An Equitable, Efficient and Implementable Scheme

to Control Global Carbon Dioxide Emissions

By

ARTHUR J. CAPLAN*
Associate Professor of Economics
Department of Economics
Utah State University
3530 Old Main Hill
Logan, Utah 84322-3530, USA
Phone: (435) 797-0775
Email: acaplan@econ.usu.edu

and

EMILSON C. D. SILVA
Professor of Economics
School of Economics
Georgia Institute of Technology
Atlanta, Georgia 30302-0615, USA
Phone: (404) 894-4919
Email: emilson@tulane.edu

An Equitable, Efficient and Implementable Scheme
to Control Global Carbon Dioxide Emissions

September 13, 2012

Abstract: We design an international scheme to control global externalities in which autonomous regions choose their own emissions levels in anticipation of interregional resource transfers implemented by an international agency. This agency follows a proportional equity principle, which preserves the status-quo ratio of regional welfare levels. We show that it is individually rational for each region to participate in the proposed international scheme and that regional environmental authorities choose policies that fully internalize the global externality. Although based on an admittedly ideal scheme, these results are especially noteworthy in light of the call for various forms of transfers in international agreements such as the Kyoto Protocol.

Keywords: proportional equity, efficiency, implementability, global externality; interregional transfers

1. Introduction

The Kyoto Protocol to the United Nations Framework Convention on Climate Change (called "Convention" hereafter), completed on December 10, 1997, will probably be remembered most for selecting emissions trading as the main mechanism to control global greenhouse gas emissions. However, it will also be remembered for the promulgation of another type of incentive: international transfers between developed and developing regions. These transfers are intended to: (1) "provide new and additional financial resources to meet agreed full costs incurred by (developing) countries in advancing the implementation of existing commitments;" and (2) "provide such financial resources, including for the transfer of technology, needed by (developing) countries to meet the agreed full incremental costs of advancing the implementation of existing commitments" (Article 12 of the protocol). The Conference of the Parties to the Convention, the Convention's supreme body, has delegated to the Global Environment Facility (GEF) the responsibility of transferring resources from developed to developing regions. However, the details about how the GEF might implement such a transfer scheme have not yet been formalized in any agreement. The absence of formal details on this issue therefore begs the question of how to ideally design an efficient and individually rational (implementable) transfer scheme.

In this paper, we investigate the efficiency and implementability properties of an ideal international scheme designed to control global externalities such as carbon dioxide emissions. In the proposed scheme, participating regions select their most desirable contributions to the global externality fully anticipating that an international agency, say the GEF, will implement resource transfers from wealthy to poor regions in accordance with a particular equity principle. We demonstrate that if the GEF's objective function obeys a proportional equity principle, which preserves the status-quo relative ranking of international welfare levels, both wealthy and poor regions will have incentives to: (1) voluntarily participate in the international scheme; and (2) efficiently control their global emissions. The equilibrium allocation of resources for the global economy in the presence of the international scheme is proportionally equitable and Pareto efficient.[1]

Intuitively, the resource transfers promoted by the GEF in the proposed scheme align the incentives of the participating regions because the welfare of one region can rise only if the welfare of the other region rises. Since the autonomous regional governments are aware of this fact when they select their contributions to the global externality, they have incentives to choose regional environmental policies that fully internalize the externality. Previous research has shown that competing autonomous regional governments may behave efficiently when they anticipate that regional welfare levels will be equalized in equilibrium (Boadway (1982), Caplan and Silva (1999), Myers (1990), Silva (1997), Silva and Caplan (1997), Nagase and Silva (2000), and Wellisch (1994)).[2] Myers (1990) referred to this phenomenon as "perfect incentive equivalence." Analogously, research of the family has shown that children may behave efficiently (in terms of maximizing total family income) when they anticipate transfers from a benevolent parent that are linked directly to their economic choices (Bergstrom (1989), Cornes and Silva (1999), and Chiappori and Werning (2002)). Becker (1982) coined this phenomenon the "Rotten Kid Theorem."

This paper contributes to this literature by enlarging the set of circumstances under which one observes perfect incentive equivalence, or a Rotten Kid Theorem. Our results demonstrate that equalization of regional welfare levels is not *necessary* for perfect incentive equivalence. Rather, we claim that perfect incentive equivalence occurs in equilibrium whenever the welfare levels of competing autonomous regional governments are positively related (i.e., they are complements).

Although the focus of our analysis is not on emissions trading, this paper is closely related to the emerging literature on environmental markets. Chichilnisky, et al. (2000), for example, show that equity and efficiency go hand-in-hand whenever carbon dioxide emissions are traded. They demonstrate that emission markets will allocate resources efficiently if and only if international transfers are made in order to equalize social marginal utilities of consumption. They also show that this resource redistribution condition can be satisfied by an appropriate initial distribution of emission permits. In a similar vein, Caplan, et al. (2003) show that emissions trading can achieve an efficient resource allocation irrespective of the initial permit distribution, as long as interregional income transfers are determined by the GEF *after* the regions have chosen their respective permit demands.

Our analysis provides support for this recent view that equity and efficiency play complementary roles in a solution to problems such as global warming. Our results suggest that it is possible to achieve efficiency if international transfers are promoted to advance some equity principle. In our framework, the GEF's optimal strategy entails equalization of social marginal utilities of consumption in equilibrium. As in Caplan, et al. (2003), but unlike in Chichilnisky, et al. (2000), the resource redistribution takes place *after* the autonomous regional governments select their most desirable emission quantities. Indeed, it is this sequencing of moves that induces efficient behavior in the absence of a market for carbon emissions.

We introduce the basic model in the next section. In Section 3, we characterize the benchmark Pareto efficient solution. Section 4 presents two environmental policy games, a fully decentralized benchmark scenario in which the regions behave non-cooperatively, and the case where the regions participate in the proportional equity transfer scheme. Section 5 provides an example with three countries in order to demonstrate some of the limitations of our proposed international scheme as the number of potential participants is expanded. Section 6 concludes.

2. The Model

To begin with, consider an economy consisting of two regions indexed by j, j = 1,2, e.g., two coalitions of wealthy and poor regions, respectively. Each region has an autonomous government. There is one marketed commodity whose production results in emission of carbon dioxide (e.g., an industrial good). Let X_j be region j's industrial product and E be the total quantity of carbon dioxide emitted in the atmosphere. We assume that $E = X_1 + X_2$; that is, production of a unit of the industrial good leads to the emission of a unit of carbon dioxide.

The industrial sector in region j is competitive and consists of a large number of identical producers. Let k_j be the (fixed) number of industrial producers in region j. Each industrial producer utilizes an input quantity $z_{Xj} \leq 0$ of a numeraire good to produce $f^j\left(z_{Xj}\right)$ units of the industrial good. We assume that f^j is decreasing and strictly concave. Define $Z_{Xj} = k_j z_{Xj}$ as the total amount of the numeraire good demanded as input by region j's industrial sector and $F^j\left(Z_{Xj}\right) = k_j f^j\left(Z_{Xj}/k_j\right)$ as this sector's production function. Hence, $X_j = F^j\left(Z_{Xj}\right)$. If we normalize the price of the numeraire good to one and let *p* denote the price of the industrial good, the profit of the industrial sector in region j is denoted $pX_j + Z_{Xj}$.

Region j has a potentially large population n_j of identical residents. The utility of each consumer in region j is denoted $U^j\left(x_j, z_j, E\right)$, where x_j and z_j are the quantities consumed of the industrial and numeraire goods, respectively. We assume that U^j is strictly quasi-concave. It increases in the first two arguments and decreases in the last. Carbon dioxide emissions are, for example, harmful to each individual's health.

Let $I_j = pX_j^0 + Z_j^0 + pX_j + Z_{Xj} + T_j$ be region j's total income. The quantities X_j^0 and Z_j^0 denote region j's initial endowments of industrial and numeraire goods, respectively. The quantity T_j represents the total amount of income that this region receives from the other region (if positive) or remits to the other region (if negative) in order to satisfy the proportional equity principle underlying the international scheme examined below. Since the interregional transfers are purely redistributive, $T_1 + T_2 = 0$. Each consumer in region j faces a budget constraint $px_j + z_j = I_j/n_j$. The industrial good is freely traded in an international market. In any equilibrium for the global economy, $\sum_{j=1}^{2}\left(n_j x_j - X_j^0 - X_j\right) = 0$; namely, the international market must clear.

3. Pareto Efficiency

Before we analyze the making of environmental policy, it is useful to consider the conditions that characterize a Pareto efficient allocation. A Pareto efficient allocation can be obtained by choosing $\left\{x_j, z_j, X_j, Z_{Xj}\right\}_{j=1,2}$ to maximize $U^1\left(x_1, z_1, X_1 + X_2\right)$ subject to: $U^2\left(x_2, z_2, X_1 + X_2\right) \geq \overline{U}^2$ and

$X_j \leq F^j\left(Z_{Xj}\right)$, $\sum_{j=1}^{2}\left(n_j x_j - X_j^0 - X_j\right) \leq 0$, $x_j \geq 0$, $z_j \geq 0$, $\mathrm{X_j} \geq 0$, $Z_{Xj} \leq 0$, $j = 1,2$.

An interior Pareto efficient allocation satisfies:

$$U^2\left(x_2, z_2, X_1 + X_2\right) = \overline{U}^2 > 0, \tag{1a}$$

$$X_j = F^j\left(Z_{Xj}\right) > 0, \quad j = 1,2, \tag{1b}$$

$$\sum_{j=1}^{2}\left(n_j x_j - X_j^0 - X_j\right) = 0, \tag{1c}$$

$$\frac{U_x^1}{U_z^1} = \frac{U_x^2}{U_z^2} > 0, \tag{1d}$$

$$\frac{U_x^j}{U_z^j} = -\frac{1}{F_Z^j} - \sum_{i=1}^{2}\left(n_i \frac{U_E^i}{U_z^i}\right), \quad j = 1,2. \tag{1e}$$

Conditions (1a) – (1d) require no comment because they are fairly standard. Equations (1e) tell us that the marginal rate of substitution between industrial and numeraire goods for the representative consumer of region j must be equal to region j's social marginal rate of transformation between industrial and numeraire goods, which includes the global marginal negative effects brought upon by production of the industrial good in the region. From equations (1d) and (1e), it follows that $F_Z^1 = F_Z^2$; that is, we observe equalization of marginal products of the industrial good across regions.

4. Environmental Policy Making

Remember that the amount of carbon dioxide emitted in a region corresponds to the regional quantity of the industrial good produced. It is therefore reasonable to think that each regional government regulates the regional industrial product. Since, in our model, having control over a region's industrial product is equivalent to having control over a region's quantity of input demanded by the industrial sector, the policy instrument controlled by the regulator in region j is effectively Z_{Xj}. Regional regulators make their policy choices knowing how their consumers will behave. As such, it becomes imperative that we first consider the problem facing consumers.

Each consumer in region j chooses nonnegative $\left\{x_j, z_j\right\}$ to maximize $U^j\left(x_j, z_j, E\right)$ s.t. $px_j + z_j = I_j / n_j$, taking $\left\{p, I_j, E\right\}$ as given. An interior solution satisfies the budget constraint and $U_x^j / U_z^j = p$. Let $x_j\left(p, I_j, E\right)$ and $z_j\left(p, I_j, E\right)$ be the consumer's demand functions. It is convenient, however, to express these demands as functions of the policy variables and the industrial good's price:

$$x_j\left(p, Z_{X1}, Z_{X2}, T_j\right) \equiv x_j\left(p, Z_j^0 + pX_j^0 + T_j + pF^j\left(Z_{Xj}\right) + Z_{Xj}, \sum_{i=1}^{2} F^i\left(Z_{Xi}\right)\right),$$

$$z_j\left(p, Z_{X1}, Z_{X2}, T_j\right) \equiv z_j\left(p, Z_j^0 + pX_j^0 + T_j + pF^j\left(Z_{Xj}\right) + Z_{Xj}, \sum_{i=1}^{2} F^i\left(Z_{Xi}\right)\right).$$

Let $V^j\left(p, Z_{X1}, Z_{X2}, T_j\right) \equiv U^j\left(x_j\left(p, Z_{X1}, Z_{X2}, T_j\right), z_j\left(p, Z_{X1}, Z_{X2}, T_j\right), \sum_{i=1}^{2} F^i\left(Z_{Xi}\right)\right)$ be the consumer's indirect utility function. We therefore obtain:[3]

$$\partial V^j / \partial p = \left(X_j^0 + X_j - n_j x_j\right)\left(U_z^j / n_j\right), \tag{2a}$$

$$\partial V^j / \partial Z_{Xj} = U_E^j F_Z^j + \left(pF_Z^j + 1\right)\left(U_z^j / n_j\right), \tag{2b}$$

$$\partial V^j / \partial Z_{X_{-j}} = U_E^j F_Z^j > 0, \tag{2c}$$

$$\partial V^j / \partial T_j = U_z^j / n_j > 0. \tag{2d}$$

The market clearing condition for the industrial good can now be written as $\sum_{j=1}^{2} n_j x_j\left(p, Z_{X1}, Z_{X2}, T_j\right) = \sum_{j=1}^{2}\left(X_j^0 + F^j\left(Z_{Xj}\right)\right)$. Regional regulators control the input quantities demanded by the regional industrial sectors. The market clearing condition for the industrial good can therefore be used to implicitly define $p\left(Z_{X1}, Z_{X2}, T_1, T_2\right)$.

We examine environmental policy making in two policy settings, one in which both regions participate in the international scheme to control global carbon dioxide emissions and another in which there is no international scheme. When both regions participate in the scheme, they do so voluntarily. Therefore, each region's utility from participation must be no less than its utility from nonparticipation (i.e., its reservation utility).[4] Since these participation constraints can be adequately written only after we compute the reservation utilities, we first consider the setting in which there is no international scheme.

4.1. The Decentralized Policy Game

While it is reasonable to assume that individual producers and consumers are price takers, it is equally reasonable to assume that the regional regulators are endowed with considerable market power. Hence, regulator j chooses non-positive $\{Z_{Xj}\}$ to maximize $V^j\left(p, Z_{X1}, Z_{X2}\right)$ s.t. $p = p\left(Z_{X1}, Z_{X2}\right)$, taking $\{Z_{X_{-j}}\}$ as given, $j = 1, 2$.[5] Assuming interior solutions, the first order conditions are

$$\frac{\partial V^j}{\partial p}\frac{\partial p}{\partial Z_{Xj}} + \frac{\partial V^j}{\partial Z_{Xj}} = 0, \; j = 1, 2 . \tag{3}$$

Given (2a), (2b) and $U_x^j / U_z^j = p$, we may rewrite equations (3) as follows:

$$\frac{U_x^j}{U_z^j} = -\left[1 + \left(X_j^0 + X_j - n_j x_j\right)\left(\frac{\partial p}{\partial Z_{Xj}}\right)\right]\left(\frac{1}{F_Z^j}\right) - n_j \frac{U_E^j}{U_z^j}, \; j = 1, 2 , \tag{4}$$

where

$$\frac{\partial p}{\partial Z_{Xj}} = \frac{F_Z^j}{n_j \dfrac{\partial x_j}{\partial p}} - \frac{n_j \dfrac{\partial x_j}{\partial Z_{Xj}}}{n_j \dfrac{\partial x_j}{\partial p}}, \; j = 1, 2 , \tag{5}$$

follows from the regional regulators' partial differentiation of the market-clearing condition for the industrial good with respect to their respective choices of Z_{Xj}. Comparing equations (4) with equations (1e), we notice that the decentralized policy equilibrium involves two sources of distortion, an "externality" distortion and a "market power" distortion. The externality distortion comes from the fact that each regulator ignores the negative effects that his region's production of the industrial good generates in the other region, thus inducing excessive production of the global externality. The market power distortion arises because each regulator's choice influences the price of the industrial good and hence the international terms of trade. As equations (5) clearly illustrate, the market power distortion consists of two components. The first is the marginal effect on the net global supply of the industrial good originating with regulator j's market intervention. The second is the marginal effect on the global demand of the industrial good caused by such an intervention. Although the net effect is nonzero in general, it cannot be unambiguously signed.

Let V^{jD} represent the level of per capita utility obtained in region j in the decentralized policy equilibrium. In the case of two regions, V^{jD} is therefore region j's reservation utility level. Also, since region 1 is wealthier than region 2, we observe $V^{1D} > V^{2D}$.[6]

4.2. The Proportional Equity Scheme

There are three players in our proposed international scheme: two regional regulators and the GEF. Assuming they agree to participate in the scheme, the regulators are free to choose their own environmental policies. Their environmental policy choices are simultaneous and occur prior to the choice made by the GEF concerning the amount of resources that should be transferred from one region to the other. The GEF lacks political and economical powers to directly punish or reward the regional regulators for their actions. We postulate that the GEF's objective function obeys a proportional equity principle whereby the status quo proportion of regional per capita utility levels is held constant. This function is described below.

As noted above, the international scheme comes to existence only if both regions voluntarily decide to participate in it. Let V^{jS} denote the per-capita utility level obtained by region j if it participates in the international scheme. The regulators play the following two games:

Participation Game

Stage 0: Each regulator decides whether or not to participate, taking each other's decision as given. The binary strategies are "Yes (Y)" or "No (N)". Regulator j chooses Y if and only if $V^{jS} \geq V^{jD}$, $j = 1, 2$.

Proportional Equity Game

Stage 1: Taking $\{Z_{X2}\}$ as given, regulator 1 chooses non-positive $\{Z_{X1}\}$ to maximize $V^1(p, Z_{X1}, Z_{X2}, T_1)$ s.t. $p = p(Z_{X1}, Z_{X2}, T_1, T_2)$ and $T_1 = T_1(Z_{X1}, Z_{X2})$, taking the actions of regulator 2 as given. Similarly, regulator 2 chooses non-positive $\{Z_{X2}\}$ to maximize $V^2(p, Z_{X1}, Z_{X2}, T_1)$ s.t. $p = p(Z_{X1}, Z_{X2}, T_1, T_2)$ and $T_2 = T_2(Z_{X1}, Z_{X2})$, taking the actions of regulator 1 as given.

Stage 2: Having observed $\{Z_{X1}, Z_{X2}\}$, the GEF chooses $\{T_1, T_2\}$ to maximize

$$Min\left\{\frac{V^1(p, Z_{X1}, Z_{X2}, T_1)}{V^{1D}}, \frac{V^2(p, Z_{X1}, Z_{X2}, T_2)}{V^{2D}}\right\} \text{ s.t. } \sum_j T_j = 0 \text{ and}$$

$$p = p(Z_{X1}, Z_{X2}, T_1, T_2).$$

The Participation Game is a simultaneous non-cooperative game. We describe it as occurring in Stage 0 to emphasize the fact that it happens prior to the Proportional Equity Game ("PEG" hereafter). The regulators play the PEG if and only if the Nash equilibrium of the Participation Game is {Y,Y}. In such a case, we say that the proportional equity scheme is implementable. If the PEG is not played, each region receives its reservation utility level. In what follows, we demonstrate that the proportional equity scheme is implementable whenever there are two potential participant regions. With a larger set of potential participants, the proportional equity scheme may not be implementable.[7] We provide an example that illustrates this fact in Section 5.

The PEG consists of two sequential stages. The regulators are Stackelberg leaders and the GEF is a common Stackelberg follower. The leaders anticipate the responses of the GEF when they simultaneously select their most desirable environmental policies. The GEF then observes the leaders' choices and afterward makes its own choices. The equilibrium concept used for the PEG is subgame perfection.

In the second stage of the PEG, the GEF's best-response functions, $T_j\left(Z_{X1},Z_{X2}\right)$, $j=1,2$, satisfy:

$$\frac{V^1\left(p\left(Z_{X1},Z_{X2},T_1\left(Z_{X1},Z_{X2}\right),T_2\left(Z_{X1},Z_{X2}\right)\right),Z_{X1},Z_{X2},T_1\left(Z_{X1},Z_{X2}\right)\right)}{V^{1D}}$$

$$=\frac{V^2\left(p\left(Z_{X1},Z_{X2},T_1\left(Z_{X1},Z_{X2}\right),T_2\left(Z_{X1},Z_{X2}\right)\right),Z_{X1},Z_{X2},T_2\left(Z_{X1},Z_{X2}\right)\right)}{V^{2D}}, \tag{6a}$$

$$T_1\left(Z_{X1},Z_{X2}\right)+T_2\left(Z_{X1},Z_{X2}\right)=0. \tag{6b}$$

In the first stage, we obtain the following first order conditions if we assume interior solutions:

$$\frac{\partial V^j}{\partial p}\frac{\partial p}{\partial Z_{Xj}}+\frac{\partial V^j}{\partial Z_{Xj}}+\frac{\partial V^j}{\partial p}\left(\frac{\partial p}{\partial T_j}\frac{\partial T_j}{\partial Z_{Xj}}+\frac{\partial p}{\partial T_{-j}}\frac{\partial T_{-j}}{\partial Z_{Xj}}\right)+\frac{\partial V^j}{\partial T_j}\frac{\partial T_j}{\partial Z_{Xj}}=0,\ j=1,2. \tag{7}$$

Given (7), differentiation of (6a) and (6b) yields:

$$\frac{\partial V^{-j}}{\partial p}\frac{\partial p}{\partial Z_{Xj}}+\frac{\partial V^{-j}}{\partial Z_{Xj}}+\frac{\partial V^{-j}}{\partial p}\left(\frac{\partial p}{\partial T_j}\frac{\partial T_j}{\partial Z_{Xj}}+\frac{\partial p}{\partial T_{-j}}\frac{\partial T_{-j}}{\partial Z_{Xj}}\right)+\frac{\partial V^{-j}}{\partial T_{-j}}\frac{\partial T_{-j}}{\partial Z_{X_j}}=0,\ j=1,2. \tag{8a}$$

$$\sum_{i=1}^{2}\frac{\partial T_i}{\partial Z_{Xj}}=0,\quad j=1,\,2. \tag{8b}$$

For $j=1,2$, if we use (2a) – (2d), we may rewrite (7) and (8a), respectively, as follows:

$$\left(X_j^0+X_j-n_jx_j\right)\left(\frac{\partial p}{\partial Z_{Xj}}+\frac{\partial p}{\partial T_j}\frac{\partial T_j}{\partial Z_{Xj}}+\frac{\partial p}{\partial T_{-j}}\frac{\partial T_{-j}}{\partial Z_{Xj}}\right)+F_Z^j\left(\frac{U_x^j}{U_z^j}+n_j\frac{U_E^j}{U_z^j}\right)+1+\frac{\partial T_j}{\partial Z_{Xj}}=0, \tag{9a}$$

$$\left(X_{-j}^0+X_{-j}-n_{-j}x_{-j}\right)\left(\frac{\partial p}{\partial Z_{Xj}}+\frac{\partial p}{\partial T_j}\frac{\partial T_j}{\partial Z_{Xj}}+\frac{\partial p}{\partial T_{-j}}\frac{\partial T_{-j}}{\partial Z_{Xj}}\right)+n_{-j}\frac{U_E^{-j}}{U_z^{-j}}F_Z^{-j}+\frac{\partial T_{-j}}{\partial Z_{Xj}}=0. \tag{9b}$$

Given (8b) and the market clearing condition for the industrial good, adding up (9a) and (9b) yields

$$F_Z^j \left(\frac{U_x^j}{U_z^j} + n_j \frac{U_E^j}{U_z^j} \right) + 1 + n_{-j} \frac{U_E^{-j}}{U_z^{-j}} F_Z^{-j} = 0, \; j = 1, 2 . \tag{10}$$

Since $U_x^1 / U_z^1 = U_x^2 / U_z^2$, equations (10) imply that $F_Z^1 = F_Z^2$. Given this, we may rewrite (10) as follows:

$$\frac{U_x^j}{U_z^j} = -\frac{1}{F_Z^j} - \sum_{i=1}^{2} n_i \frac{U_E^i}{U_z^i}, \; j = 1, 2 . \tag{11}$$

Equations (11) inform us that the regional regulators behave efficiently since their regulations fully account for all external effects and do not distort the international terms of trade. The GEF's income transfer functions are powerful enough to nullify the incentives of both regulators of behaving inefficiently. The transfer functions induce both regulators to face the "correct" price for the industrial good.

Given equations (6a), (6b), and (11), it is straightforward to show that the resulting equilibrium for the global economy is Pareto efficient. Besides (6a), (6b), and (11), the global equilibrium allocation satisfies the market clearing condition for the industrial good and the marginal conditions that characterize the behavior of consumers.[8] Hence, we obtain:

Theorem 1: Assume that both regions participate in the international scheme based on proportional equity. Then, the resulting equilibrium for the global economy is Pareto efficient.

Theorem 1 informs us that an ideal allocation mechanism – one in which a central governing authority enacts interregional transfers based on a pre-determined proportional-equity principle after observing the unilateral input choices of the regions – can indeed induce an efficient allocation of resources. Turning now to the Participation Game played in Stage 0, we show that its Nash equilibrium is {Y,Y}; namely, it implies that the proportional equity scheme is implementable.

Theorem 2: The proportional equity scheme is implementable, since the subgame perfect equilibrium for the PEG yields $V^{jS} > V^{jD}$, $j = 1, 2$.

Proof. Assume that the Nash equilibrium for Stage 0 is {Y,Y}. Both regions play the PEG and the resulting outcome is such that $\sum_{j=1}^{2} V^{jS} > \sum_{j=1}^{2} V^{jD}$ because the equilibrium for the global economy is Pareto efficient. Thus, $V^{1D}\left(V^{1S}/V^{1D}\right) + V^{2D}\left(V^{2S}/V^{2D}\right) > V^{1D} + V^{2D}$. Since $\left(V^{1S}/V^{1D}\right) = \left(V^{2S}/V^{2D}\right)$ from (6a), we have $\left(V^{1D} + V^{2D}\right)\left(V^{1S}/V^{1D}\right) > V^{1D} + V^{2D}$ or $V^{1S} > V^{1D}$. A similar reasoning proves that $V^{2S} > V^{2D}$. Now, we show that indeed the Nash equilibrium for Stage 0 is {Y,Y}. Suppose region 1 chooses Y. Region 2's best response to this choice is to choose Y as well. Region 1's best response to region 2's choice is to maintain its Y-choice, yielding a Nash equilibrium {Y,Y}. Now, suppose that region 1 chooses N. As region 2's choice of Y or N yields the same payoff for this region if region 1 chooses N, it should choose Y in response to region 1's choice. Now, region 1's best response to region 2's Y-choice is to choose Y. Furthermore, region 2's best response to region 1's Y-choice is to maintain its Y-choice. Hence, the Nash equilibrium is again {Y,Y}. ■

Theorem 2 tells us that each region is unilaterally better off by participating in the proportional equity scheme.[9] Therefore, Theorems 1 and 2 together indicate that since the equilibrium for the PEG satisfies both efficiency and individual rationality, the proposed international scheme yields a "win-win" situation (i.e., both an efficient and implementable solution to the global externality problem). Intuitively, the "win-win" scenario emerges because the proportional equity principle aligns the incentives of both regions. The GEF's optimal strategy implies that the welfare of the poorer region rises if and only if the welfare of the wealthier region rises. Knowing this, the regional regulators make efficient choices. In fact, it is easy to show that this perfect incentive equivalence phenomenon occurs whenever the GEF's preferences can be represented by a function $Min\{w^1V^1, w^2V^2\}$, where w^1 and w^2 are positive weights.

As mentioned previously, the efficiency and implementability results for the proportional equity scheme are particularly relevant to the rotten-kids literature, where Bergstrom (1989), Cornes and Silva (1999), and Chiappori and Werning (2002) have derived efficient allocations for games in which the rotten kids possess quasi-linear preferences or identical preferences over a pure public good. Because the proportional equity scheme is not restrictive with respect to the preferences displayed by the rotten kids (i.e. the regional governments), our Theorems 1 and 2 therefore expand the set of circumstances under which the Rotten Kid Theorem applies. The restriction here is for the benevolent parent's preferences, since they must satisfy a particular equity principle.

To shed some additional light on these efficiency and individual-rationality results, Figure 1 depicts Theorems 1 and 2 and compares the proportional equity scheme with an alternative transfer scheme, horizontal equity (i.e., egalitarianism). Point A depicts the status quo welfare levels for regions 1 and 2, i.e., the welfare levels V^{jD}, $j=1,2$ that are obtained in the decentralized policy game. Note that because point A is on a steeper ray from the origin than the 45° line, $V^{1D} > V^{2D}$. Theorem 1 and the proportional equity constraint (6a) imply that, by choosing to participate in Stage 0, the regions move along the ray from point A to point B. Point B is located on the Pareto frontier and, based on its position relative to point A, it is consistent with Theorem 2.

If instead of participating in the proportional equity scheme the regions had agreed to participate in a scheme that obeys the horizontal equity principle, their move would have been from point A to point C.[10] Because point C corresponds to $V^{1D} > V^{1S}$ (i.e., it lies outside the shaded triangular area with vortex A), it would not be individually rational for region 1 to participate in the horizontal equity scheme. Thus, although the horizontal equity scheme results in a Pareto-efficient solution, this solution (for this particular example) is not implementable.[11] Indeed, there are several conceivable transfer schemes such as the horizontal equity scheme that would result in a Pareto-efficient but non-implementable allocation. The beauty of the proportional equity scheme is that it is both Pareto efficient and implementable for *any* initial, status quo allocation.

[INSERT FIGURE 1 HERE]

We conclude this section by briefly considering whether a fully participatory proportional equity scheme would remain implementable as the number of potential participants increases from two. Considering more than just two regions enables us to examine the effects that non-participation in the proportional equity scheme by one region has on the payoffs of the remaining regions. Consider, for example, a setting in which there are three regions that may decide to participate in the proportional equity scheme, namely, regions 1, 2 and 3. There are eight possible equilibrium outcomes in Stage 0: (i) no region decides to participate – outcome {N,N,N}; (ii) all regions decide to participate – outcome {Y,Y,Y}; (iii) regions 1 and 2 choose in favor and region 3 chooses against participation – outcome {Y,Y,N}; (iv) regions 1 and 3 choose in favor and region 2 chooses against participation – outcome {Y,N,Y}; (v) regions 2 and 3 choose in favor and region 1 chooses against participation – outcome {N,Y,Y}; (vi) region 1 chooses in favor and regions 2 and 3 choose against participation – outcome {Y,N,N}; (vii) region 2 chooses in favor and regions 1 and 3 choose against participation – outcome {N,Y,N}; and (viii) region 3 chooses in favor and regions 1 and 2 choose against participation – outcome {N,N,Y}. Clearly, outcomes {N,N,N}, {Y,N,N}, {N,Y,N} and {N,N,Y} yield identical payoffs – i.e., those associated with the status quo – since they imply that no proportional equity scheme is implementable. Partially participatory schemes are implied by outcomes {Y,Y,N}, {Y,N,Y} and {N,Y,Y}.

Let us suppose that regions 2 and 3 choose Y. Then, the outcome {Y,Y,Y} will be a Nash equilibrium in the participation game if and only if region 1's best reply to these choices is also to choose Y and each of regions 2 and 3 best reply to the other two regions choosing Y is also to choose Y. Hence, no region has an incentive to unilaterally deviate from its participation choice. It is possible, however, that the best reply of a region to the other two regions Y-choice is N. In such a case, there will not be a fully participatory Nash equilibrium. The participation game may be characterized by a unique Nash equilibrium, multiple Nash equilibria or even by the non-existence of Nash equilibrium in pure strategies. In sum, as the number of potential participants expands from two, the participation game becomes very complex and there is no certainty that a fully participatory equilibrium will be implementable. [12]

The potential for non-implementability is represented by point D in Figure 1. To see this, we can think of Figure 1 as displaying a Pareto frontier derived for any two of *n* possible regions in the global economy that are deciding whether to participate in the PEG, say regions 1 and 2. Suppose that, for whatever reason, one or more of the remaining *n-2* regions has chosen N – i.e., against participation – in Stage 0. As is common with global environmental agreements, there may exist a prisoner's dilemma for regions 1 and 2, implying that their participation in the proportional equity scheme leads to a reduction in their regional welfares – to a point such as D in Figure 1 – relative to the status quo of complete non-participation at point A. If enough regions similarly perceive themselves as "victims" of a prisoner's dilemma, then they too will choose N in Stage 0, potentially resulting in the status quo of non-implementability.[13]

5. A Three-Region Example of the PEG

Let us now turn to a simple numerical analysis to illustrate that (i) there are indeed circumstances under which the fully participatory proportional equity scheme is implementable with more than two potential participants, but that (ii) non-implementability may nevertheless occur under different circumstances.

Consider three regions – regions 1, 2, and 3 – and two scenarios. In the first scenario, the distribution of initial endowments across the regions is assumed to be relatively "tight", i.e.,

$\left(X_1^0 + Z_1^0\right) > \left(X_2^0 + Z_2^0\right) > \left(X_3^0 + Z_3^0\right)$ but the magnitudes of these differences are relatively small. In the second scenario, the magnitude of the difference between $\left(X_2^0 + Z_2^0\right)$ and $\left(X_3^0 + Z_3^0\right)$ remains small, however the magnitude of the difference between $\left(X_1^0 + Z_1^0\right)$ and $\left(X_2^0 + Z_2^0\right)$ is "loose", i.e., large. The second scenario is included for two reasons. First, it is perhaps a more accurate reflection of the difficulties faced by signatories to international environmental agreements such as the Kyoto Protocol, where one large nation – the USA – has a priori decided against participation. Second, it allows us to explore the relationship between relative endowment (or wealth) levels and the incentives of a relatively large nation to decide not to participate in the international scheme.

To focus on the issue of potential non-participation, we assume a very simplistic model. Each region has one consumer and one industrial firm, and preferences and technology are identical across regions. The firm's production function is $X_j = \alpha \ln\left(-Z_{Xj}\right)$ and the consumer's utility function is $U^j = \beta \ln\left(x_j\right) + \ln\left(z_j\right) - \gamma E^2$, $\alpha > 0$, $\beta > 0$, $\gamma > 0$. Table 1 provides the specific parameter values adopted for this example. Note that Scenario 1(2) refers to the tight (loose) distribution of initial endowments across regions. Specifically, the distribution of aggregate initial endowments of X and Z_X in Scenario 1 is 100, 90, and 80 for regions 1 through 3, respectively, and in Scenario 2 the distribution is 125, 90, and 80.

[INSERT TABLE 1 HERE]

Three simulations are run – one depicting the status quo decentralized policy game, one depicting the fully participatory PEG, and one depicting the case where regions 2 and 3 participate in the PEG but region 1 chooses not to participate.[14] Equations (2a) – (2d) hold for each simulation. Equations (3) – (5) hold for the decentralized policy game simulation, as well as for region 1 when it does not participate. Equations (6a), (6b), and (11) hold for the fully participatory PEG simulation, as well as for regions 2 and 3 when region 1 does not participate. The simulations are performed using GAMS, version 2.0.13.

Table 2 provides the simulation results for the regional welfare levels obtained under the different scenarios and policy regimes. First, note that in both Scenarios 1 and 2 the fully participatory proportional equity scheme improves welfare for all regions relative to the welfare levels obtained in the decentralized policy game, while maintaining the status quo proportion of aggregate welfare for each region. For example, the status quo proportions of aggregate welfare under Scenario 1 – 0.342, 0.334, and 0.325 for regions 1, 2, and 3, respectively – are the same under the fully participatory PEG. However, in the situation where region 1 does not participate in the PEG, the outcomes across the two scenarios diverge.

[INSERT TABLE 2 HERE]

In Scenario 1, where the distribution of initial endowments across regions is tight, region 1's welfare is smaller when it does not participate in the PEG than when it does. Thus, region 1 has no incentive to decide against participating in the PEG and therefore no incentive to free-ride on regions 2 and 3. However, in Scenario 2, where the distribution of initial endowments is loose, region 1's welfare increases with non-participation. So, let us say that in Stage 0 it chooses N. In this case, the best responses of regions 2 and 3 to region 1's N-choice would also be to choose N. Since region 1 does not gain anything from changing its N-choice in response to the N-choices made by the other two regions, the Nash equilibrium is {N,N,N}.

In comparing the two scenarios, we therefore find evidence to suggest that a wealthier nation's incentive to decide against participating in the PEG and to free-ride increases in the relative size of its initial endowment. One possible explanation for this result is that by joining a scheme that includes budget-balanced interregional transfers, a wealthy nation can expect the size of its transfer payment (to the GEF, which in turn is transferred to the poorer nations) to be positively related to the relative size of its initial endowment.[15] Thus, a critical initial-endowment differential may exist, beyond which the gain to the wealthy nation from deciding not to participate in the PEG outweighs the corresponding loss, which results in an overall increase in the level of the global externality.[16]

However, such a motivation to not join the proportional equity scheme may be overcome if the wealthy nation is able to commit to its decision of whether or not to participate before the other nations make their own choices concerning participation. By anticipating that the other nations may follow its lead, the wealthy nation may choose to participate in the PEG. For example, suppose now that the participation decisions are made sequentially, with region 1 being the leader and regions 2 and 3 being followers. Suppose, in addition, that regions 2 and 3 decide whether or not to participate simultaneously. The participation game, therefore, consists of two stages. The equilibrium concept used for this game is subgame perfection. Consider the second stage of the game. Suppose that region 1 chose N, against participation, in the first stage. Then, regions 2 and 3 choose N in the second stage. The resulting outcome is thus {N,N,N}. Now, suppose instead that region 1 chose Y in the first stage. Having observed this choice, regions 2 and 3 choose Y in the second stage. The resulting outcome is {Y,Y,Y}. Anticipating how regions 2 and 3 will behave in the second stage, region 1 will surely choose Y in the first stage. Hence, the subgame perfect equilibrium for the participation game is {Y,Y,Y}. This example illustrates that sequential decision-making in the participation stage may improve the chances of the fully participatory PEG becoming implementable.[17]

6. Conclusion

Our main results suggest that regions may behave efficiently in the presence of an international scheme in which resource transfers from wealthy to poor regions are implemented to satisfy a proportional equity principle. Such transfers promote perfect incentive equivalence (i.e., satisfy the Rotten Kid Theorem) since regional welfare levels become complements: the welfare of the wealthy region can rise if and only if the welfare of the poor region rises. In addition to perfect incentive equivalence, our proposed international scheme induces both wealthy and poor regions to participate; that is, it yields a potential "win-win" scenario. The efficiency and individual rationality properties of such international scheme are especially noteworthy in light of the call for international transfers in global agreements such as the Kyoto Protocol.

We claim that our results are robust to regional policy decentralization provided that the structure of our proposed international scheme remains intact. It is straightforward to show that, as long as the GEF implements interregional transfers to satisfy the proportional equity principle *after* the regional regulators choose their environmental policies, the results of our analysis would remain unchanged in a setting where either the regional regulators control Pigouvian taxes or initial endowments of regional quantities of carbon dioxide permits for trade among the regions' agents in a global permit market. The reason is that the complementarity of regional welfare levels under the proportional equity principle, which provides the rationale for our results, would still be present in these alternative and more complex economic settings.

Notes

*Corresponding author.

[1] As we show in Section 4, the proportional equity scheme is certainly not the only transfer scheme that can result in a Pareto-efficient resource allocation. However, because it is both efficient *and* implementable, the proportional equity scheme merits further investigation. As we demonstrate, the scheme is always implementable in the case of two regions and is more likely to be implementable with greater than two regions when the regions make their participation decisions sequentially as opposed to simultaneously.

[2] These papers all assume that interregional transfers act as a mechanism to align the incentives of regions that have pre-committed to participate in the federation. For an example of how transfers and commitments can be used to enlarge the number of regions participating in the federation see Carraro and Siniscalco (1993).

[3] Derivations of the results in equations (2a) – (2d) are available upon request from the authors.

[4] In a game with more than two regions, a given region's reservation utility level is the maximum utility level obtainable from any possible sub-coalition to which the given region could possibly belong.

[5] $V^j\left(p, Z_{X1}, Z_{X2}\right) \equiv \left\{V^j\left(p, Z_{X1,} Z_{X2}, T_1, T_2\right) \middle| T_1 = T_2 = 0\right\}$
and $p\left(Z_{X1}, Z_{X2}\right) \equiv \left\{p\left(Z_{X1}, Z_{X2}, T_1, T_2\right) \middle| T_1 = T_2 = 0\right\}$.

[6] This assumes identical preferences and technologies across regions.

[7] Although beyond the scope of this study, the issue of sub-coalition formation naturally arises in the cooperative game theory literature. See, for example, Barrett (1994), Carraro and Siniscalco (1993), Chander and Tulkens (1997), Hoel and Schneider (1997), Jeppesen and Anderson (1998), Finus (2004), and Finus and Rundshagen (2005) on the design of cooperative international environmental agreements.

[8] The modified Samuelson conditions could also be obtained by differentiating the proportional equity constraint with respect to each of the choice variables controlled by the regional governments while accounting for the overall resource constraint. We thank an anonymous reviewer for pointing this out. Although this alternative way of deriving the efficiency conditions highlights the importance of the proportional equity constraint, it does not immediately conform to the formulation of PEG as described above.

[9] With just two regions, if one region chooses (for whatever reason) not to participate in Stage 0 the PEG trivially collapses.

[10] In an HEG, the equity constraint (6a) becomes

$$V^1\left(p\left(Z_{X1},Z_{X2},T_1\left(Z_{X1},Z_{X2}\right),T_2\left(Z_{X1},Z_{X2}\right)\right),Z_{X1},Z_{X2},T_1\left(Z_{X1},Z_{X2}\right)\right)=$$

$$V^2\left(p\left(Z_{X1},Z_{X2},T_1\left(Z_{X1},Z_{X2}\right),T_2\left(Z_{X1},Z_{X2}\right)\right),Z_{X1},Z_{X2},T_2\left(Z_{X1},Z_{X2}\right)\right).$$ Thus, the solution to the HEG is likewise Pareto efficient.

[11] Interestingly, the HEG is implied by a particular case of the quasi-linear preferences that Bergstrom (1989) proposes as being sufficient for the Rotten Kid Theorem to hold.

[12]Barrett (1994) and Carraro and Siniscalco (1993) discuss more formally the possibilities of multiple equilibria (i.e., various configurations of sub-coalitions) that might result for environmental agreements with larger numbers of potential participants.

[13] Interestingly enough, there may be circumstances under which the sequentially withdrawing regions might instead "bribe" the initial non-participating region(s), or lead region(s), to participate and thus obtain an efficient allocation. In this case, however, the proportional equity principle would be violated. Although potentially useful for increasing participation in the PEG, bribing non-participants due to a prisoner's dilemma represents a limitation of our analysis as we extend the model to accommodate more than two regions. To the degree that the regions are forward-looking, the potential for bribes would likely be endogenized by both the participating and non-participating regions, therefore eliminating the possibility of ex post side-payments

altogether. This is an interesting research question that is beyond the scope of our paper. It is also possible that by having the option in a sequential participation game stage to preempt the other regions with its participation decision, a leader region may decide not to withdraw from the PEG in the first place. This would be an expected outcome if the leader nation (i) predicts that the remaining regions would reciprocate with non-participation given its decision not to participate, and (ii) determines that it is better off with a fully participatory PEG than with the non-cooperative outcome. To the contrary, with simultaneous rather than sequential decision-making in the participation stage a prisoner's dilemma is more likely to ensue, resulting in the non-cooperative outcome. We explore these issues further in Section 5. Finus and Rundshagen (2005) provide a more formal treatment. See also Sandler (1997) for a more in-depth discussion of the prisoner's dilemma problem in international environmental agreements.

[14] The GAMS programs for each simulation are available upon request from the authors. We have not performed simulations for the situations where regions 2 and 3, respectively, choose not to participate in order to focus on the key issue currently facing signatories to international agreements such as the Kyoto Protocol, namely the large region's decision not to participate in the agreement.

[15] This will always be the case for a purely public externality when the wealthy nation is identical to the poorer nations in every respect except the size of its initial endowment.

[16] The gain to the wealthy nation materializes via (i) reduced contributions to the global externality by the other nations to compensate for its increased contribution and (ii) not having to make a transfer payment to the GEF.

[17] We thank an anonymous reviewer for raising some interesting points about the potential for non-implementability that led us to consider this possibility. Discussing the multitude of other possibilities is obviously beyond the scope of this paper.

References

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

Becker, G.S. (1982). "A Theory of Social Interactions." Journal of Political Economy 82, 1063-1094.

Bergstrom, T.C. (1989). "A Fresh Look at the Rotten-Kid Theorem – and Other Household Mysteries." Journal of Political Economy 97(5), 1138-1159.

Boadway, R. (1982). "On the Method of Taxation and the Provision of Local Public Goods: Comment." American Economic Review 72, 846-851.

Caplan, A.J. and E.C.D. Silva (1999). "Federal Acid Rain Games." Journal of Urban Economics 46, 25-52.

Caplan, A.J., R. C. Cornes, and E.C.D. Silva (2003). "An Ideal Kyoto Protocol: Emissions Trading, Redistributive Transfers and Global Participation." Oxford Economic Papers 55, 216-234.

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

Chander, P. and H. Tulkens (1997). "The Core of an Economy with Multilateral Environmental Externalities." International Journal of Game Theory 26(3), 379-401.

Chiappori, P.A. and I. Werning (2002). "Comment on Rotten Kids, Purity, and Perfection." Journal of Political Economy 110(2), 475-480.

Chichilnisky, G., G. Heal and D. Starrett (2000). "Equity and Efficiency in Environmental Markets: Global Trade in Carbon Dioxide Emissions." In G. Chilchinisky and G. Heal (eds.), Environmental Markets: Equity and Efficiency. New York: Columbia University Press.

Cornes, R.C. and E.C.D. Silva (1999). "Rotten Kids, Purity, and Perfection." Journal of Political Economy 107, 1034-1040.

Cornes, R.C. and T. Sandler (1996). The Theory of Externalities, Public Goods, and Club Goods. Cambridge: Cambridge University Press.

Finus, M. (2004). "Modesty Pays: Sometimes!" Working Paper 68.04, Fondazione Eni Enrico Mattei (FEEM).

Finus, M. and B. Rundshagen (2005). "Participation in International Environmental Agreements: The Role of Timing and Regulation." Working Paper 45.05, Fondazione Eni Enrico Mattei (FEEM).

Hoel, M. and K. Schneider (1997). "Incentives to Participate in an International Environmental Agreement." Environmental and Resource Economics 9 (2), 153-170.

Jeppesen, T. and P. Andersen (1998). "Commitment and Fairness in Environmental Games', in N. Hanley and H. Folmer (eds.), Game Theory and the Environment. Cheltenham: Edward Elgar Publishers.

Myers, G.M. (1990). "Optimality, Free Mobility, and the Regional Authority in a Federation." Journal of Public Economics 43, 107-121.

Nagase, Y. and E.C.D. Silva (2000). "Optimal Control of Acid Rain in a Federation with Decentralized Leadership And Information." Journal of Environmental Economics and Management 40, 164-180.

Sandler, T. (1997). Global Challenges: An Approach to Environmental, Political, and Economic Problems. Cambridge: Cambridge University Press.

Silva, E.C.D. (1997). "Decentralized and Efficient Control of Transboundary Pollution in Federal Systems." Journal of Environmental Economics and Management 32, 95-108.

Silva, E.C.D. and A.J. Caplan (1997). "Transboundary Pollution Control in Federal Systems." Journal of Environmental Economics and Management 34, 173-186.

Wellisch, D. (1994). "Interregional Spillovers in the Presence of Perfect and Imperfect Household Mobility." Journal of Public Economics 55, 167-184.

Table 1. Parameter Values for a Numerical Example Containing Three Regions.

Parameter	Value
↦	0.2
↕	2
•	0.1
$\left(X_1^0, Z_{X1}^0\right)_1, \left(X_1^0, Z_{X1}^0\right)_2$	(100,100), (125,125)
$\left(X_2^0, Z_{X2}^0\right)_1, \left(X_2^0, Z_{X2}^0\right)_2$	(90,90), (90,90)
$\left(X_3^0, Z_{X3}^0\right)_1, \left(X_3^0, Z_{X3}^0\right)_2$	(80,80), (80,80)

Table 2. Regional Welfares for Scenarios 1 and 2.

Region	Status Quo	PEG	Region 1 Does Not Participate
		Scenario 1	
1	13.373	13.398	13.396
2	13.057	13.082	13.038
3	12.705	12.729	12.686
		Scenario 2	
1	14.046	14.068	14.072
2	13.062	13.083	13.051
3	12.709	12.730	12.699

Figure 1. The Proportional Equity Transfer Scheme