian tax on its emissions, or, in other words, a tax equal to marginal damage in all countries taken together.

The model is presented in Section II. Each country is induced by the other to follow a cooperative strategy, under threat of the imposition of a positive "pollution tariff" if either party fails to internalize

[4] For critical treatments of an uncritical predilection for international agreements, see Lal (1989), Hahn and Richards (1989), and Robertson (1990). In the field of monetary policy, there has been a fair amount of work, generally showing that the gains from coordination are small. See: Canzoneri and Minford (1986, 1987), Hughes-Hallett (1986a, 1986b, 1987), Oudiz and Sachs (1984) and Currie, Levine and Vidalis (1987).

[5] The use of a "pollution tariff" against offending imports has been commonly used in the literature. See, for example, Baumol and Oates (1975).

appropriately its pollution externality. Note that both countries are assumed to possess a credible threat by way of the "pollution tariff", and this paper does not address the nature of international cooperation on environmental matters between large and small countries. The standard Pigouvian tax structure is derived for each country in Section III, and the optimal tariff in the face of externalities (and conjectural variations) is determined in Section IV. In Sections IV and V, the Nash (noncooperative) and efficient (cooperative) solutions are calculated in the context of a one-shot game. This analysis leads to the standard Prisoners' Dilemma result, and noncooperation produces an equilibrium with a positive level of tariffs. Both countries would be better off if they cooperated with one another, but the incentive for unilateral action (cheating) exceeds that offered by cooperation. This analysis, based on a one-shot game, would give justification for binding international agreements.

But one-shot games do not allow countries to take advantage of quid pro quo strategies. They eliminate any possibility of the players learning to cooperate, in the knowledge that they will be accountable in a subsequent period for their behavior in the current period. Sections VI and VII develop a multi-period scenario, in which an intertemporal trigger mechanism can be applied to punish deviant behavior in a previous period. It is the presence of the trigger mechanism (credible threat), taking the form in this case of a punitive "pollution tariff", that allows an efficient outcome to be achieved in a noncooperative game.

Two trigger mechanisms are developed in the paper. These punitive responses are triggered when transfrontier pollution exceeds a certain determinate level "h" in a given period. When h is reached, the temptation to defect from the efficient solution outweighs deterrence, and the trigger mechanism will fail to create the right incentive. The first trigger mechanism involves reversion to the Nash noncooperative equilibrium, such that the value of temptation is reduced to zero whenever the temptation to cheat exceeds the deterrence from cheating. Under the second trigger mechanism, tariffs are continuously adjusted in order to keep temptation and deterrence equal. This amounts to better than a Nash reversion, and is associated with a higher h, which may be conceptualized as a higher reward for not cheating.

In theory, the presence of a trigger mechanism allows players to achieve the same efficient outcome that would result from cooperation when the game is played is infinitely repeated.[6] Strictly speaking, this equivalence only holds in a deterministic game, where rational players are operating with full information. As soon as a stochastic, or random, element is introduced into the game, such that outcomes are not known in advance, then a high level of transfrontier pollution in one period could provoke the application of punisment that is embeded in the trigger mechanism. Since a binding agreement is a proxy for the trigger mechanism that would supposedly ensure that h was never reached, such an agreement may be shown to be superior to a noncooperative situation in which a punishment is likely to be triggered.

The point at which a binding agreement becomes superior to a noncooperative repeated game depends crucially on the level of h and the likelihood that it will be attained in a noncooperative game. The value of h is determined by several factors, including the discount rate in the country concerned (a high discount rate is associated with a high h), the spread between gains from cheating and gains from cooperation, and the spread between the efficient solution and the Nash equilibrium. In addition, the behavior of players in terms of their policy actions will determine where games will tend to be played on the spectrum between entirely random and fully deterministic outcomes. Noncooperative players will have a better chance of avoiding outcomes above h the more they know about its value and about the nature of polluting activities that produce outcomes tending towards h.

The two central conclusions that this paper comes to are, first, that there exists a range of circumstances in which a noncooperative policy setting may be at least as good and perhaps better than a binding international agreement. Second, given the disadvantages and the costs of binding agreements that make this so, there are grounds for considering different forms of interaction among countries that will raise the likelihood of noncooperative games producing efficient outcomes. One such form of cooperation could involve consultations and the exchange of information among governments about their national policies and their expectations of the policies of other countries.

---

[6] A reputational approach could also be used. See Kreps, Milgrom, Roberts and Wilson (1987) for a discussion of this approach.

II- The Model

The model presented here is an extension of that of Vandendorpe (1972), Markusen (1975) and Thursby and Jensen (1983). Vandendorpe (1972) developed optimal tax structures in a model with traded and non-traded goods without any externalities. Markusen (1975) calculated optimal tax structures in the presence of externalities but did not consider any strategic interactions. Finally, the model borrows from the works of Thursby and Jensen (1983a) and (1983b) the idea of the existence of a consistent conjectural variation in the strategic tariff equilibria.

The world is made up of two countries, the home country and the foreign country. Variables with asterisks are foreign country variables. The two countries are assumed to be producing and trading two goods (x,y) and (x^*, y^*), and pollution is a function of the domestic and foreign production of good x. We ignore the waste being released into the environment by the consumption sector. We further assume that pollution does not affect the production functions and there is no possibility of substitution among inputs or outputs such that the amount of pollution resulting from a given level of production can be varied. Finally, we assume pollution, G, to be additive and given by:

$$
\begin{aligned}
G &= G(x,x^*) \quad and \\
dG &= G_1 dx + G_2 dx^*
\end{aligned}
\tag{1}
$$

where G_i represents the partial derivative of G with respect to its ith argument. The special case of $G_2 = 0$ refers to the case in which the externality is purely domestic.

Home and foreign production possibility curves are assumed to be concave and given by:

$$
\begin{aligned}
F(x,y) &= 0 \\
F^*(x^*,y^*) &= 0
\end{aligned}
\quad or \quad
\begin{aligned}
y &= T(x) \\
y^* &= T^*(x^*)
\end{aligned}
\tag{2}
$$

where F^i and T^i (i=home, foreign) are twice continuously differentiable and concave and their partial derivatives are negative.

Let p and p^* be the producer prices of good x in terms of y in home and foreign countries respectively. Along the production transformation curves, it must be the case that:

$$p = \frac{F_x}{F_y} = -T_x$$

and

$$p^* = \frac{F_x^*}{F_Y^*} = -T_x^* \qquad (3)$$

$$pdx = -dy$$
$$p^*dx^* = -dy^* \qquad (4)$$

Following Vandendorpe (1972), the differentials of equation (4) can be written as:

$$dp = -T_{xx}dx \qquad \text{or} \qquad dx = (-T_{xx})^{-1}dp = Rdp$$
$$dp^* = -T_{xx}^*dx^* \qquad\qquad dx^* = (-T_{xx}^*)^{-1}dp^* = R^*dp^* \qquad (5)$$

Let q and q^* be the domestic consumers' prices of good x in terms of y in home and foreign countries respectively. These are related to domestic producers' prices by:

$$q = (1+\theta)p$$
$$q^* = (1+\theta^*)p^* \qquad (6)$$

Define excess demand (E) as the difference between consumption and production. For the import good E > 0, while for the export good E < 0. Without loss of generality, assume that the home country exports good x in return for good y:

$$x + E_x = C_x \qquad y + E_y = C_y$$
$$x^* + E_x^* = C_X^* \qquad y^* + E_y^* = C_y^* \qquad (7)$$

Assuming that each country may levy a tariff on its imports, then the balance of trade conditions in the two countries are given by:

$$E_y + q\tau_y E_x = 0$$
$$\tau_x E_y^* + q^* E_x^* = 0 \qquad (8)$$

where

$$\tau_x = 1 + t_x$$
$$\tau_y = 1 + t_y \qquad (9)$$

and (t_x, t_y) are the tariffs that the foreign and the home countries impose on imports of goods x and y respectively. Assuming transport costs to be zero, then the world price ratio (π) is given by:

$$\pi = \frac{q^*}{\tau_x} = q\tau_y \quad and$$
$$\hat{q}^* = \hat{p}^* = \hat{\pi} + \hat{\tau}_x \qquad (10)$$
$$\hat{q} = \hat{p} = \hat{\pi} - \hat{\tau}_y$$

where a hat over a variable denotes percentage change.

The market clearing conditions in the world market for the two goods are given by:

$$E_x + E_x^* = 0$$
$$E_y + E_y^* = 0 \qquad (11)$$

Using the relationships described in equation (10), it is clear from equation (11) that the goods' markets are not independent; by Walras' law, equilibrium in one implies equilibrium in the other, so that one equation, say the first one, can be dropped.

Finally, each government is assumed to implement a social welfare function in which lump sum payments redistribute income such that we can express the welfare of each country as a function of the total quantity of each good consumed and pollution. It is assumed that pollution can be characterized as a pure public bad. Home and foreign welfare are given by:

$$W = W(C_x, C_y, G)$$
$$W^* = W^*(C_x^*, C_y^*, G) \qquad (12)$$

where the functions W and W^* are assumed to be twice differentiable and quasi-concave with $W_i > 0$ ($i = x, y, x^*, y^*$), $W_G, W_G^* < 0$, $W_{ii}, W_{ii}^* < 0$, $W_{GG}, W_{GG}^* < 0$, and $W_{ij}, W_{ij}^* = 0$. For the sake of simplicity all cross derivatives are assumed to be zero. Marginal utility of consumption of each commodity is positive but decreasing, whereas marginal disutility of pollution is increasing.

III- Derivation of The Optimal Production Tax

The presence of an externality creates a divergence between the domestic rate of substitution in consumption (RSC) and the domestic rate of transformation in production (RTP). Under these circumstances Bhagwati and Ramaswami (1963) demonstrated that the Paretian first-best policy is to intervene with a tax at the point at which the distortion occurs. Thus a Pareto optimal state in the home country is defined where:

$$\frac{\dfrac{\partial W}{\partial C_x} + \dfrac{\partial W}{\partial G}\dfrac{\partial G}{\partial C_x}}{\dfrac{\partial W}{\partial C_y}} = p \tag{13}$$

$$or \quad q + q_G G_1 = p, \quad q_G < 0$$

Similarly, in the foreign country:

$$\frac{\dfrac{\partial W^*}{\partial C_x^*} + \dfrac{\partial W^*}{\partial G}\dfrac{\partial G}{\partial C_x^*}}{\dfrac{\partial W^*}{\partial C_y^*}} = p^* \tag{14}$$

$$or \quad q^* + q_G^* G_2 = p^*, \quad q_G^* < 0$$

where q_G (q_G^*) is the social marginal rate of substitution between pollution and good y in the home (foreign) country, and where we have made use of the competitive equilibrium condition $W_x^i/W_y^i = q^i$ (i= home, foreign).

Equations (13) and (14) reveal that the optimal production tax is a simple joint product optimality condition. The home (foreign) producers receive a price p (p^*) per unit of production which equals the sum of the social values of the joint products produced.

The left-hand side of equation 13 (and 14) is by definition the RSC in the presence of externality, while the right-hand side represents the RTP which is given in equation (3).

Using equation (6), the optimal production tax in the home and foreign countries is given by:

$$\theta = -\frac{q_G G_1}{p} > 0 \quad and \quad \theta^* = -\frac{q_G^* G_2}{p^*} > 0 \tag{15}$$

IV- Determination of the Optimal Tariff with Externalities and Conjectural Variations

In this section we characterize the set of Nash equilibria when countries choose tariffs as punishment. These equilibria will serve as credible punishments (i.e., they are subgame perfect) in the dynamic game considered in Section VI, the threat of which can support tacit cooperation in a repeated setting. The results derived here are illustrated in Figure 1 below.

Consider first the home country. Substituting equations (4), (5), (7), and (13) in its welfare function (equation 12) and totally differentiating it yields the first order condition for an optimum:

$$\frac{dW}{\partial W / \partial C_y} = q dE_x + dE_y + q_G G_2 R^* dp^* = 0 \tag{16}$$

Next, using the differential of the home balance of payments (equation 8) and grouping terms yields:

$$\begin{aligned} &(\tau_y - 1)\hat{E}_y + \hat{\tau}_y + \hat{q} + \gamma^* \hat{q}^* = 0 \\ &where \quad \gamma^* = -\frac{q_G G_2 R^*}{(1+\theta^*)} > 0 \end{aligned} \tag{17}$$

Next, differentiate the foreign balance of payments equation and using equation (11) and the relation between domestic prices and the world price (equation 10), equation (17) becomes:

$$\hat{\pi}\tau_y + \tau_y \hat{E}_x^* - \hat{E}_x^* + \hat{\pi}\gamma^*(1 + \frac{\hat{\tau}_x}{\hat{\pi}}) = 0 \tag{18}$$

Finally, using equation (9), the optimal tariff formula is given by:

$$\begin{aligned} &t_y^N = \frac{1}{\epsilon^* - 1} + \frac{\gamma^*}{\epsilon^* - 1} + \frac{\gamma^*}{\epsilon^* - 1}\frac{\hat{\tau}_x}{\hat{\pi}} \\ &where \quad \epsilon^* = -\frac{\hat{E}_x^*}{\hat{\pi}} \end{aligned} \tag{19}$$

with ε^* the elasticity of the foreign country's demand for imports along its offer curve. This equation, except for the last two terms, is identical to the one derived by Thursby and Jensen (1983a).

An analogous derivation for the foreign country yields:

$$t_x^N = \frac{1}{\epsilon - 1} + \frac{\gamma}{\epsilon - 1} - \frac{\gamma}{\epsilon - 1}\frac{\hat{\tau}_y}{\hat{\pi}}$$
$$where \quad \epsilon = \frac{\hat{E}_y}{\hat{\pi}} \quad and \quad \gamma = -\frac{q_G^* G_1 R}{(1+\theta)} > 0 \tag{20}$$

with ε the elasticity of the home country's demand for imports along its offer curve.

Using Slutsky's decomposition of ε^* and ε as in Thursby and Jensen (1983a):

$$\epsilon^* = \eta^* + \frac{1}{\beta}\frac{\hat{\tau}_x}{\hat{\pi}}(\sigma^* + \frac{\xi^*}{\tau_x}) \quad and$$
$$\epsilon = \eta - \frac{1}{\rho}\frac{\hat{\tau}_y}{\hat{\pi}}(\sigma + \frac{\xi}{\tau_y}) \tag{21}$$

where

$$\eta^* = [\sigma^* + \frac{\xi^*}{\tau_x} + \frac{m^*}{\tau_x}]\frac{1}{\beta} \qquad \eta = [\sigma + \frac{\xi}{\tau_y} + \frac{m}{\tau_y}]\frac{1}{\rho}$$
$$\beta = 1 - \frac{m^*(\tau_x - 1)}{\tau_x} \qquad \rho = 1 - \frac{m(\tau_y - 1)}{\tau_y} \tag{22}$$

and

σ^i = compensated relative price elasticity of demand for the importable good

m^i = marginal propensity to consume the importable

ξ^i = elasticity of supply of the exportable, i = home, foreign.

Equations (19) and (20) define respectively the home and foreign country's reaction functions (RR and R^*R^*) in the (t_x,t_y) space of Figure 2.[7] The Nash reaction function tells a country how it should set its tariff to maximize its welfare, given the tariff of the other country and the size of the spillover. The

[7] A similar figure is used by Mayer (1981).

intersection of these two reaction functions gives the Nash equilibrium of the noncooperative tariff game. Since ε^* enters the home country optimization problem, we can interpret $\hat{\tau}_x/\hat{\pi}$ as the home country's conjecture about how the foreign country will react to a change in its terms of trade induced by a change in t_y. As equation (19) stands, in order for the home country to solve for its optimal tariff, it has to make a conjecture about the form of $\hat{\tau}_x/\hat{\pi}$. Likewise, the calculation of the foreign country's optimal tariff necessitates a conjecture on the part of the foreign country about how the home country will react to a change in its terms of trade induced by a change in t_x ($\hat{\tau}_y/\hat{\pi}$). In what follows, we will assume, as in Thursby and Jensen, that these "conjectural variations" are constant with C and C^* denoting the home and foreign country, respectively. Note that setting $C = C^* = 0$ in equations (19) and (20), respectively, (i.e., no retaliation is anticipated) yields the traditional Cournot optimal tariff. However, in the more general case where C and C^* are nonzero constants, each country's optimal tariff is given by substitution of equations (21) and (22) into equations (19) and (20) which, after some manipulation, yields [8]:

$$
\begin{aligned}
& t_y = \frac{[1+\gamma^*(1+C)][1+\mu^* t_x]}{v^*(C)+\delta^*(C)t_x} \qquad t_x = \frac{[1+\gamma(1+C^*)][1+\mu t_y]}{v(C^*)+\delta(C^*)t_y} \\
& \textit{where} \qquad v^*(C) = (\sigma^*+\theta^*)(1+C)-(1-m^*) \\
& \qquad\qquad\quad \delta^*(C) = \sigma^*(1+C)-(1-m^*) \\
& \qquad\qquad\quad \mu^* = 1-m^* \\
& \textit{and} \qquad\ \ v(C^*) = (\sigma+\theta)(1-C^*)-(1-m) \\
& \qquad\qquad\quad \delta(C^*) = \sigma(1-C^*)-(1-m) \\
& \qquad\qquad\quad \mu = 1-m
\end{aligned}
\tag{23}
$$

Now consider the case where the home country contemplates raising its tariff. It expects $\hat{\pi}/\hat{\tau}_y > 0$ if the foreign country does not retaliate. However, the foreign country will have to increase its tariff t_x in order to offset the worsening of its terms of trade ($\hat{\pi} > 0$) caused by the increase in the home country tariff rate. Therefore, if the home country believes that the foreign country is maximizing its utility, then the home country should conjecture $C = \hat{\tau}_x/\hat{\pi} > 0$ where the change in π is induced by the change in t_y.

[8] Similar games with non-zero conjectural variations have been studied by Cornes and Sandler (1984, 1985).

Likewise, if the foreign country expects $\hat{\pi}/\hat{\tau}_x < 0$ in the absence of retaliation but believes that the home country will increase t_y in an attempt to offset the worsening of its terms of trade, then the foreign country should conjecture $C^* = \hat{\tau}_y/\hat{\pi} < 0$. Thus, given our assumption of constant conjectural variations, these must satisfy $C \geq 0 \geq C^*$.

The first term in either equation (19) or (20) is the well known optimal tariff for the two good case in the absence of externalities.[9] The last two terms in either of these equations result from the environmental spillover that each country inflicts on the other. Both terms are positive (ε and ε^* are both greater than one [10] and $C \geq 0 \geq C^*$). These two terms will vanish if the spillovers are zero (this is equivalent to setting γ^* in equation (19) and γ in equation (20) to zero). Finally, note that if the victimized country faces a foreign offer curve which is infinitely elastic, then its optimal tariff is equal to zero.

Before concluding this section, we need to investigate how changes in each country's conjecture affect the equilibrium tariff levels. It turns out that a sufficient, though not necessary, condition for the equilibrium tariffs to decline as C and $|C^*|$ increase is that neither country's best response for a given conjecture is to increase its tariff when the other country lowers its tariff.[11] A graphical representation of the above results is given in Figure 1.

The offer curve construction is particularly useful in describing optimal tariffs with externalities. The curve OCBA represents the home offer curve, while that labelled OC'B'A represents that of the foreign country. The free trade relative price of good x is given by line OA and thus amount (a) of good

[9] See for example, Bhagwati and Srinivasan (1983) or Takayama (1974).

[10] Recall that the optimal tariff of one country refers to a point on the elastic portion of the other country's offer curve.

[11] For a formal proof of this proposition, see Thursby and Thursby (1983a).

x is imported by the foreign country. The curves I_iI_i ($I^*_iI^*_i$) and J_iJ_i ($J^*_iJ^*_i$) represent home (foreign) trade indifference curves, with the former, but not the latter, taking the externality into account. Each country is better off the higher the value of the subscript.

The locus of tangency points between the home and foreign trade indifference curves (II and I^*I^*) traces the contract curve PP, while that between JJ and J^*J^* traces the contract curve P'P'. Thus P'P' represents the "true" locus of Pareto optimal points. The contract curve P'P' intersects the foreign offer curve at point D, which represents the market equilibrium point that would prevail if market prices were adjusted to reflect the full value of the social costs associated with the production of good x. Each country is better off the higher the value of the subscript.

Assume initially that the foreign country does not take the externality into account and that it does not anticipate any retaliation from the home country. Then the foreign country optimal tariff is at point B where the home offer curve is tangent to the foreign trade indifference curve ($I^*_2I^*_2$), whose slope is equal to W_x/W_y. Point B' must lie to the right of point D since at B' the full social cost of producing good x is not taken into account. Point B' corresponds to levying a tariff that will rotate the terms of trade line from OO^1 to OO^2. In terms of equation (19), this is equivalent to setting a tariff $t_x = 1/(\varepsilon-1)$ where $\varepsilon = \eta$. Thus, the restricted volume of imports is now at point b.

The inclusion of the externality that the home country inflicts on the foreign one generates an indifference curve (JJ) that is flatter than the II curve (compare the slope of the II curve given above with the slope of the JJ curve which is given by equation (13)). The trade indifference curve $J^*_1J^*_1$, the J curve through B, lies above the offer curve to the left of point B, and so better points for the importer must also lie to the left of point B. Point C' must lie to the left of point D for, as was stated earlier, D is the market equilibrium point that would prevail if market prices were adjusted to reflect the full value of the social costs of producing good x, and C' is the point that corresponds to B' in these circumstances. Thus, the level of imports that best serves the foreign country's interest is smaller than D, the internationally optimal level of imports. The tariff corresponding to C is indicated by the relative slope of OO^3 and is given by equation (19), where again $\varepsilon = \eta$ (with $C^* = 0$).

Finally, allowing the home country to retaliate against the foreign country's imposition of a positive tariff implies that the new t_x is smaller than where retaliation was assumed to be absent. This requires a flatter price line than that implied by line OO^3. The new price line OO^4 (not drawn) should intersect the foreign offer curve at a point like E, which lies between points C' and B'. As already noted, point B' must lie to the left of point A. For exactly the same reason, point C' must lie to the left of point E. The exact location of point E with respect to point D, however, depends on whether the threat of retaliation can force the foreign country to adopt a taxing policy that is efficient from an international perspective.

V- The Efficient Outcome and Cooperation

We begin by transforming the welfare functions given in equation (12) from the commodity space into the tariff space.[12] This is because we are interested in the relation between economic welfare and tariff rates, thus [13]

---

[12] This transformation is possible only if trading equilibria are unique, which we will assume.

[13] One can transform $\partial W/\partial t_y$ and $\partial W^*/\partial t_x$ into the optimal tariff formulae that are specified in equations (19) and (20). This is the approach taken by Mayer (1981), though his model does not contain any externalities.

$$W = W(t_y, t_x; G) \qquad W^* = W^*(t_y, t_x; G) \tag{24}$$

the functions W and W^* are twice differentiable and quasi-concave with the first derivative with respect to own tariff being positive and otherwise negative. Equation (24) gives the home (foreign) country's maximum attainable utility, given the home and the foreign country's tariff rates (t_x, t_y) for every G. These W and W^* contours are described in Figure 2 below.

The Pareto-efficient combinations of tariff policies require that the relative price of x be the same in the two countries. Equating q and q^* in equation (10) yields:

$$\tau_y^E \tau_x^E = 1 \qquad or \qquad (1 + t_y^E)(1 + t_x^E) = 1 \tag{25}$$

Equation (25) reveals that free trade (i.e. $t_x = t_y = 0$) is on the Pareto-efficient loci. In terms of Figure 1, this is equivalent to both countries setting production taxes that are internationally optimal and which will shift the home offer curve outward and the foreign offer curve inward so that they intersect at point A' where the contract curve P'P' crosses the price line OO^1. Notice that equation (25) gives rise to a continuum of efficient policies where one country or the other subsidizes imports. Thus we have the traditional multiplicity of Pareto- efficient equilibria.[14] The set of Nash and Pareto equilibria are described in Figure 2 below.

Figure 2 traces three different sets of contours. The first set refers to the tariff indifference curves which are derived from the offer curve diagram of Figure 1. Under free trade, the equilibrium trading point in Figure 1 (point A' that takes externalities into account) corresponds to the tangency of the home and foreign trade indifference curves (the JJ curves). As both countries' tariffs are zero, the corresponding

[14] It is well known that multi-period games generally have a multiplicity of Pareto-improving solutions. See for example Rogoff (1987).

point in the tariff plane of Figure 2 is O'. The other points on the tariff indifference curve W_2 are obtained by varying the tariff combinations t_x, t_y while holding the foreign trade indifference curve through point A constant. Using equation (23), the slope of this tariff indifference curve is given by:

$$\frac{\partial t_x}{\partial t_y}\Big| = \frac{[1+\gamma(1+C)]}{\{\delta^*(C)t_y - \mu^*[1+\gamma(1+C)]\}^2}[\mu^* v^*(C) - \delta^*(C)] \gtreqless 0 \quad as \quad \mu^* v^*(C) \gtreqless \delta^*(C) \tag{26}$$

and that of the foreign country is given by:

$$\frac{\partial t_x}{\partial t_y}\Big|^* = \frac{[1+\gamma^*(1-C^*)]}{[v(C^*)+\delta(C^*)t_y]^2}[\mu v(C^*) - \delta(C^*)] \gtreqless 0 \quad as \quad \mu v(C^*) \gtreqless \delta(C^*) \tag{27}$$

Notice that the slopes of the home and the foreign tariff indifference curves (equations (26) and (27) respectively) are positive, zero or negative depending on whether each country's tariffs are less than, equal to or larger than its optimal tariff. The other tariff indifference curves in Figure 2 are obtained by choosing another point (such as B' or C') and repeating the same process described above. Again, each country is better off the higher the value of the subscript.

The second set of contours refers to the home and foreign country's reaction functions RR and R^*R^*. These are derived using equations (19) and (20). The reaction functions pass through the tariff indifference curves where the latter have zero slopes. This is so because the reaction functions indicate one country's optimal tariff rates corresponding to the tariff rate prevailing in the other country, which is exactly what zero-sloped tariff indifference curves indicate.

The third and last contour in Figure 2 refers to the locus of the optimal combinations of t_x and t_y that are described by equation (24). This is the E^*E curve where the welfare of the home country falls steadily, and that of the foreign country rises, as we move from E to E^*. Points along the segment KL are preferred by both countries to the Nash equilibrium at N. Notice that the free trade equilibrium (point O') lies in the range KL.

If the tariff game between the two countries is played only once, then the efficient outcome (any point within the segment KL) can only be attained if the two countries precommit to the policies (τ_x^E, τ_y^E) in a binding agreement, because these tariffs are not on the countries' reaction functions. In Figure 2, if the foreign country plays a tariff equal to point P, then the home country will play the tariff on its reaction function equal to F, and thus obtain a higher welfare ($W_3 > W_1$). The foreign country will thus end up at a tariff indifference curve $W_0^* < W_3^*$. The home tariffs are higher at F than at P and welfare is greater. The foreign country knows that the home country will be so tempted, and therefore will not play P in the absence of a binding agreement.

VI- Dynamic Tariff Strategies

The analysis presented in the previous section suffers from one crucial limitation. It views countries as interacting only once: they are portrayed as players in a one-shot game. In applications of the theory of one-shot games, it is common to find that noncooperative behavior leads to inefficient outcomes. This prediction of inefficient outcomes has led some game theorists to question the appropriateness of one-shot game theory for analyzing many interactions. Noting that the same agents interact over and over again, they developed the theory of repeated games which predicts more efficient outcomes than the theory of one-shot games. This section applies recent development in repeated games to the question of environmental policy cooperation. The section draws on Friedman's (1971) work on trigger strategies.[15]

Trigger strategy equilibria explain how rational, self-interested agents who interact repeatedly can achieve Pareto-efficient outcomes even though binding agreements are not possible. Trigger strategy equilibria may be considered as self-enforcing agreements in which the enforcement mechanism is a threat made by all players that, in the event the agreement is violated, they will change their actions to those corresponding to a single period noncooperative equilibrium (the Nash outcome) with payoffs that are worse for every player than the cooperative payoffs.[16] Thus trigger mechanisms can eliminate the need for binding agreements.

Let the game described in the previous section be repeated again and again. The timing of events is as follows: at the beginning of each period the home and foreign countries learn the realization of G. They then simultaneously choose the levels of their tariffs, and these choices determine the outcome for that period. The strategic choices of the home and the foreign countries become common knowledge, and this one-period game is repeated after a new G is realized. In any given period, the policymakers in each country know that the game will be repeated by them or their successors. If, in this sense, policymakers live forever, we have an infinitely repeated game.

In period t, the home and the foreign countries maximize V_t and V_t^*, the discounted sums of current and expected future welfare:

---

[15] For a textbook exposition of trigger mechanisms, see Friedman (1990).

[16] *ibid*, p. 126.

$$V_t = W_t + E_t \sum_{i=1}^{\infty} \delta^i W_{t+i}$$
$$V_t^* = W_t^* + E_t \sum_{i=1}^{\infty} \delta^i W_{t+i}^* \tag{28}$$

where δ is the discount rate.

Both countries have already observed G_t and the tariff rates and can therefore calculate current welfare; future welfare depends upon the G's (and tariffs) that have yet to be realized.

One might try to apply Friedman's trigger mechanism in this setting, but the countries' intertemporal strategies will be complicated by the fact that the temptation to defect from the efficient solution fluctuates from period to period.[17] If the foreign country plays t_x^E the home country can gain:

$$Temptation = W^C(t_y^C, t_x^E; G_t) - W^E(t_y^E, t_x^E; G_t) \tag{29}$$

by playing t_y^C instead of t_y^E. The temptation to cheat is stochastic since a new G is realized in each period. If the support of G is unbounded, then so is temptation. Trigger mechanisms are intended to deter cheating, with deterrence measured by the value of obtaining the efficient outcome E rather than a bad outcome, say N, over all future periods. Since future values of G are not known, deterrence must be defined in terms of expected values. These expected values are bounded; so in periods in which a large G is drawn, temptation will be greater than deterrence, and the Friedman trigger mechanism will fail.

The trigger mechanism will not support the efficient point E in any period in which the level of G is too big. Even if no country had cheated, E can only be played in period t if

[17] A similar type of trigger mechanism has been developed in the macro literature by Canzoneri and Henderson (1988), and Safadi (1990).

$$G_t \leq h \tag{30}$$

where h is a parameter that will be determined below. Once we recognize this, we need to define cooperation less ambitiously; namely, play the efficient tariffs so long as the G is not too large; otherwise, in periods of extreme values of G, play the one-shot Nash tariff levels. Thus punishment is defined in the usual way, i.e., as Nash reversion. With this modified version of cooperation, the trigger mechanism just introduced is then the standard one: "cooperate as long as my opponent does likewise; otherwise punish."

More formally, consider the following strategies for repeated games. Strategy S_1 is to play t_y^N (t_x^N) each period. S_1 is a Nash strategy for the repeated game; if one country adopts it the other country can do no better than adopt it as well. In the Nash equilibria that results, N will be the outcome in every period. However, the trigger mechanism introduced above suggests that a better equilibrium also exists. Define strategy S_2 as follows: (1) in period 1, observe G_t and play t_y^E (t_x^E) if inequality (30) holds; otherwise play t_y^N (t_x^N). (2) In periods t = 2,3,..., play t_y^N (t_x^N) if the other country has cheated in the past; if not, play t_y^E (t_x^E) if inequality 30 holds and t_y^N (t_x^N) if it does not. This strategy recognizes that the simplest trigger mechanism does not provide the right incentives when G_t is large; it simply says to revert to N in those periods.

For a suitable value of h, S_2 is a Nash strategy of the repeated game. Suppose that the foreign country has adopted S_2, and ask whether the home country can do any better than adopt S_2 in response. If the home country does not adopt S_2, then E will be the outcome when $G \leq h$ and N will be the outcome otherwise. The only alternative worth considering is to cheat when $G \leq h$. But if h is chosen as the value of G with the property that temptation and deterrence are equal, and if $G \leq h$, the gain from cheating is less than the discounted expected future gain. Thus the home country will adopt strategy S_2. By symmetry, if the home country adopts S_2, the foreign country can do no better than adopt S_2 in response. S_2 is therefore a Nash strategy for the repeated game.

What remains is to characterize h. In particular, we need to ask whether, and under what circumstances, there exists a $h > 0$ that equalizes temptation and deterrence.

If the home country does not cheat between periods t and t+i , its expected welfare in period t+i is:

$$E_t W_{t+i} = \int_0^h W(t_y^E, t_x^E, G) f(G) dG + \int_h^\infty W(t_y^N, t_x^N, G) f(G) dG \tag{31}$$

where f(.) is the probability density function of pollution. We can now calculate the deterrence against cheating that is embodied in the trigger mechanism. If the home country cheats, N will be played in all future periods; if it does not cheat, E will be played whenever $G \leq h$. So, deterrence is the discounted sum:

$$K = [\frac{\delta}{1-\delta}][\int_0^h W^E f(G) dG + \int_h^\infty W^N f(G) dG - \int_0^\infty W^N f(G) dG]$$
$$or$$
$$K = [\frac{\delta}{1-\delta}]\int_0^h [W^E - W^N] f(G) dG \tag{32}$$

Since h is defined as the value of G with the property that the temptation to cheat against the efficient outcome equals the deterrence embodied in Nash reversion, we can write:

$$W^C(t_y^C, t_x^E; G_t) - W^E(t_y^E, t_x^E; G_t) = K \tag{33}$$

and the value of h is then determined by:

$$W^C - W^E = [\frac{\delta}{1-\delta}] \int_0^h [W^E - W^N] f(G) dG \qquad (34)$$

VII- A More Efficient Trigger Mechanism

Though the trigger mechanism developed in the previous section is better than a repeated play of the noncooperative equilibrium, it is not the best solution that one could find. Rotemberg and Saloner's (1986) work on an oligopoly example suggests that the players might be able to revert to a better point than N when temptation outweighs deterrence.

The basic idea in this section is that we can improve on the definition of cooperation that was given in the previous section. The modified definition of cooperation is: "Play the efficient tariffs so long as G is not too large; otherwise in periods of extreme G, choose tariff rates with the property that the temptation to cheat is just offset by the deterrence value of the punishment. " Note, of course, that what one means by a "large" value of G need not be the same as the one developed in the previous section. Indeed, as we show below, the efficient solution can be supported for higher values of G under this mechanism than under the one developed in the previous section.

With this modification, matters proceed as before. Punishment is defined as Nash reversion, and the new trigger strategy is: "Cooperate so long as my opponent does likewise; otherwise punish."

As before, let h be the value of G with the property that the temptation to cheat against the efficient outcome equals the deterrence embodied in Nash reversion; that is:

$$[\, W^C(t_y^C, t_x^E; G) - W^E(t_y^E, t_x^E; G)\,] = K \tag{35}$$

Consider separately the cases in which G is below and above h. In the former case neither country has the incentive to cheat on the efficient solution. Therefore, if we define $W^S(G;h)$ to be the maximum welfare the country can obtain,

$$W^S(G;h) = W^E(G), \qquad G \leq h \tag{36}$$

If G > h, the efficient outcome is not achievable since both countries have an incentive to cheat. In this case, define the maximum sustainable welfare as:

$$W^S(G;h) = W^R(G), \qquad G > h \tag{37}$$

where $W^S(G)$ is the level of welfare associated with the pair of tariffs (t_y^R, t_x^R) such that the temptation to cheat against an opponent's tariff equals K. Thus the tariff t_y^R is defined by:

$$W(t_y^C, t_x^R, h) - W(t_y^R, t_x^R, h) = K \tag{38}$$

where $t_y^C = \text{argmax}\ \{W(t_y, t_x^R, h)\}$; that is t_y^C is the best response that the home country will play when the foreign country is playing t_x^R. In summary:

$$W^S(G;h) = \begin{matrix} W^E(G) & for\ G \leq h \\ W^R(G) & for\ G > h \end{matrix} \tag{39}$$

Now we can turn to the punishment. The punishment equals the present discounted value of the extra welfare that the country would have earned had it not cheated, or:

$$K = (W^C - W^E)\,h = \frac{\delta}{1-\delta}\int_0^{\infty}[W^S(G) - W^N(G)]f(G)dG$$
$$\textit{that is}$$
$$W^C - W^E = \kappa\{\int_0^h[W^E(t_y^E,t_x^E,G) - W^N(t_y^N,t_x^N,G)]f(G)dG + \int_h^{\infty}[W^R(t_y^R,t_x^R,G) - W^N(t_y^N,t_x^N,G)f(G)dG]\} \tag{40}$$
$$\textit{where}$$
$$\kappa = \frac{\delta}{1-\delta}$$

What remains to be done is to prove the claim made at the beginning of this section; namely that the second trigger mechanism gives rise to an efficient outcome more often than the one developed in the previous section. This is equivalent to proving that the cutoff implied by the present trigger mechanism (call it h^R) is larger than that associated with the previous one (call it h^P).

Consider a trigger strategy using the earlier definition of cooperation with a cutoff value h^P, and the property that deterrence and temptation are equalized. Now switch to the present section definition of cooperation. At h^P, the temptation to cheat is unchanged. However, the punishment should any country cheat is increased. To reequalize temptation and deterrence, the cutoff value of h^P has to rise. That is, $h^R > h^P$.

VIII- Conclusions

There are no grounds for the presumption that binding international agreements offer the best way of dealing with environmental spillovers among countries. This is as true of localized transfrontier pollution as it is of action to address threats to the global commons. Among the drawbacks of binding international agreements are that they take time to negotiate, carry transactions costs, divert attention from options for domestic action, and tend towards minimalist outcomes. Countries acting on their own, against the background of mutually credible threats of retaliatory actions once a certain unacceptable level of polluting spillover is attained, may be expected over time to cooperate implicitly towards an efficient solution.

The analytical approach adopted in this paper does not lead to the conclusion that implicit cooperation is always better than explicit commitment. It depends on what degree of uncertainty there is about how countries will behave. In the game theoretic context, it is a matter of whether the outcomes of repeated games are random, or more deterministic in nature. The case for implicit cooperation made in this paper rests on the assumption that countries individually will pursue pollution abatement and control policies, and that such policies will be mindful of the policies and objectives of other countries. In such circumstances, repeated games will not throw up random outcomes, and the likelihood of triggering retaliation at some threshold level of pollution will be small. A stable and efficient outcome along these lines would be aided by international arrangements for the exchange of information.

FIGURE 1

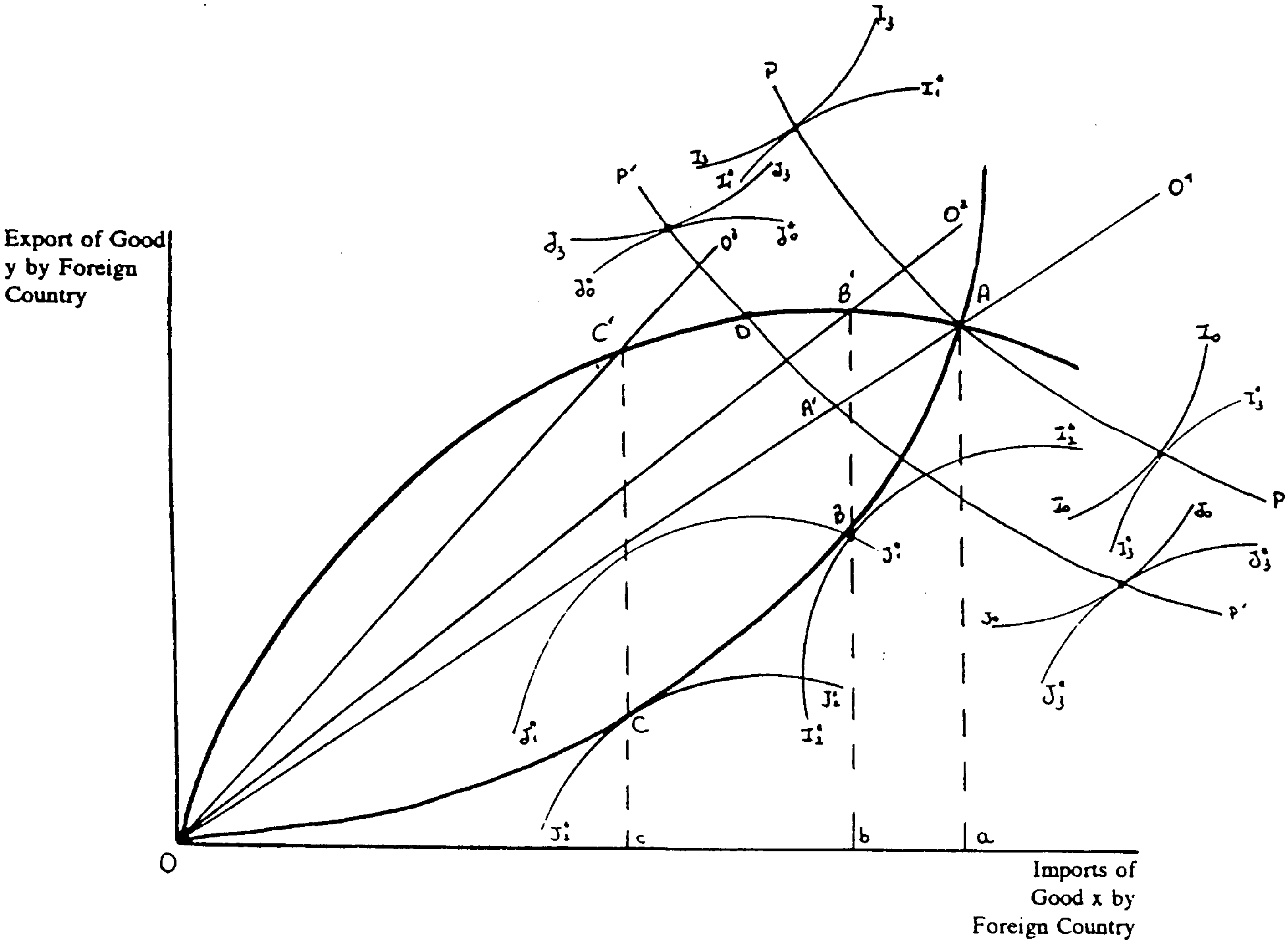
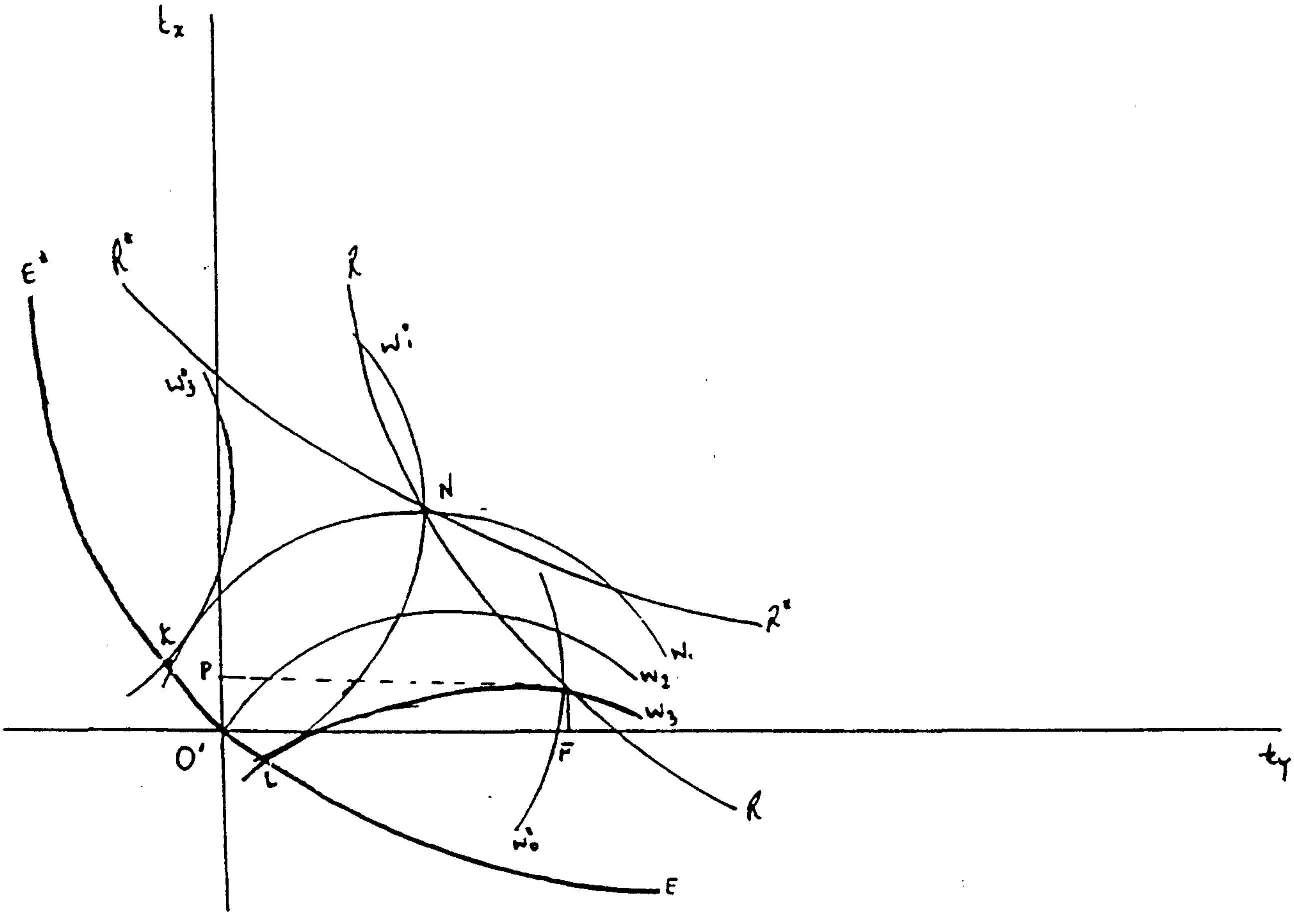

Export of Good y by Foreign Country
Imports of Good x by Foreign Country
O
A
B
C
A'
B'
C'
D
P
P'
a
b
c

Figure 2



References

Bhagwati, Jagdish and T.N. Srinivasan (1983), *Lectures on International Trade*, (MIT Press, Cambridge, MA).

Bhagwati, Jagdish and V.K. Ramaswami (1963), "Domestic Distortions, Tariffs and the Theory of Optimum Subsidy," Journal of Political Economy, Vol. 71(1), pp. 44-50.

Baumol, William J. and Wallace Oates (1975), *The Theory of Environmental Policy*, Prentice-Hall, Englewood Cliffs.

Bohm, Peter (1990), "Efficiency Issues and the Montreal Protocol on CFCs," Environment Working Paper No. 40, Sector Policy and Research Staff, The World Bank.

Canzoneri, Mathew and Dale Henderson (1988), "Is Sovereign Policymaking Bad?," *Carnegie-Rochester Conference Series on Public Policy*, Vol. 28, pp. 93-140.

Canzoneri, Mathew and Patrick Minford (1986), "When International Policy Coordination Matters: An Empirical Analysis," CEPR Discussion Paper No. 119.

Canzoneri, Mathew and Patrick Minford (1987), "Policy Interdependence: Does Strategic Behavior Pay?," CEPR Discussion Paper No. 201.

Cornes, R. and Sandler, T. (1984), "The Theory of Public Goods: Non-Nash Behaviour," Journal of Public Economics, Vol. 23, pp. 367-379.

Cornes, R. and Sandler, T. (1985), "Externalities, Expectations and Pigouvian Taxes," Journal of Environmental Economics and Management, Vol. 23, pp. 1-13.

Currie, David, Paul Levine and Nic Vidalis (1987), "Cooperative and Noncooperative Rules for Monetary and Fiscal Policy in and Empirical Two-Bloc Model," in Ralph Bryant and Richard Portes (ed.), *Global Macroeconomics: Policy Conflict and Cooperation*, (London: Macmillan).

Eskeland, Gunnar and Emmanuel Jimenez (1991), "Choosing Policy Instruments for Pollution

Control, A Review," Policy, Research and External Affairs Working Papers, The World Bank.

Friedman, James (1971), "A Non-Cooperative Equilibrium for Supergames," Review of Economic Studies, Vol. 28, pp. 1-12.

Friedman, James (1990), *Game Theory with Applications to Economics*, (New York, Oxford University Press).

Hahn, Robert and Kenneth Richards (1989), "The Internationalization of Environmental Regulation", Harvard International Law Journal, Vol. 30, No. 2, pp. 421-446.

Hughes-Hallett, Andrew (1986a), "Autonomy and the Choice of Policy in Asymmetrically Dependent Economies," Oxford Economic Papers, Vol. 38.

Hughes-Hallett, Andrew (1986b), "International Policy Design and the Sustainability of Policy Bargains," Journal of Economic Dynamics and Control, Vol. 10.

Hughes-Hallett, Andrew (1987), "The Impact of Interdependence on Economic Policy Design: The Case of the US, EEC and Japan," Economic Modelling, No. 4.

Kreps David, Paul Milgrom, John Roberts and Robert Wilson (1982), "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," Journal of Economic theory, Vol. 27, pp. 245-252.

Lal, Deepak (1989), "The Limits of International Co-Operation", University College London, Department of Economics, Discussion Papers in Economics, No. 89-21.

Markusen, James (1975), "International Externalities and Optimal Tax Structure," Journal of International Economics, Vol. 5, pp. 15-29.

Mayer, Wolfgang (1981), "Theoretical Considerations on Negotiated Tariff Adjustments," Oxford Economic Papers, Vol. 33, pp. 135-153.

Oudiz, Gilles and Jeffrey Sachs (1984), "Macroeconomic Policy Coordination Among the Industrial Countries," Brookings Papers on Economic Activity, No. 1, pp. 1-64.

Robertson, David (1990), "The Global Environment, Are International Treaties a Distraction?", The World Economy, Vol. 13, No. 1.

Rogoff, Kenneth (1987), "Reputational Constraints on Monetary Policy," in K. Brunner and A. Meltzer (eds.), *Bubbles and Other Essays*, Carnegie-Rochester Conference Series on Public Policy, Vol. 26, pp. 141-181.

Rotemberg Julio and Garth Saloner (1986), "A Super-Game Theoretic Model of Price Wars During Booms," American Economic Review, Vol. 76, pp. 390-407.

Safadi, Raed (1990), Conflict Between Monetary Authorities: Trigger Strategies in Stochastic Supergames, (Ph.D. Dissertation, Georgetown University, Department of Economics, Washington, D.C.).

Takayama, A. (1974), *International Trade*, (Holt Rinehart and Winston, New York).

Thursby, Marie and Richard Jensen (1983a), "A Conjectural Variation Approach to Strategic Tariff Equilibria," Journal of International Economics, Vol. 14, pp. 145-161.

Thursby, Marie and Richard Jensen (1983b), "Consistent Conjectural Tariff Equilibria," (Ohio State University, Department of Economics Working Paper No. 83-10).

Vandendorpe, Adolf (1972), "Optimal Tax Structures in a Model with Traded and Non-Traded Goods," Journal of International Economics, Vol. 2, pp. 235-256.

Whalley, John (1991), "The Interface Between Environmental and Trade Policies", Economic Journal, Vol. 101, No. 405, pp. 180-189.

Winter, Alan (1991), "The Trade and Welfare Effects of Greenhouse Gaz Abatement, A Survey of Current Estimates," in K. Anderson and R. Blachurst (eds.), *The Greening of World Trade Issues*, (Harvester-Wheatsheaf, for the GATT).

ECONOMIC
RESEARCH
FORUM
FOR THE ARAB COUNTRIES, IRAN & TURKEY

7 Boulos Hanna St., Dokki, Cairo, Egypt
Tel: (202) 700810, Fax: (202) 3616042, E Mail: ERF @ vrit 01 eg.