u^b

b
**UNIVERSITÄT
BERN**

Faculty of Business, Economics
and Social Sciences

Department of Economics

Sustainable Climate Treaties

Hans Gersbach
Noemi Hummel
Ralph Winkler

11-05

June 2011

DISCUSSION PAPERS

**Schanzeneckstrasse 1
Postfach 8573
CH-3001 Bern, Switzerland**
http://www.vwi.unibe.ch

Sustainable Climate Treaties [*]

Hans Gersbach

CER-ETH – Center of Economic
Research at ETH Zurich and CEPR
Zürichbergstrasse 18
8092 Zurich, Switzerland
hgersbach@ethz.ch

Noemi Hummel

CER-ETH – Center of Economic
Research at ETH Zurich
Zürichbergstrasse 18
8092 Zurich, Switzerland
nhummel@ethz.ch

Ralph Winkler

Department of Economics and Oeschger Centre
for Climate Change Research, University of Bern
Schanzeneckstrasse 1
3001 Bern, Switzerland
rwinkler@vwi.unibe.ch

First Version: July 2009

This Version: June 2011

Abstract: We examine a global refunding scheme for mitigating climate change. Countries pay an initial fee into a global fund that is invested in long-run assets. In each period, part of the fund is distributed among the participating countries in relation to the emission reductions they have achieved in this period. We identify two possible types of sustainable treaty. A first-best sustainable treaty involves varying amounts of refunded wealth and a minimal amount of initial fees inducing socially desirable abatement efforts in each period. In a second-best sustainable treaty with only two parameters – optimally selected initial fees and constant refunds equal to the interest earned on the fund – the stock of greenhouse gases converges to the socially optimal stock. Finally, we suggest ways for countries to raise money for the payment of initial fees that are neutral to tax payers and international capital markets.

Keywords: climate change mitigation, refunding scheme, international agreements, sustainable treaty

JEL Classification: Q54, H23, H41

1 Introduction

International treaties on the provision of global public goods are plagued by the fundamental free-riding problem: each country's contribution will benefit all countries in a non-exclusive and non-rival manner. This prisoner's dilemma aspect and the absence of a supranational authority makes international coordination both crucial and exceptionally difficult to achieve. Countries may either lack the incentive to sign an agreement and benefit from the signatories' contributions or they may have incentives not to comply with promises made in an agreement.

In long-run problems extending over decades or even centuries, such as mitigating anthropogenic climate change, a second problem arises. Even if the free-riding problem has been solved, little is achieved if the international community fails to agree on a subsequent agreement when the first has expired. With respect to anthropogenic climate change, this is precisely the problem we face today. Although the end of the Kyoto Protocol is nigh,[1] the international community failed to agree on a subsequent international agreement to reduce greenhouse gas emissions both in December 2009 in Copenhagen and a year later in Cancún.

In this paper we propose and analyze a treaty design involving refunding, which we call a refunding scheme (henceforth RS). The main idea of the RS is that all countries pay an initial fee into a global fund that is invested in long-run assets. Countries maintain full sovereignty over how much emissions they abate each year and what policy measures they use to do so. At the end of each year, part of the fund is paid out to countries in proportion to the relative GHG emission reductions they have achieved in that year. We show that a suitably selected RS establishes a sustainable solution to the free-rider problem, so the fund will never be exhausted. A sustainable solution can be achieved in two ways. Emissions can be at the socially optimal level in each period (first-best sustainable solution) or emissions converge to their socially optimal level in the long run (second-best sustainable solution). In the latter case, the refund is equal to the interest earned on the fund. Both these sustainable solutions share one property: once the refunding scheme is established, no further coordination is required, except in administering the system, measuring reductions, and investing and distributing money. By construction, both schemes will last forever, as the fund will never be exhausted.

Our main formal results are as follows: First, the globally optimal solution minimizing the discounted values of global abatement costs and global damages prescribes uniquely

---

[1] In the Kyoto Protocol – the first international treaty on reducing greenhouse gas emissions with binding emission targets – the industrialized countries of the world, so called Annex B countries, committed themselves to a reduction of greenhouse gas (GHG) emissions by 5.2% against 1990 levels over the period from 2008 to 2012.

determined emission abatement efforts for each period and each nation such that the global stock of GHGs converges to a steady state called long-run desired stock. If there is no treaty, countries will choose levels of GHG that fall considerably short of the globally optimal solution, while the stock of GHGs converges to a steady state well above the long-run desired stock.

Second, we show that initial fees and a feasible sequence of refunds can be devised in such a way that the RS implements socially optimal abatement levels in each period. We call this treaty the first-best sustainable RS. We also explore the potential of a particularly simple RS. In each period, the interest yields of the initial fees invested at a constant interest rate are refunded to the countries. We determine the amount of initial fees that induces convergence to the long-run desired stock. We call such a scheme second-best sustainable RS, as although countries do not choose socially optimal abatement levels in all periods, in the long run the abatement levels and also the stock of GHGs will converge to the social optimum. Both treaties provide a sustainable solution for the climate change problem. The main intuition is as follows: Nations are free to choose low or even zero abatement levels in one or more periods, but then they will forfeit refunds for that period. If the initial fees are sufficiently large, countries will choose high abatement levels, thereby benefiting from correspondingly large refunds. In the first-best sustainable RS, refunds can be adjusted so that countries choose socially optimal abatement levels in each period, whereas in the second-best sustainable RS, initial fees are chosen in such a way that the countries' emission abatement levels will converge to the socially optimal levels in the long run.

Third, as the initial amount of money the countries have to pay may be quite large, especially in the case of the second-best sustainable RS, we show that the same solution can be obtained if countries periodically pay a fixed amount into the fund, which will be smaller, the shorter the duration is between two payments made by a nation. Additionally, we suggest different ways of financing the initial fees that are neutral to tax payers and international capital markets.

Fourth, as countries may want to renounce paying the initial fee and not sign the treaty, initial participation requires that countries be pivotal for the formation of the RS. That is, if any country defects, the treaty will fail.[2] Then initial participation will be part of a subgame perfect equilibrium.

---

[2]In practice, of course, only industrial countries will be called upon to set-up the climate fund. But even the participation of industrial countries remains a thorny issue, as we will discuss further in Section 6.

The starting point for our scheme and its analysis is the large body of game-theoretic literature on the formation of international and self-enforcing environmental agreements[3] as there is no supranational authority to ensure participation and compliance. This literature has provided valuable insights into the potentialities and limitations of international environmental agreements. The literature also suggests that a large coalition will achieve only modest abatement efforts or will fail to enter into force (Asheim et al. 2006). Our approach complements this literature by suggesting a procedure that enables a coalition to achieve its emission reduction objectives after the coalition has been formed. We suggest that even large coalitions can achieve substantial emission reductions if they are able to set up an agency that has the power to administer a refunding scheme on which coalition members have previously agreed.[4]

The paper is organized as follows: In the next section, we set up our model, for which in Section 3 we derive the social optimum and the decentralized solution as benchmark cases. The refunding scheme is introduced in Section 4, where the existence of the first- and second-best sustainable RS is also established. In Section 5 we illustrate our model numerically. In Sections 6 and 7 we discuss practical aspects of the RS, such as initial participation and how to raise initial fees. Section 8 concludes.

2 The Model

We consider a world with $n \geq 2$ identical countries characterized by an emission function E, an abatement cost function C, and a damage function D over a finite time horizon T. As we consider T to be arbitrarily large, we shall also investigate the limit $T \to \infty$. Throughout the paper countries are indexed by i and j, and time is indexed by t.

Emissions of country i in period t are assumed to equal "business-as-usual" emissions ϵ (i.e., emissions arising if no abatement effort is undertaken) minus emission abatement a_t^i:

$$E(a_t^i) = \epsilon - a_t^i , \quad \text{with} \quad a_t^i \in [0, \epsilon] , \quad i = 1, \ldots, n , \quad t = 1, \ldots, T . \tag{1}$$

---

[3]Non-cooperative and cooperative approaches have been pursued. Notable contributions are Barrett (1994), Barrett (1999), Barrett (2003), Carraro and Siniscalco (1992), Carraro and Siniscalco (1993), Chander and Tulkens (1992), Hoel (1992), and Tulkens (1979). The earlier literature is comprehensively summarized by Finus (2001). Pioneering in the modelling of coalition structures are Bloch (1997) and Yi (1997), further developed by Finus and Rundshagen (2009).

[4]Taking up a suggestion by Gersbach (2005), Gersbach and Winkler (2007) focus on refunding schemes in a two-period setting in which participating countries pay emission taxes into a global fund. In this paper, we examine a refunding scheme in which countries only pay initial fees and focus on the long-run properties of such a scheme.

We assume that emission abatement a_t^i is achieved by enacting some national environmental policy, which induces convex abatement costs in country i:[5]

$$C(a_t^i) = \frac{\alpha}{2}\left(a_t^i\right)^2 \;, \quad \text{with} \quad \alpha > 0 \;, \quad i = 1,\ldots,n \;, \quad t = 1,\ldots,T \;. \tag{2}$$

Global emissions, which are the sum of the emissions by all countries, accumulate the stock of greenhouse gases, s_t, according to the following equation of motion:

$$s_{t+1} = (1-\gamma)s_t + \sum_{i=1}^{n} E(a_t^i) \;, \quad \text{with} \quad 0 < \gamma < 1 \;, \quad t = 1,\ldots,T \;, \tag{3}$$

where γ denotes the constant and positive natural decay rate of greenhouse gases in the atmosphere. The initial stock of greenhouse gases is denoted by s_1.

The global stock of greenhouse gases in period t, s_t, gives rise to strictly increasing and strictly convex damage for each country i:

$$D(s_t) = \frac{\beta}{2}s_t^2 \;, \quad \text{with} \quad \beta > 0 \;, \quad t = 1,\ldots,T \;. \tag{4}$$

Finally, countries are assumed to discount outcomes in period t with the discount factor δ^{t-1} with $0 < \delta < 1$.

3 Social Optimum and Decentralized Equilibrium

Before we introduce the refunding scheme (RS) in the next section, we characterize the global social optimum and the decentralized solution when no international agreement has been reached. As is well-known, the latter is inefficient because the emissions of each individual country impose negative externalities on all other countries that an individual country does not take into account when choosing the extend of its emission abatement.

Both the global social optimum and the decentralized outcome are important benchmarks in evaluating the performance of potential international agreements. While the decentralized outcome is realized if no agreement takes place, the social optimum is the ultimate goal an international agreement seeks to implement. Obviously, any agreement has to outperform the decentralized outcome in order to be seriously considered, and it is the "better," the closer its outcome is to the global social optimum.

---

[5]This is a standard short-cut way of capturing aggregate abatement costs in country i (see, e.g., Falk and Mendelsohn 1993).

3.1 Global social optimum

Consider a global social planner seeking to maximize *global* welfare, i.e., seeking to minimize the net present value of *total global costs* consisting of global costs of emission abatement and the sum of national environmental damages stemming from the pollution stock.

To solve the social planner's problem, we introduce the following recursive value function in period t:

$$V_t(s_t) = \max_{\{a_t^i\}_{i=1}^n} \left\{ \delta V_{t+1}(s_{t+1}) - \sum_{j=1}^n \frac{\alpha}{2} \left(a_t^j\right)^2 - \frac{n\beta}{2} s_t^2 \right\} , \tag{5}$$

where $V_t(s_t)$ represents the negative of the total global costs accruing from period t onwards discounted to period t. The social global planner's problem is to maximize (5) for $t = 1, \ldots, T$ subject to equation (3) and $V_{T+1}(s_{T+1}) \equiv 0$. Assuming an interior solution[6], i.e., $a_t^i \in (0, \epsilon)$ for all $i = 1, \ldots, n$ and $t = 1, \ldots, T$, the following first-order conditions are necessary for a global optimum:

$$\alpha a_t^i = -\delta V_{t+1}'(s_{t+1}) , \quad t = 1, \ldots, T . \tag{6}$$

Differentiating $V_t(s_t)$ with respect to s_t and applying the envelope theorem yields

$$-V_t'(s_t) = n\beta s_t - \delta(1-\gamma)V_{t+1}'(s_{t+1}) , \quad t = 1, \ldots, T . \tag{7}$$

Recursive evaluation of equation (7) implies that the negative of the first derivative of the value function, $-V_t'(s_t)$, equals the net present value in period t of all future damages from one additional marginal unit of the pollution stock s_t summed up over all countries. Then first-order condition (6) says that in the global social optimum the costs of abating an additional marginal unit of emissions have to equal the net present value of all mitigated future damages caused by this additional marginal unit by decreasing the pollution stock s_t. As abatement in period t only influences the damages in the period $t+1$ and the world ends after period T, abatement in the terminal period does not pay off, implying $a_T^i = 0$. According to the following proposition, there exists a unique global optimum:

Proposition 1 (Global Social Optimum)
For any time horizon $T \leq \infty$ there exists a unique social global optimum characterized by identical sequences of emission abatements for all countries i in all periods t, $a_t^{i\star} = a_t^\star$, and a sequence for the greenhouse gas stock $s_t^\star$ $(i = 1, \ldots, n; \; t = 1, \ldots, T)$.

---

[6]Interior solutions $a_t^i > 0$ are guaranteed by the quadratic abatement cost and damage functions.

The proof of Proposition 1 is given in the Appendix. The proof is constructive in the sense that we not only show the existence and the uniqueness of the social global optimum but also determine closed-form solutions for the sequences $a_t^\star$ and $s_t^\star$. As we are particularly interested in the long run, we state the following corollary:

Corollary 1 (Global Social Optimum in the Long Run)
For $T \to \infty$ the sequences of emission abatement and the greenhouse gas stock in the global social optimum converge to their steady state levels

$$a^{SO} = \frac{n^2 \beta \delta \epsilon}{\alpha \gamma \left[1 - \delta(1 - \gamma)\right] + n^2 \beta \delta} \ , \tag{8a}$$

$$s^{SO} = \frac{n \alpha \epsilon \left[1 - \delta(1 - \gamma)\right]}{\alpha \gamma \left[1 - \delta(1 - \gamma)\right] + n^2 \beta \delta} \ . \tag{8b}$$

The proof of Corollary 1 is given in the Appendix.

3.2 Decentralized solution

Next we examine a decentralized system in the absence of an international treaty, where a local planner in each country (e.g., a government) seeks to minimize the *total local costs* consisting of local abatement costs and local environmental damages. We are looking for subgame perfect Nash equilibria in pure strategies for this game.

We solve the game by backward induction, starting from period T. It is useful to consider a typical step in this procedure. To this end, suppose that there exists a unique subgame perfect equilibrium for the subgame starting in period $t+1$ with a stock of greenhouse gases s_{t+1}. For the moment, this is assumed to hold in all periods $t+1$ and will be verified in the proof of Proposition 2. Other details of the history of the game apart from the level of the greenhouse gas stock s_{t+1} do not matter, as only s_{t+1} influences the payoffs of the subgame starting in period $t+1$ and the equilibrium is assumed to be unique.

Given the unique subgame perfect equilibrium for the subgame starting in period $t+1$ with the associated equilibrium payoff $W_{t+1}^i(s_{t+1})$, country i's best response in period t, $\bar{a}_t^i$, is determined by the solution of the optimization problem

$$V_t^i(s_t)|A_t^{-i} = \max_{a_t^i} \left\{ \delta W_{t+1}^i(s_{t+1}) - \frac{\alpha}{2} \left(a_t^i\right)^2 - \frac{\beta}{2} s_t^2 \right\} \ , \tag{9}$$

subject to equation (3), $W_{T+1}^i(s_{T+1}) \equiv 0$, and given the sum of abatement efforts by all other countries $A_t^{-i} = \sum_{j \neq i} a_t^j$.

As we will show in the proof of Proposition 2, country i's optimization problem in period t is strictly concave. Thus, differentiating equation (9) with respect to a_t^i and restricting our attention to interior solutions, we obtain an implicit function for country i's best response

$$\alpha \bar{a}_t^i = -\delta W_{t+1}^{i}{}'(\bar{s}_{t+1}) \ , \tag{10}$$

where $\bar{s}_{t+1} = (1-\gamma)s_t + n\epsilon - \bar{a}_t^i - A_t^{-i}$. In addition, applying the envelope theorem to equation (9) yields

$$-V_t^{i}{}'(s_t)|A_t^{-i} = \beta s_t - \delta(1-\gamma)W_{t+1}^{i}{}'(\bar{s}_{t+1}) \ . \tag{11}$$

We observe that $W_t^i(s_t) = V_t^i(s_t)|\hat{A}_t^{-i}$ with $\hat{A}_t^{-i} = \sum_{j \neq i} \hat{a}_t^j$ denoting the sum of emission abatement in period t in the subgame perfect Nash equilibrium for all countries $j \neq i$. Recursive evaluation of equation (11) implies that the negative of the first derivative of the payoff function in period t, $-W_t^{i}{}'(s_t)$, equals the net present value of all future damages in country i from one additional marginal unit of the pollution stock s_t in period t. Then, the best-response function (10) says that country i's costs of abating an additional marginal unit of emissions have to equal the net present value of all mitigated local future damages caused by this additional marginal unit by decreasing the pollution stock s_t. Note that country i's best-response function only depends on the sum of the emission abatement of all other countries A_t^{-i} and not on the contribution of individual countries a_t^j.

The following proposition establishes the existence and uniqueness of a subgame perfect Nash equilibrium:

Proposition 2 (Decentralized Solution)
For any time horizon $T < \infty$, there exists a unique subgame perfect Nash equilibrium characterized by identical sequences of emission abatements for all countries i in all periods t, $\hat{a}_t^i = \hat{a}_t$, and a sequence for the greenhouse gas stock $\hat{s}_t$ ($i = 1, \ldots, n$; $t = 1, \ldots, T$).

The proof of Proposition 2 is given in the Appendix. In the proof we also determine closed-form solutions for the sequences $\hat{a}_t$ and $\hat{s}_t$.

Again, we are interested in the long run and take the limit for $T \to \infty$. The reason is that this equilibrium approximates the equilibrium for very large, but still finite time horizons T.[7]

---

[7]In infinite horizon models, further equilibria and even a continuum of equilibria can occur (Tsutsui and Mino 1990, Rowat 2007). However, it can be shown that the equilibrium we achieve by taking the limit $T \to \infty$ is the unique Markov perfect equilibrium in affine strategies (Lockwood 1996).

Corollary 2 (Decentralized Solution in the Long Run)

For the limit $T \to \infty$, the sequences of emission abatement and the greenhouse gas stock in the unique subgame perfect Nash equilibrium of the decentralized solution for finite time horizons T converge to the steady state levels

$$a^{DS} = \frac{n\beta\delta\epsilon}{\alpha\gamma\left[1 - \delta(1 - \gamma)\right] + n\beta\delta} \; , \tag{12a}$$

$$s^{DS} = \frac{n\alpha\epsilon\left[1 - \delta(1 - \gamma)\right]}{\alpha\gamma\left[1 - \delta(1 - \gamma)\right] + n\beta\delta} \; . \tag{12b}$$

The proof of Corollary 2 is given in the Appendix.

From Corollaries 1 and 2 we observe that $a^{SO} > a^{DS}$ and $s^{SO} < s^{DS}$, which reflects the well-known underprovision of emission abatement in the decentralized case due to the incentives for each country to free-ride on the emission abatements of all other countries. As these incentives increase with the number of countries n, the underprovision of abatement becomes more severe, the higher n is.

4 Refunding Scheme

In the following, we introduce a refunding scheme (RS) and analyze its potential for improving on the decentralized solution. The essential idea is that a global fund be established that refunds interest earnings to member countries in each period, proportionally to their relative emission reductions.

4.1 Rules of the RS

We consider a three-step procedure. First, participating countries negotiate the parameters of the RS. In particular, this includes duration T of the treaty, the level of an initial fee f_0^i payable into a global fund by each participating country, and reimbursements R_t for all periods $t = 1, \ldots, T - 1$. Second, in each period $t = 1, \ldots, T - 1$ the fraction R_t of the fund is reimbursed to the participating countries in proportion to the emission reductions they have achieved relative to overall emission abatements in this period. The remaining assets of the fund are invested at the constant interest rate ρ per period, and the returns add to the global fund in the next period $t+1$. Finally, in period T the fund is equally redistributed to the participating countries.

Thus, the fund at the end of period t reads

$$f_t = (1 + \rho)f_{t-1} - R_t , \quad t = 1, \ldots, T - 1 , \tag{13}$$

with an initial fund $f_0 = \sum_{i \in S} f_0^i$, where S denotes the set of countries participating in the RS. Note that $f_T = 0$, or equivalently $R_T = (1 + \rho)f_{T-1}$. In addition, the refund r_t^i a member country i receives in period t yields

$$r_t^i = \begin{cases} R_t \dfrac{a_t^i}{\sum_{j \in S} a_t^j} , & t = 1, \ldots, T - 1 , \\ \dfrac{R_t}{|S|} , & t = T . \end{cases} \tag{14}$$

A set of initial fees $f_0^i \geq 0$ and refunds $R_t \geq 0$ is *feasible* if $f_t \geq 0$ for all $t = 1 \ldots, T$ holds.

In order to analyze the potential of an RS to mitigate climate change, we proceed as follows: First, assuming that all countries participate in the RS in step one, we show that for any feasible set of initial fees f_0^i and refunds R_t and any time horizon $T < \infty$ there exists a unique and symmetric subgame perfect Nash equilibrium for steps two and three. Second, we show that there exists a feasible set of initial fees f_0^i and refunds R_t for which the subgame perfect Nash equilibrium resembles the global social optimum. This is called a first-best sustainable RS. Third, we show that a refunding scheme in which all returns of the fund are fully redistributed in each period converges to the social global optimum in the long run for an appropriate choice of initial fees f_0^i. This treaty is called second-best sustainable RS.

4.2 Subgame perfect equilibrium

Given that all n countries have joined the RS and given an arbitrary but finite time horizon T and a feasible set of initial fees f_0^i and refunds R_t, we now analyze the subgame perfect equilibria of the RS. We assume that all countries set a sequence of local abatement efforts a_t^i so as to minimize local abatement costs and environmental damages and to maximize refunds r_t^i.

Proceeding as in Subsection 3.2, we assume momentarily that there exists a unique subgame perfect equilibrium for the subgame starting in period $t+1$ with a greenhouse gas stock s_{t+1}. Thus, we can write $W_{t+1}^i(s_{t+1})$ for country i's equilibrium payoff for this subgame. Then country i's best response in period t, $\bar{a}_t^i$, is determined by the solution of the optimization problem

$$V_t^i(s_t)|A_t^{-i} = \max_{a_t^i} \left\{ \delta W_{t+1}^i(s_{t+1}) - \frac{\alpha}{2}(a_t^i)^2 - \frac{\beta}{2}s_t^2 + r_t^i \right\} , \tag{15}$$

subject to equation (3), $W_{T+1}^i(s_{T+1}) \equiv 0$, and given the sum of the abatement efforts of all other countries $A_t^{-i} = \sum_{j \neq i} a_t^j$. In the proof of Proposition 3 we will establish that the optimization problem of country i in period t is strictly concave. Thus, differentiating equation (15) with respect to a_t^i and restricting our attention to interior solutions, we obtain an implicit function for country i's best response

$$\alpha \bar{a}_t^i = -\delta W_{t+1}^i{}'(\bar{s}_{t+1}) + \left. \frac{\partial r_t^i}{\partial a_t^i} \right|_{a_t^i = \bar{a}_t^i} , \tag{16}$$

where $\bar{s}_{t+1} = (1-\gamma)s_t + n\epsilon - \bar{a}_t^i - A_t^{-i}$ and

$$\frac{\partial r_t^i}{\partial a_t^i} = \begin{cases} R_t \dfrac{A_t^{-i}}{(a_t^i + A_t^{-i})^2} , & t = 1, \ldots, T-1 , \\ 0 , & t = T . \end{cases} \tag{17}$$

Applying the envelope theorem yields equation (11), as in the decentralized solution. Thus, the best response function (16) implies that country i's costs for abating an additional marginal unit of emissions have to equal the net present value of all mitigated local future damages caused by this additional marginal unit by decreasing the pollution stock s_t plus the refunds induced by abating this additional marginal unit. In case of the RS as well, the best-response function of country i only depends on the sum of the emission abatement of all other countries A_t^{-i} and not on the contribution of individual countries $a_t^j, j \neq i$.

Again, we can show the existence and uniqueness of a subgame perfect Nash equilibrium:

Proposition 3 (Refunding Scheme)
Given a time horizon $T < \infty$ and a feasible set of initial fees f_0^i and refunds R_t, there exists a unique subgame perfect Nash equilibrium characterized by identical sequences of emission abatements for all countries i in all periods t, $\tilde{a}_t^i = \tilde{a}_t$, and a sequence for the greenhouse gas stock $\tilde{s}_t$ $(i = 1, \ldots, n; \ t = 1, \ldots, T)$.

The proof of Proposition 3 is given in the Appendix.

4.3 First-best sustainable RS

We now show that there exists a feasible set of initial fees f_0^i and refunds R_t such that the unique Nash equilibrium of Proposition 3 is identical to the global social optimum characterized in Proposition 1. We call a treaty that exhibits this property *first-best sustainable refunding scheme*.

To prove the existence of a first-best sustainable RS, we first look for the sequence of refunds R_t that ensures that the unique Nash equilibrium under the refunding scheme equals the outcome of the social global optimum. In a second step, we calculate the minimal initial fund f_0 for which this sequence of refunds is feasible.

From equation (A.31) in the proof of Proposition 3, we know that in the unique and symmetric subgame perfect Nash equilibrium the following condition for the refund in period t, R_t, holds:

$$R_t = \frac{n^2 \tilde{a}_t}{n-1} \left[\alpha \tilde{a}_t + \delta W'_{t+1}(\tilde{s}_{t+1}) \right] , \tag{18a}$$

where $\tilde{a}_t$ and $\tilde{s}_t$ denote the levels of abatement and the pollution stock in period t in the subgame perfect Nash equilibrium. Recursively applying condition (A.33), we can rewrite equation (18a) to yield

$$R_t = \frac{n^2 \tilde{a}_t}{n-1} \left\{ \alpha \tilde{a}_t - \delta \beta \sum_{k=t+1}^{T} [\delta(1-\gamma)]^{k-(t+1)} \tilde{s}_k \right\} . \tag{18b}$$

Inserting the sequences $a_t^\star$ and $s_t^\star$ of the global social optimum, as characterized by Proposition 1, yields the sequence of refunds $R_t^\star$ for which the unique subgame perfect Nash equilibrium of the RS and the social global optimum coincide:

$$R_t^\star = \frac{n^2 a_t^\star}{n-1} \left\{ \alpha a_t^\star - \delta \beta \sum_{k=t+1}^{T} [\delta(1-\gamma)]^{k-(t+1)} s_k^\star \right\} . \tag{19}$$

To determine the minimal initial global fund $f_0^\star$ for which the sequence of refunds $R_t^\star$ is feasible, we re-write the recursive equation (13) to yield

$$f_0 = \sum_{t=1}^{T} \left[\frac{R_t}{(1+\rho)^t} \right] + \frac{f_T}{(1+\rho)^T} . \tag{20}$$

Thus, the minimal fund necessary to support the sequence of refunds $R_t^\star$ is given by

$$f_0^\star = \sum_{t=1}^{T-1} \left[\frac{R_t^\star}{(1+\rho)^t} \right] . \tag{21}$$

Note that it is optimal to empty the fund as early as period $T-1$, as the equal distribution of the remainder of the fund in period T does not influence the countries' abatement decisions in the last period. Hence $R_T^\star = 0$.

As a consequence, the following proposition holds:

11

Proposition 4 (Existence of First-best Sustainable RS)

For any time horizon $T < \infty$, the unique and symmetric subgame perfect Nash equilibrium of the RS coincides with the social global optimum for $f_0 \geq f_0^\star$ and $R_t = R_t^\star$ ($t = 1, \ldots, T$), where $f_0^\star$ and $R_t = R_t^\star$ are given by equations (21) and (19).

Proposition 4 says that for any initial global fund equal to or exceeding $f_0^\star$ the social global optimum can be implemented by setting $R_t = R_t^\star$. Note that the levels of abatement $a_t^\star$ and the greenhouse gas stock $s_t^\star$ are analytically solvable, as shown in the proof of Proposition 1. As a consequence, the levels of the minimal initial global fund $f_0^\star$ and the sequence of refunds $R_t^\star$ are also analytically solvable.

4.4 Second-best sustainable RS

As it may be difficult to agree ex ante on both an initial global fund f_0 and a sequence of refunds R_t, we now analyze a simplified version of the RS in which the total initial fund f_0 is invested and the refunds in each period equal the interest payments per period

$$R_t = \rho f_0 \equiv R , \quad t = 1, \ldots, T - 1 . \tag{22}$$

Then, the fund in period t is given by the initial fund, i.e. $f_t = f_0$ for all periods $t = 1, \ldots, T - 1$. Note that in the last period, the remainder of the fund is redistributed equally to the participating countries, so that $f_T = 0$ or $R_T = (1 + \rho)f_0$.

We show that under these conditions we can construct an RS for which the level of the greenhouse gas stock, $\tilde{s}_t$, converges to the global social optimal level $s_t^\star$ in the long run. We call such a treaty *second-best sustainable refunding scheme* as the global social optimum is only reached in the limit $t \to \infty$.

First, we calculate the long-run levels of abatement and the greenhouse gas stock under the RS that correspond to a particular level of the global fund f_0. Again, we analyze the unique Nash equilibrium of the RS for $T < \infty$ in the limit $T \to \infty$, which approximates the equilibrium of arbitrarily large but still finite time horizons T.

Proposition 5 (RS in the Long Run)
For the limit $T \to \infty$, the sequences of emission abatement and the greenhouse gas stock in the unique subgame perfect Nash equilibrium of the RS for finite time horizons T converge to the steady state levels

$$a^{RS} = \frac{n\beta\delta\epsilon + \sqrt{n^2\beta^2\delta^2\epsilon^2 + 4\gamma R \frac{n-1}{n^2}[1 - \delta(1 - \gamma)]\{\alpha\gamma[1 - \delta(1 - \gamma)] + n\beta\delta\}}}{2\{\alpha\gamma[1 - \delta(1 - \gamma)] + n\beta\delta\}} \quad , \quad (23a)$$

$$s^{RS} = \frac{n}{\gamma}(\epsilon - a^{RS}) \ . \quad (23b)$$

The proof of Proposition 5 is given in the Appendix.

Second, comparing the steady state levels (a^{RS}, s^{RS}) with the corresponding levels in the global social optimum (a^{SO}, s^{SO}), we obtain

Corollary 3 (Second-best Optimal Level of the Global Fund f)
The steady state of the RS (a^{RS}, s^{RS}) coincides with the steady state of the global social optimum (a^{SO}, s^{SO}) if and only if

$$f_0^{SB} = \frac{\alpha n^5 \beta^2 \delta^2 \epsilon^2}{\rho\{\alpha\gamma[1 - \delta(1 - \gamma)] + n^2\beta\delta\}^2} \ . \quad (24)$$

Proof: Solving $a^{RS} = a^{SO}$ for the fund f_0^{SB} yields (24). $\qquad \square$

Proposition 5 together with Corollary 3 immediately imply

Corollary 4 (Existence of Second-best Sustainable RS)
Given that a treaty with an initial global fund f_0^{SB} has been signed, it holds that

$$\lim_{t \to \infty} \tilde{s}_t = s^{RS} = s^{SO} \ . \quad (25)$$

Corollary 4 shows that there exists a RS that induces a convergence to the socially desired stock of greenhouse gases in the long run. Hence, the second-best sustainable RS provides a sustainable solution for the provision of the global public good of mitigating climate change in the long run. However, the path of abatement levels does not, in general, coincide with the socially optimal level. Although this difference in abatement levels between the second-best sustainable RS and the social global optimum vanishes over time, it may be substantial in the short run. The following proposition examines this difference:

13

Proposition 6 (Difference Second-best Sustainable RS – Social Optimum)

In the linear approximation around the steady state of the second-best sustainable RS, the following statements hold:

(i) The difference in the levels of the greenhouse gas stock between the second-best sustainable RS and the social global optimum is given by

$$\Delta s_t = \tilde{s}_t - s_t^\star = (s_1 - s^{SO}) \left[\nu_2^{t-1} - \lambda_2^{t-1} \right] \ , \tag{26}$$

with

$$\nu_2 = \frac{1}{2} \left[1 - \gamma + \frac{\alpha(2n-1) + n^2\beta\delta}{\alpha(2n-1)\delta(1-\gamma)} - \sqrt{\left(1 - \gamma + \frac{\alpha(2n-1) + n^2\beta\delta}{\alpha(2n-1)\delta(1-\gamma)} \right)^2 - \frac{4}{\delta}} \right] \ , \tag{27a}$$

$$\lambda_2 = \frac{1}{2} \left[1 - \gamma + \frac{\alpha + n^2\beta\delta}{\alpha\delta(1-\gamma)} - \sqrt{\left(1 - \gamma + \frac{\alpha + n^2\beta\delta}{\alpha\delta(1-\gamma)} \right)^2 - \frac{4}{\delta}} \right] \ . \tag{27b}$$

(ii) We have

$$\Delta s_t \begin{cases} < 0 \ , & if \quad s_1 < s^{SO} \ , \\ > 0 \ , & if \quad s_1 > s^{SO} \ . \end{cases} \tag{28}$$

Proposition 6 says that (at least in a sufficiently small neighborhood around the steady state) the saddle point path of greenhouse gas emissions in the second-best sustainable RS is below (above) the global social optimum if the initial stock of greenhouse gases is below (above) the long run steady state. Thus, the second-best sustainable RS induces inefficiently high (low) abatement levels for $s_1 < s^{SO}$ ($s_1 > s^{SO}$).

5 Numerical Illustration

To give an idea of the order of magnitude needed for the initial fund f_0 to implement the first- and second-best sustainable RS, we run a little numerical exercise. However, due to the highly stylized model, the results are rather a numerical illustration than a quantitative analysis.

Following similar numerical illustrations, such as those provided by Goulder and Mathai (2000), we adopt the slightly more complex but also slightly more realistic stock equation

14

by Nordhaus (1994):

$$s_{t+1} = (1 - \gamma)s_t + \theta \sum_{i=1}^{n} (\epsilon_t^i - a_t^i) \ , \tag{29}$$

where s_t denotes the CO_2 stock above the preindustrial level of 278 ppm, $\gamma = 0.008$, and $\theta = 0.64$. Thus, only the fraction θ accumulates the atmospheric stock. Moreover, we assume time-dependent business-as-usual emissions ϵ_t^i. Global business-as-usual emissions rise exponentially from approximately 8 GtC (3.75 ppm) in 2010 until they peak at 26 GtC (12.2 ppm) in 2125 and flatten out to 18 GtC by 2200.[8] As countries are identical, all countries exhibit the same business-as-usual emissions, which are $1/n$-th of the global business-as-usual emissions.

Of course, our assumption of identical countries drastically oversimplifies real-world affairs. For the numerical illustration, we assume that the world's economic activity is symmetrically distributed among ten identical countries. This is driven by the observation that the ten largest economies emitted more than 70% of global greenhouse gas emissions in 2006. This ratio is likely to rise further.[9] Thus, an international agreement among the 10 largest greenhouse gas emitters would essentially solve the problem of anthropogenic climate change. The parameters α and β are calibrated such that (a) the CO_2 stock in the global social optimum equals 2.5 times the preindustrial concentration (which is close to the optimal scenario in Nordhaus 1994) and (b) a doubling of the CO_2 concentration in the atmosphere against preindustrial levels amounts to environmental damages of 1.33% of world GDP per year (Nordhaus 1994, Goulder and Mathai 2000). According to the CIA World Factbook this equaled $6.5 \cdot 10^{13}$ US \$ in 2007. The values of the interest rate ρ and the discount factor δ are in line with Nordhaus (1994). Table 1 summarizes the parameter values used for our numerical example.

Figure 1 illustrates the business-as-usual path and shows total abatement, relative abatement, marginal abatement costs, the atmospheric stock for the decentralized solution, the first-best, and the second-best sustainable RS. In addition, yearly total refunds in % of 2007 world GDP are shown for the first-best and the second-best sustainable RS. Figure 1 gives rise to the following observations: First, we see that with respect to abatement levels and atmospheric CO_2 concentrations the decentralized solution falls dramatically short of the social global optimum (which is identical to the first-best sustainable RS).

Second, we observe that while in the long run the second-best sustainable RS converges to the social global optimum, abatement levels and marginal abatement costs differ dra-

---

[8]This business-as-usual emission scenario is similar to Goulder and Mathai (2000).

[9]The ten economies are China, USA, European Union, Indonesia, India, Russia, Brazil, Japan, Canada, and Mexico.
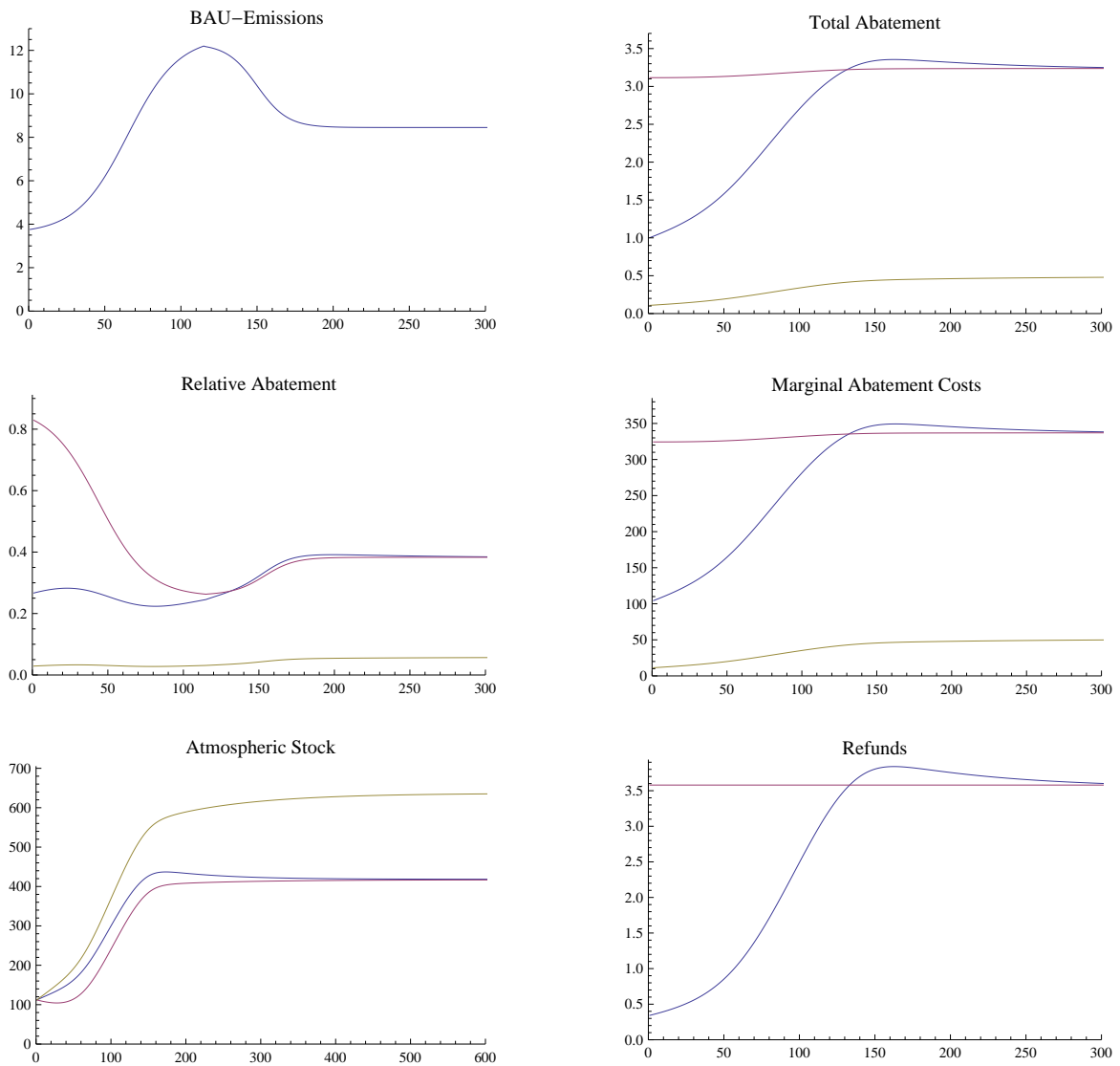
Figure 1: Global business-as-usual emissions in ppm are shown in the top left corner. Total abatement in ppm (top right), relative abatement (middle left), marginal abatement costs in US $ per tC (middle right) and the atmospheric stock in ppm above preindustrial level (bottom left) for the decentralized solution (green), the fist-best sustainable RS (blue) and the second-best sustainable RS (magenta). The bottom right corner shows total yearly refunds in % of world GDP for the first-best sustainable RS (blue) and the second-best sustainable RS (magenta).

Parameter		Value	Units
Initial stock (above pre-industrial level)	s_1	112	ppm
Abatement cost parameter	α	$2.21757 \cdot 10^{12}$	$\frac{\$/a}{(ppm/a)^2}$
Damage cost parameter	β	$9.94312 \cdot 10^6$	$\frac{\$/a}{ppm^2}$
Decay rate of GHGs	γ	0.008	
Discount factor	δ	0.952381	
Fraction of GHGs that accumulates s_t	θ	0.64	
Interest rate	ρ	0.05	
Number of countries	n	10	

Table 1: Summary of the parameter values used in the numerical example

matically for the first 125 years. Compared to the social global optimum, the second-best sustainable RS is in fact overambitious with respect to emission reductions. This is in line with Proposition 6, as the initial atmospheric CO_2 concentration is below the long run levels. As a consequence, the initial fund of the second-best sustainable RS amounts to 71.6 % of the world GDP in 2007 compared to 10.8 % for the first-best sustainable RS. Thus, from the current perspective the second-best sustainable RS seems rather unattractive.

However, it is possible, or even plausible, that countries may fail to reduce greenhouse gas emissions in the near future. Countries might also become more ambitious with respect to the long-run level of greenhouse gases in the atmosphere. For example, limiting global warming to a maximum of 2 °C roughly translates into a long-run concentration of greenhouse gases in the atmosphere amounting to 450 ppm. In both cases, the greenhouse gas concentration will move above the long-run level of the socially optimal stock. Then the second-best sustainable RS would ensure achievement of the long-run goal at lower abatement costs than the first-best sustainable RS. Third, we observe that while the total initial fund in the case of the first-best sustainable RS is also substantial (almost 11 % of world GDP), the yearly refunds for the first-best sustainable RS start at a moderate level of around 0.3 % of world GDP in 2010 and rise to around 3.75 % in 2170.

6 Initial Participation

So far, we have focused on the capacity of an RS to induce countries to follow (first-best sustainable RS) or to converge to (second-best sustainable RS) a socially optimal abatement path. To achieve this, all countries have to agree on the appropriate parameters (initial fees and, in the case of the first-best sustainable RS, a sequence of refunds in all periods) and on initial monetary commitments. We observe that a sustainable RS, as developed in this

paper, transforms the intertemporal climate-policy problem into a standard, static public-goods problem. Once all countries have made their initial contribution, a first-best allocation (or at least a long-run convergence to it in the case of the second-best sustainable RS) is ensured for all later occasions, as countries would be worse off by forfeiting refunds. In the following, we discuss how solution procedures developed in the literature on the private supply of public goods can be applied to motivate countries to make initial payments.

6.1 The ideal solution

At the initial level, when countries are pondering whether to sign the treaty and to pay the initial fee, the free-rider problem remains present. Especially if the number of countries n is large, each country may be better off by not signing the treaty. If all other countries participate, the country would benefit from all other countries' abatement efforts, without having to pay the initial fee and to compete for refunds.

To solve this free-rider problem, the RS could be incorporated in a two-stage game. In the first stage, countries decide whether to participate in a sustainable RS by paying the initial fee f_0^i. The treaty only becomes effective if all countries sign and pay the initial fee. Otherwise, the treaty is cancelled and initial fees already paid are returned. If all countries have signed, we can proceed to the second stage, in which the treaty is executed as outlined in Section 4.

It is straightforward to see that it is a weakly dominant strategy for all countries to sign the treaty and pay the initial fee. All countries are better off with the social global optimum achieved if the treaty becomes effective than they are with the decentralized solution that results if the treaty is cancelled.[10]

6.2 Difficulties in achieving initial participation

In practice, the preceding conceptual solution has to be supplemented by additional considerations. This holds, in particular, when countries are not identical, as this paper assumes.

Making larger countries pivotal

The ideal solution lies in making countries – and in particular large countries – pivotal for the formation of the RS. In order to achieve such a scenario, about ten to twenty of the

---

[10]While an improvement on the decentralized solution is obvious for the first-best sustainable RS, it is not clear *a priori* whether the second-best sustainable RS will improve on the decentralized solution in initial periods. However, there always exists a time period at which the second-best sustainable RS will be a Pareto improvement over the decentralized system for the entire remaining lifetime.

largest greenhouse gas emitters must coordinate on the agreement that the RS will fail if any of them defects.[11]

Sequential procedures

As full participation by all countries at once is unlikely, it is useful to resort to sequential procedures where a subset of countries makes a start and the others follow later (see Andreoni 1998, Lange 2006, Varian 1994). For the RS we might envision four steps. First, as suggested in the last paragraph, a set of large and mainly wealthy countries could initiate the system by paying initial fees. Second, smaller rich countries could follow, which would increase the initial wealth. In the third and fourth steps, larger and smaller developing countries could be invited to join the RS. Regarding the payment of initial fees, they should be treated differently, as we will discuss next.

Renouncing initial fees for developing countries

Developing countries lack the necessary wealth to pay the initial fees.[12] To induce participation, payment of initial fees could be forgone. Indeed, the RS works in the same way if a subset of k countries ($1 < k < n$) pays an initial fee equal to $f_0^\star/k$ and n countries are eligible to the refunds. Once the RS has been initiated, incentives to abate are independent of initial contributions. In such circumstances, developing countries would voluntarily join the system, as they can always choose the same emission reduction policy under the RS as without, but they can benefit from the refunds if it is in their interest. Indeed, the prospect of earning refunds will motivate them to abate much more. Moreover, the discrimination of initial fees is a powerful tool in implementing transfers to the RS. If countries are heterogeneous, that may be necessary to render the RS a Pareto improvement.

7 Raising Initial Fees

The sustainable refunding scheme relies on the payment of initial fees. As such fees tend to be quite large, especially in the case of the second-best sustainable RS, we outline two ways in which such fees might be financed.

---

[11] In practice countries must be mutually stubborn and insist on full participation by this core group before going ahead.

[12] On average, the economic and social consequences of climate change also tend to be more severe in developing countries.

7.1 Repeated payments

Paying the initial fees in full at the beginning of the treaty is not necessary. We can also achieve the first- and second-best sustainable RS by repeatedly paying a smaller amount of money.

Proposition 7 (RS with Repeated Payments)

(i) For every time span $\Delta > 0$ there exists a sequence of fees $f_t(\Delta)$ defined by

$$f_t(\Delta) = \sum_{\tau=1}^{\Delta} \frac{R^\star_{t+\tau}}{(1+\rho)^\tau} \tag{30}$$

such that this RS implements the same solution as the first-best sustainable RS with initial payments larger or equal to $f_0^\star$ if $f_t(\Delta)$ is paid into the fund at times $t = 0, \Delta, 2\Delta, \ldots$.

(ii) For every time span $\Delta > 0$ there exist fees $f(\Delta)$ defined by

$$f(\Delta) = \frac{(1+\rho)^\Delta - 1}{(1+\rho)^\Delta} f_0^{SB} \tag{31}$$

such that this RS implements the same solution as the second-best sustainable RS with initial payments f_0^{SB} if $f(\Delta)$ is paid into the fund each Δ periods.

The proof of Proposition 7 is given in the Appendix.

With the repeated payments scheme we face a trade-off between high initial fees and the property of the sustainable RS to transform an intertemporal climate-policy problem into a static public-goods problem. In particular, if the time span Δ is short, the solution of the climate-change problem relies on the repeated commitment of all countries, as the initial participation problem would have to be solved whenever new payments have to be made. Therefore Δ should not be too small.[13]

7.2 Borrowing and capital markets

If the repeated solution to the initial participation problem turns out to be a major obstacle to international cooperation, raising the initial fees by allowing countries to borrow money may be more advisable. Countries could then borrow either from the international capital market or directly from the administering agency of the RS. As the administering agency

---

[13]Sustained participation can also be fostered by not completely depleting the fund and hence making exit costly.

invests the wealth of the climate fund on the international capital market, both ways by countries to borrow are equivalent when capital markets are perfect, as the following proposition shows:

Proposition 8 (Borrowing Initial Fees)
Suppose capital markets are perfect and all countries borrow the entire amount of the initial fees f_0^i required for the sustainable RS. Use μ ($0 \leq \mu \leq 1$) to denote the fraction of f_0^i borrowed from the international capital market, implying that the remainder $(1 - \mu)f_0^i$ is borrowed from the administering agency running the RS. Then the following statements hold:

(i) *For all values of μ, borrowing by countries does not affect international capital markets.*

(ii) *For $\mu = 0$, the first- and second-best sustainable RS can be implemented without the flow of money.*

Proof: Recall that perfect capital markets imply that countries may borrow or lend freely at the per-period interest rate ρ. To see part (i), observe that borrowing by countries increases demand for capital on the international capital market by a total of μf_0. As the administering agency lends a total of $(1 - \mu)f_0$ to the countries, it can invest μf_0 in the capital market and thus supply increases by the same amount. Hence, the equilibrium in the capital market is not affected.[14] Part (ii) follows from the observation that in the (first-best and second-best) sustainable RS the initial fee f_0^i is the net present value of all future refunds the country receives in the unique subgame perfect equilibrium. So in each time period t, the net present value of the two claims exactly offset each other. No actual money flow is necessary in the polar case $\mu = 0$. □

Proposition 8 says that raising the money needed for the initial payments is no problem under the assumption of perfect capital markets. In practice, two types of deviations from perfect capital markets have to be taken into account. A country may default against the administering agency or default in general. First, if μ is small, countries might be tempted to renounce high abatement efforts and to default on their interest-rate obligations to the administering agency. The country would lose all claims to refunds. However, as such refunds are small when abatement efforts are small, such a strategy may be profitable. For say $\mu = 0$, a country could choose to default against the administering agency and could free-ride on the abatement efforts of other countries even if it has signed the treaty and has borrowed from the administering agency. Such considerations suggest that countries should borrow mainly on the international capital market.

---

[14]As the administering agency needs to invest its wealth in the capital market, borrowing by countries to pay the initial fee does not crowd out private investments.

Second, if countries borrow a large amount on international capital markets, the default risk may rise if outstanding government debt is already at a high level. If the country needs to pay a larger interest rate than the risk-free rate, as investors demand a positive risk premium, further borrowing may increase the default risk as refunds are insufficient to cover interest-rate payments. In such cases, it is useful for part of the initial fund to be raised by taxes so as to foster abatement. Then the additional interest-rate burden can be kept smaller than refunds, thereby keeping or reducing default risk on capital markets.

8 Conclusion

The RS provides a simple blueprint for an international treaty on climate change. It is governed by a very small number of parameters. The RS is no panacea, as free-rider problems have no perfect solutions, but it might be wiser to focus attention on systems like the RS than on Kyoto-style treaties in which little can be done to induce countries to fulfill their promises.

The practical implementation of the refunding schemes developed in this paper requires a variety of additional considerations. In the last two sections, we have discussed how to achieve initial participation, and we have outlined several ways of raising initial fees. Other issues, such as governance of the administering agency, uncertainty regarding damage and abatement costs, or heterogeneous countries will need thorough investigation in future research.

Appendix

Proof of Proposition 1

We show existence and uniqueness of the social global optimum by solving the optimization problem (5) for any time horizon T.

In each period t, the right-hand side of equation (6) is identical for all countries i. This implies that abatement in period t is equal among all countries, i.e., $a_t^i = a_t$ for all $i \in \{1, \ldots, n\}$ and $t = 1, \ldots, T$. We rewrite (6) to yield

$$V_{t+1}'(s_{t+1}) = -\frac{\alpha}{\delta} a_t . \tag{A.1}$$

Inserting into (7), we eliminate the value function and obtain

$$a_{t-1} = \delta(1-\gamma) a_t + \frac{n\beta\delta}{\alpha} s_t . \tag{A.2}$$

Equation (A.2) together with the equation of motion (3) yields the system of linear first-order difference equations

$$\begin{pmatrix} a_{t+1}^\star \\ s_{t+1}^\star \end{pmatrix} = \begin{pmatrix} \frac{\alpha + n^2\beta\delta}{\alpha\delta(1-\gamma)} & -\frac{n\beta}{\alpha} \\ -n & 1-\gamma \end{pmatrix} \begin{pmatrix} a_t^\star \\ s_t^\star \end{pmatrix} + \begin{pmatrix} -\frac{n^2\beta\epsilon}{\alpha(1-\gamma)} \\ n\epsilon \end{pmatrix} \tag{A.3}$$

the general solution of which is given by

$$\begin{pmatrix} a_{t+1}^\star \\ s_{t+1}^\star \end{pmatrix} = \begin{pmatrix} a^{SO} \\ s^{SO} \end{pmatrix} + B_1(T) v_1 \lambda_1^t + B_2(T) v_2 \lambda_2^t , \tag{A.4}$$

where (a^{SO}, s^{SO}) are the stationary states obtained by setting $a_{t+1}^\star = a_t^\star = a^{SO}$ and $s_{t+1}^\star = s_t^\star = s^{SO}$, hence

$$a^{SO} = \frac{n^2\beta\delta\epsilon}{\alpha\gamma \left[1 - \delta(1-\gamma)\right] + n^2\beta\delta} , \tag{A.5a}$$

$$s^{SO} = \frac{n\alpha\epsilon \left[1 - \delta(1-\gamma)\right]}{\alpha\gamma \left[1 - \delta(1-\gamma)\right] + n^2\beta\delta} , \tag{A.5b}$$

and v_i and λ_i, $i = 1, 2$ are the Eigen vectors and the Eigen values, respectively, of the 2×2 matrix in (A.3). $B_i(T)$, $i = 1, 2$ are constants that have to be determined in such a way that the initial and final conditions are satisfied.

Calculating Eigen values and vectors yields

$$\lambda_{1/2} = \frac{1 + \delta(1-\gamma)^2 + \frac{n^2\beta\delta}{\alpha} \pm \sqrt{\left[1 + \delta(1-\gamma)^2 + \frac{n^2\beta\delta}{\alpha}\right]^2 - 4\delta(1-\gamma)^2}}{2\delta(1-\gamma)} \tag{A.6a}$$

$$v_1 = \begin{pmatrix} 1 \\ \frac{n}{1-\gamma-\lambda_1} \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ \frac{n}{1-\gamma-\lambda_2} \end{pmatrix}. \tag{A.6b}$$

Setting $C = 1 - \delta + \delta\gamma^2 + \frac{n^2\beta\delta}{\alpha}$ and $D = \left[1 + \delta(1-\gamma)^2 + \frac{n^2\beta\delta}{\alpha}\right]^2 - 4\delta(1-\gamma)^2$, the Eigen values read

$$\lambda_{1/2} = 1 + \frac{C \pm \sqrt{D}}{2\delta(1-\gamma)}. \tag{A.7}$$

As $C, D > 0$ and $\sqrt{C} > D$, we obtain $\lambda_1 > 1 > \lambda_2 > 0$.

Then the general solution is given by

$$a_t^\star = a^{SO} + B_1(T)\lambda_1^t + B_2(T)\lambda_2^t, \tag{A.8a}$$

$$s_t^\star = s^{SO} + \frac{nB_1(T)}{1-\gamma-\lambda_1}\lambda_1^t + \frac{nB_2(T)}{1-\gamma-\lambda_2}\lambda_2^t. \tag{A.8b}$$

We determine $B_1(T)$ and $B_2(T)$ from the initial greenhouse gas stock, s_1, and the terminal condition for emission abatement, $a_T = 0$:

$$B_1(T) = \frac{(1-\gamma-\lambda_1)\left[-\frac{na^{SO}}{\lambda_1^T} - \frac{\lambda_2^{T-1}}{\lambda_1^T}(1-\gamma-\lambda_2)(s_1 - s^{SO})\right]}{n\left[(1-\gamma-\lambda_1) - \frac{\lambda_2^{T-1}}{\lambda_1^{T-1}}(1-\gamma-\lambda_2)\right]}, \tag{A.9a}$$

$$B_2(T) = \frac{(1-\gamma-\lambda_2)\left[(1-\gamma-\lambda_1)(s_1 - s^{SO}) + \frac{na^{SO}}{\lambda_1^{T-1}}\right]}{n\left[(1-\gamma-\lambda_1)\lambda_2 - \frac{\lambda_2^T}{\lambda_1^{T-1}}(1-\gamma-\lambda_2)\right]}. \tag{A.9b}$$

Inserting back into equations (A.8) yields the unique global social optimum. □

Proof of Corollary 1

For large T we obtain for $B_1(T)$ and $B_2(T)$ from the proof of Proposition 1

$$B_1^\infty \equiv \lim_{T \to \infty} B_1(T) = 0, \quad B_2^\infty \equiv \lim_{T \to \infty} B_2(T) = \frac{(1-\gamma-\lambda_2)(s_1 - s^{SO})}{n\lambda_2}, \tag{A.10}$$

implying for the solution (A.8) in the limit case $T \to \infty$

$$a_t^\star = a^{SO} + \frac{(1 - \gamma - \lambda_2)(s_1 - s^{SO})}{n}\lambda_2^{t-1} \, , \tag{A.11a}$$

$$s_t^\star = s^{SO} + (s_1 - s^{SO})\lambda_2^{t-1} \, , \tag{A.11b}$$

which we can also write as a policy rule $a_t(s_t)$:

$$a_t^\star(s_t^\star) = a^{SO} + \frac{(1 - \gamma - \lambda_2)(s_t^\star - s^{SO})}{n} \, . \tag{A.12}$$

From the proof of Proposition 1 we know that $0 < \lambda_2 < 1$. Thus we obtain from (A.11)

$$\lim_{t\to\infty} a_t^\star = a^{SO} \, , \qquad \lim_{t\to\infty} s_t^\star = s^{SO} \, . \tag{A.13}$$

$\square$

Proof of Proposition 2

Before we show the existence of a unique and symmetric subgame perfect Nash equilibrium by backward induction, note that the optimization problem of country i in period t is strictly concave if and only if

$$\delta W_{t+1}^i{}''(s_{t+1}) - \alpha < 0 \, . \tag{A.14}$$

Starting in period T, recall that $W_{T+1}^i(s_{T+1}) \equiv 0$, which implies that $\bar{a}_T^i = 0$ is the best response of all countries independently of the emission abatement choices of all other countries. As a consequence, $\hat{a}_T = \hat{a}_T^i$ is the unique and symmetric Nash equilibrium for the subgame starting in period T given the level of the greenhouse gas stock s_T. The equilibrium pay-off $W_T^i(s_t) = V_T^i(s_T)|\hat{A}_T^{-i}$ is identical for all countries and is strictly concave:

$$W_T(s_T) \equiv W_T^i(s_T) = -\frac{\beta}{2}s_T^2 \qquad \Rightarrow \qquad W_T''(s_T) = -\beta \, . \tag{A.15}$$

Now assume there exists a unique and symmetric subgame perfect Nash equilibrium for the subgame starting in period $t + 1$ with a greenhouse gas stock of s_{t+1} yielding symmetric equilibrium pay-offs $W_{t+1}(s_{t+1}) = W_{t+1}^i(s_{t+1})$ with $W_{t+1}''(s_{t+1}) < 0$. Then the optimization problem in period t is strictly concave for all countries i, implying there exists a unique best response $\bar{a}_t^i$ for all countries i given the emission abatements of all other countries $j \neq i$, which is given implicitly by

$$\alpha \bar{a}_t^i = -\delta W_{t+1}^i{}'(\bar{s}_{t+1}) \, . \tag{A.16}$$

As the right-hand side is identical by assumption for all countries, so is the left-hand side as well. This implies that emission abatement in equilibrium is symmetric and unique, $\hat{a}_t = \hat{a}_t^i$. As a consequence, the equilibrium pay-off is also identical for all countries i, $W_t(s_t) = W_t^i(s_t)$. Differentiating (11) with respect to s_t, we obtain

$$V_t^{i''}(s_t)|A_t^{-i} = \delta(1-\gamma)^2 W_{t+1}^{i''}(\bar{s}_{t+1}) - \beta \ . \tag{A.17}$$

As $W_T^{i''}(s_t) = V_t^{i''}(s_t)|\hat{A}_T^{-i}$, this implies that the equilibrium pay-off $W_t^i(s_t) = W_t(s_t)$ is strictly concave. Working backwards to $t = 1$ yields a unique symmetric sequence of emission abatements $\hat{a}_t^i = \hat{a}_t$ and the corresponding sequence of the greenhouse gas stock $\hat{s}_t$ ($i = 1, \ldots, n$; $t = 1, \ldots, T$) that constitute the unique and symmetric subgame perfect Nash equilibrium of the decentralized system.

As the subgame perfect Nash equilibrium is symmetric and unique, we obtain the following system of first-order difference equations from equations (10), (11) and (3):

$$\begin{pmatrix} \hat{a}_{t+1} \\ \hat{s}_{t+1} \end{pmatrix} = \begin{pmatrix} \frac{\alpha+n\beta\delta}{\alpha\delta(1-\gamma)} & -\frac{\beta}{\alpha} \\ -n & 1-\gamma \end{pmatrix} \begin{pmatrix} \hat{a}_t \\ \hat{s}_t \end{pmatrix} + \begin{pmatrix} -\frac{n\beta\epsilon}{\alpha(1-\gamma)} \\ n\epsilon \end{pmatrix} \ . \tag{A.18}$$

Following the same solution technique as described in the proof of Proposition 1, we derive for the sequences of emission abatement and the greenhouse gas stock in the subgame perfect Nash equilibrium

$$\hat{a}_t = a^{DS} + B_1(T)\mu_1^t + B_2(T)\mu_2^t \ , \tag{A.19a}$$

$$\hat{s}_t = s^{DS} + \frac{nB_1(T)}{1-\gamma-\mu_1}\mu_1^t + \frac{nB_2(T)}{1-\gamma-\mu_2}\mu_2^t \ , \tag{A.19b}$$

where (a^{DS}, s^{DS}) denote the steady state of (A.18) given by

$$a^{DS} = \frac{n\beta\delta\epsilon}{\alpha\gamma\left[1-\delta(1-\gamma)\right]+n\beta\delta} \ , \tag{A.20a}$$

$$s^{DS} = \frac{n\alpha\epsilon\left[1-\delta(1-\gamma)\right]}{\alpha\gamma\left[1-\delta(1-\gamma)\right]+n\beta\delta} \ , \tag{A.20b}$$

μ_1 and μ_2 equal

$$\mu_{1/2} = \frac{1+\delta(1-\gamma)^2+\frac{n\beta\delta}{\alpha} \pm \sqrt{\left[1+\delta(1-\gamma)^2+\frac{n\beta\delta}{\alpha}\right]^2 - 4\delta(1-\gamma)^2}}{2\delta(1-\gamma)}$$

$$= 1 + \frac{1-\delta+\delta\gamma^2+\frac{n\beta\delta}{\alpha} \pm \sqrt{\left[1-\delta(1-\gamma)^2+\frac{n\beta\delta}{\alpha}\right]^2 + \frac{4\delta^2(1-\gamma)^2\beta n}{\alpha}}}{2\delta(1-\gamma)} \ , \tag{A.21}$$

which immediately implies $\mu_1 > 1 > \mu_2 > 0$, and $B_1(t)$ and $B_2(T)$ depend on the time horizon T and the initial stock of greenhouse gases s_1

$$B_1(T) = \frac{(1 - \gamma - \mu_1)\left[-\frac{na^{DS}}{\mu_1^T} - \frac{\mu_2^{T-1}}{\mu_1^T}(1 - \gamma - \mu_2)(s_1 - s^{DS})\right]}{n\left[(1 - \gamma - \mu_1) - \frac{\mu_2^{T-1}}{\mu_1^{T-1}}(1 - \gamma - \mu_2)\right]} , \tag{A.22a}$$

$$B_2(T) = \frac{(1 - \gamma - \mu_2)\left[(1 - \gamma - \mu_1)(s_1 - s^{DS}) + \frac{na^{DS}}{\mu_1^{T-1}}\right]}{n\left[(1 - \gamma - \mu_1)\mu_2 - \frac{\mu_2^T}{\mu_1^{T-1}}(1 - \gamma - \mu_2)\right]} . \tag{A.22b}$$

$\square$

Proof of Corollary 2

For large T we obtain for $B_1(T)$ and $B_2(T)$ from the proof of Proposition 2

$$B_1^\infty \equiv \lim_{T \to \infty} B_1(T) = 0 , \qquad B_2^\infty \equiv \lim_{T \to \infty} B_2(T) = \frac{(1 - \gamma - \mu_2)(s_1 - s^{DS})}{n\mu_2} , \tag{A.23}$$

implying for solution (A.19) in the limit case $T \to \infty$

$$\hat{a}_t = a^{DS} + \frac{(1 - \gamma - \mu_2)(s_1 - s^{DS})}{n}\mu_2^{t-1} , \tag{A.24a}$$

$$\hat{s}_t = s^{DS} + (s_1 - s^{DS})\mu_2^{t-1} , \tag{A.24b}$$

which we can also write as a policy rule $a_t(s_t)$:

$$\hat{a}_t(\hat{s}_t) = a^{DS} + \frac{(1 - \gamma - \mu_2)(\hat{s}_t - s^{DS})}{n} . \tag{A.25}$$

From the proof of Proposition 2 we know that $0 < \mu_2 < 1$. Thus, we obtain from (A.24)

$$\lim_{t \to \infty} \hat{a}_t = a^{DS} , \qquad \lim_{t \to \infty} \hat{s}_t = s^{DS} \tag{A.26}$$

$\square$

Proof of Proposition 3

Again, we start by noting that the optimization problem of country i in period t is strictly concave if and only if

$$\delta W_{t+1}^i{}''(s_{t+1}) - \alpha + \frac{\partial^2 r_t^i}{(\partial a_t^i)^2} < 0 , \tag{A.27}$$

with

$$\frac{\partial^2 r^i_t}{(\partial a^i_t)^2} = \begin{cases} -2R_t \frac{A_t^{-i}}{(a^i_t + A_t^{-i})^3} \ , & t = 1, \ldots, T-1 \ , \\ 0 \ , & t = T \ . \end{cases} \tag{A.28}$$

Solving for the subgame perfect Nash equilibria by backward induction, we start in period T. By virtue of the first-order condition (16), $W^i_{T+1}(s_{T+1}) \equiv 0$ implies that $\bar{a}^i_T = 0$ is the best response for all countries independently of the emission abatement choices of all other countries. As a consequence, $\tilde{a}_T = \tilde{a}^i_T$ is the unique and symmetric Nash equilibrium for the subgame starting in period T, given the level of the greenhouse gas stock s_T. The equilibrium pay-off $W^i_T(s_T) = V^i_T(s_T)|\tilde{A}^{-i}_T$ is identical for all countries and is strictly concave:

$$W_T(s_T) \equiv W^i_T(s_T) = -\frac{\beta}{2}s_T^2 + \frac{R_T}{n} \qquad \Rightarrow \qquad W''_T(s_T) = -\beta \ . \tag{A.29}$$

In addition, $W'_T(s_T) = -\beta s_T < 0$.

Now assume there exists a unique and symmetric subgame perfect Nash equilibrium for the subgame starting in period $t+1$, with a greenhouse gas stock of s_{t+1} yielding symmetric equilibrium pay-offs $W_{t+1}(s_{t+1}) \equiv W^i_{t+1}(s_{t+1})$ with $W'_{t+1}(s_{t+1}) < 0$ and $W''_{t+1}(s_{t+1}) < 0$. Then the optimization problem in period t is strictly concave for all countries i, implying there exists a unique best response $\bar{a}^i_t$ for all countries i given the emission abatements of all other countries $j \neq i$, which is given implicitly by

$$\alpha \bar{a}^i_t - R_t \frac{A_t^{-i}}{(\bar{a}^i_t + A_t^{-i})^2} = -\delta W'_{t+1}(\bar{s}_{t+1}) \ . \tag{A.30}$$

As the right-hand side is identical by assumption for all countries, so is the left-hand side as well. This implies that emission abatement in equilibrium is symmetric, $\tilde{a}_t \equiv \tilde{a}^i_t$ for all $i = 1, \ldots, n$. Summing up over all n countries and multiplying by the total amount of abatement A_t (which is strictly positive, as corner solutions are ruled out) yields a necessary condition that has to hold in the subgame perfect Nash equilibrium

$$\alpha \tilde{A}^2_t + \delta n W'_{t+1}(\tilde{s}_{t+1})\tilde{A}_t - (n-1)R_t = 0 \ . \tag{A.31}$$

This equation yields a unique solution, as total emissions have to be non-negative and, in addition, $W'_{t+1}(s_{t+1}) < 0$ holds:

$$\tilde{A}_t = \frac{-\delta n W'_{t+1}(\tilde{s}_{t+1}) + \sqrt{(\delta n W'_{t+1}(\tilde{s}_{t+1}))^2 + 4\alpha(n-1)R_t}}{2\alpha} \ . \tag{A.32}$$

We already know that emission abatement levels in the subgame perfect Nash equilibrium

are symmetric, so this implies unique and symmetric abatement levels $\tilde{a}_t = \tilde{A}_t/n$ and also identical equilibrium pay-offs for all countries i, $W_t(s_t) = W_t^i(s_t)$. The envelope theorem yields equation (11), from which we obtain

$$W_t'(s_t) = V_t^{i'}(s_t)|\tilde{A}_t^{-i} = -\beta s_t + \delta(1-\gamma)W_{t+1}'(\bar{s}_{t+1}) < 0 . \tag{A.33}$$

In addition, differentiating with respect to s_t, we observe

$$W_t''(s_t) = V_t^{i''}(s_t)|\tilde{A}_t^{-i} = \delta(1-\gamma)^2 W_{t+1}''(\bar{s}_{t+1}) - \beta < 0 , \tag{A.34}$$

implying that the equilibrium pay-off $W_t(s_t)$ is strictly concave.

Working backwards to $t = 1$ yields a unique symmetric sequence of emission abatements $\tilde{a}_t^i = \tilde{a}_t$ and the corresponding sequence of the greenhouse gas stock $\tilde{s}_t$ $(i = 1, \ldots, n; t = 1, \ldots, T)$ that constitute the unique and symmetric subgame perfect Nash equilibrium of the RS. $\qquad \square$

Proof of Proposition 5
In the subgame perfect Nash equilibrium, the first-order condition (16) reads

$$\alpha a_t = -\delta W_{t+1}'(s_{t+1}) + \frac{(n-1)R}{n^2 a_t} , \tag{A.35}$$

where we drop the tilde notation '~' for presentational convenience. Together with equation (A.33), this implies that the sequences a_t and s_t in the subgame perfect Nash equilibrium of the RS are given by the solution of the following system of first-order difference equations:

$$R\frac{n-1}{n^2 a_t} - \alpha a_t = \delta(1-\gamma)(R\frac{n-1}{n^2 a_{t+1}} - \alpha a_{t+1}) - \delta\beta s_{t+1} , \tag{A.36a}$$

$$s_{t+1} = (1-\gamma)s_t + n(\epsilon - a_t) , \tag{A.36b}$$

with the boundary conditions s_1 and $a_T = 0$ for finite time horizons T. Solving equation (A.36a) with respect to a_{t+1} yields

$$a_{t+1} = \frac{g(a_t) - \beta\delta(1-\gamma)s_t + \sqrt{[g(a_t) - \beta\delta(1-\gamma)s_t]^2 + 4\alpha\delta^2(1-\gamma)^2 R\frac{n-1}{n^2}}}{2\alpha\delta(1-\gamma)} , \tag{A.36c}$$

with $g(a) \equiv (\alpha + n\beta\delta)a - R(n-1)/(n^2 a) - n\beta\delta\epsilon$. Solving for the steady state and taking into account the fact that negative abatement levels are infeasible yields equations (23).

We now analyze the system dynamics for $T \to \infty$. We show that the system dynamics splits into three different regimes. (i) For any initial stock s_1 there exists a corresponding

initial level of abatement a_1^{SP} such that the sequences a_t and s_t converge to their steady state values (a^{RS}, s^{RS}). (ii) If the initial level of abatement is above a_1^{SP}, then the system dynamics will converge to the corner solution $(\epsilon, 0)$ in which emission abatement is maximal $a_{max} = \epsilon$, i.e. all emissions are abated, and the greenhouse gas stock converges to 0. (iii) If the initial level of abatement is below a_1^{SP}, then the system dynamics will converge to the corner solution $(0, n\epsilon/\gamma)$ in which no emissions are abated and the greenhouse gas stock converges to its maximum value $s_{max} = n\epsilon/\gamma$.

To see (i), we show that the steady state (a^{RS}, s^{RS}) is a saddle point. Denoting the Jacobian of the system of difference equations evaluated at the steady state by

$$J(a^{RS}, s^{RS}) = \begin{bmatrix} \dfrac{\partial a_{t+1}}{\partial a_t}(a^{RS}, s^{RS}) & \dfrac{\partial a_{t+1}}{\partial s_t}(a^{RS}, s^{RS}) \\ \dfrac{\partial s_{t+1}}{\partial a_t}(a^{RS}, s^{RS}) & \dfrac{\partial s_{t+1}}{\partial s_t}(a^{RS}, s^{RS}) \end{bmatrix} = \begin{bmatrix} J_{11} J_{12} \\ J_{21} J_{22} \end{bmatrix} , \tag{A.37}$$

we obtain

$$J_{11} = \frac{\alpha + R\frac{n-1}{n^2(a^{RS})^2} + \delta\beta n}{\delta(1-\gamma)\left(\alpha + R\frac{n-1}{n^2(a^{RS})^2}\right)} , \tag{A.38a}$$

$$J_{12} = -\frac{\beta}{\alpha + R\frac{n-1}{n^2(a^{RS})^2}} , \tag{A.38b}$$

$$J_{21} = -n , \tag{A.38c}$$

$$J_{22} = 1 - \gamma , \tag{A.38d}$$

where we have derived (A.38a) and (A.38b) by differentiating (A.36a) with respect to a_t and s_t respectively and inserting the steady state. Then the characteristic equation in the linearization around the steady state reads

$$\nu^2 - (J_{11} + J_{22})\nu + J_{11}J_{22} - J_{12}J_{21} = 0 , \tag{A.39}$$

which yields the eigenvalues

$$\nu_{1/2} = \frac{J_{11} + J_{22} \pm \sqrt{(J_{11} + J_{22})^2 - 4(J_{11}J_{22} - J_{12}J_{21})}}{2} . \tag{A.40}$$

It can be shown that

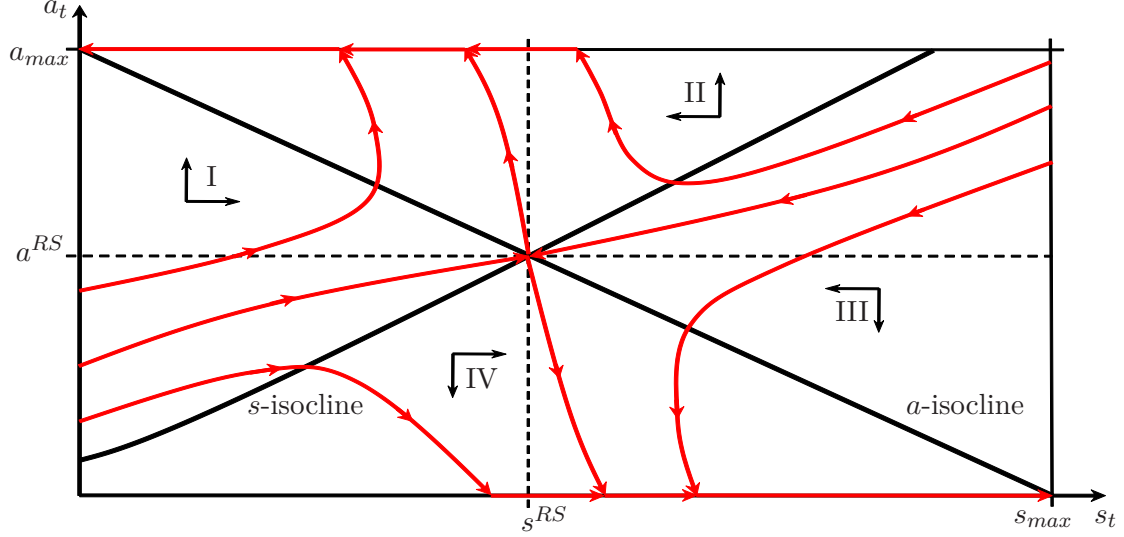$$J_{11} + J_{22} > \sqrt{(J_{11} + J_{22})^2 - 4(J_{11}J_{22} - J_{12}J_{21})} > J_{11} - J_{22} , \tag{A.41}$$

Figure 2: Phase diagram of the system dynamics of the system of difference equations (A.36) for $T \to \infty$. The steady state (a^{RS}, s^{RS}) is a saddle point. The a- and s-isoclines divide the feasible space into four areas I–IV.

implying that

$$\nu_1 > J_{11} > 1 , \qquad 0 < \nu_2 < J_{22} = 1 - \gamma < 1 . \tag{A.42}$$

Thus, the steady state (a^{RS}, s^{RS}) is a saddle point. As a consequence, for any initial value of the greenhouse gas stock s_1 there exists a corresponding initial level of abatement a_1^{SP} such that the sequences a_t and s_t converge to the steady-state values (a^{RS}, s^{RS}).

Figure 2 sketches the system dynamics. The isoclines are given by

$$a_{t+1} - a_t = 0 \quad \Leftrightarrow \quad a_t = \frac{n\beta\delta s_t + \sqrt{(n\beta\delta s_t)^2 + 4\alpha\rho f \frac{n-1}{n^2}[1 - \delta(1-\gamma)]^2}}{2n\alpha[1 - \delta(1-\gamma)]} \tag{A.43a}$$

$$s_{t+1} - s_t = 0 \quad \Leftrightarrow \quad a_t = \epsilon - \frac{\gamma}{n}s_t , \tag{A.43b}$$

They determine the combinations of s_t and a_t for which $a_{t+1} - a_t = 0$ and $s_{t+1} - s_t = 0$, respectively, and divide the a_t-s_t-plane into four areas. In area I, which is above the a- and below the s-isocline, $a_{t+1} > a_t$ and $s_{t+1} > s_t$. In area II, given by the segment above both isoclines, $a_{t+1} > a_t$ and $s_{t+1} < s_t$. Area III is below the a- and above the s-isocline, in which $a_{t+1} < a_t$ and $s_{t+1} > s_t$ hold. Finally, area IV is below both isoclines, so $a_{t+1} < a_t$ and $s_{t+1} > s_t$. The saddle point path lies in area I for $s_1 < s^{RS}$ and in area III for $s_1 > s^{RS}$. Thus, the saddle point path starts with abatement levels a_1^{SP} which are below (above) the steady-state abatement level a^{RS} and increase (decrease) over time to converge to a^{RS} for

31

$s_1 < s^{RS}$ ($s_1 > s^{RS}$).

We also observe that all, although they may start in area I, II or IV, paths for which $a_1 > a_1^{SP}$ eventually reach the maximum abatement level $a_{t'} = a_{max} = \epsilon$. In this case the system dynamics changes regimes, as no abatement levels above ϵ are feasible. As a consequence, $a_t = \epsilon$ for all $t \geq t'$. Then the stock of greenhouse gases, as determined by equation (A.36b), converges to 0 for $t \to \infty$. In a similar vain, all paths for which the initial abatement level is smaller than the initial saddle point path abatement level a_1^{SP} eventually hit the lower boundary $a_{t'} = 0$. Again, the system dynamics changes regimes, as abatement levels below 0 are infeasible, so $a_t = 0$ for all $t \geq t'$. As a consequence, the greenhouse gas stock converges to its maximum value $s_{max} = (n\epsilon)/\gamma$ for $t \to \infty$.

By assumption, the corner solutions $a_t = 0$ and $a_t = \epsilon$ cannot be best responses for any country i to all feasible greenhouse gas stocks s_t and any given abatement levels of all other countries $j \neq i$. As a consequence, all paths that either converge to $(\epsilon, 0)$ or $(0, (n\epsilon)/\gamma)$ are not subgame perfect Nash equilibria, leaving the saddle point path as the only subgame perfect Nash equilibrium. $\qquad \square$

Proof of Proposition 6
For $f_0 = f_0^{SB}$ as given by equation (24), the linear approximation around the stationary state of the system dynamics of the second-best sustainable RS is given by

$$\tilde{a}_t = a^{SO} + \frac{(1 - \gamma - \nu_2)(s_1 - s^{SO})}{n}\nu_2^{t-1} \ , \tag{A.44a}$$

$$\tilde{s}_t = s^{SO} + (s_1 - s^{SO})\nu_2^{t-1} \ , \tag{A.44b}$$

where ν_2 denotes the smaller eigenvalue of the Jacobian (A.37)

$$\nu_2 = \frac{J_{11} + J_{22} - \sqrt{(J_{11} + J_{22})^2 - 4(J_{11}J_{22} - J_{12}J_{21})}}{2} \ . \tag{A.45}$$

Evaluating ν_2 for f_0^{SB} yields equation (27a). Comparing equation (A.44b) for the greenhouse gas stock in the second-best sustainable RS with the corresponding equation (A.11b) of the social global optimum yields equation (26), where (27b) states the corresponding eigenvalue λ_2 in the social global optimum.

It is easy to see that $\nu_2 > \lambda_2$ for $n > 1$ which implies (28). $\qquad \square$

Proof of Proposition 7

(i) A treaty with repeated payments $f_t(\Delta)$ implements the same solution as the first-best sustainable RS with initial payments $f_0^\star$ if the same refund $R_t^\star$ is distributed among the countries in period t. At the end of each Δ-th period, the fund is reduced to zero and will be refilled by $f_{t+\Delta}(\Delta)$. Hence, it must hold that

$$0 = f_{t+\Delta} = (1+\rho)f_{t+\Delta-1} - R_{t+\Delta}^\star \ . \tag{A.46}$$

Inserting recursively back to period t and solving for f_t, we obtain the amount of money that should be in the fund in period t, which equals (30).

(ii) A treaty with repeated payments $f(\Delta)$ implements the same solution as the second-best sustainable RS with initial payments f_0^{SB} if all countries receive $(\rho f_0^{SB})/n$ from the fund in each period. At the end of each Δ-th period, the fund is reduced to zero and will be refilled by $f(\Delta)$. At the beginning of each period, the remainder of the fund is invested and earns interest ρ. Therefore it holds that

$$0 = f(\Delta t)(1+\rho)^{\Delta t} - (1 + (1+\rho) + \ldots + (1+\rho)^{\Delta-1})\rho f_0^{SB} \ . \tag{A.47}$$

Applying the formula for the finite geometric series

$$\sum_{k=0}^{\Delta-1}(1+\rho)^k = \frac{1 - (1+\rho)^\Delta}{1 - (1+\rho)} \ , \tag{A.48}$$

we obtain (31).

$\square$

References

J. Andreoni. Toward a theory of charitable fund-raising. *Journal of Political Economy*, 106(6): 1186–1213, 1998.

G. B. Asheim, C. B. Froyn, J. Hovi, and F. C. Menz. Regional versus global cooperation for climate control. *Journal of Environmental Economics and Management*, 51:93–109, 2006.

S. Barrett. Self-enforcing international environmental agreements. *Oxford Economic Papers*, 46: 878–894, 1994.

S. Barrett. A theory of full international cooperation. *Journal of Theoretical Politics*, 11:519–41, 1999.

S. Barrett. *Environment and Statecraft: The Strategy of Environmental Treaty-Making.* Oxford University Press, Oxford, 2003.

F. Bloch. *in: New directions in the economic theory of the environment.* Cambridge University Press, Cambridge, 1997.

C. Carraro and D. Siniscalco. The international dimension of environmental policy. *European Economic Review*, 36(2-3):379–387, 1992.

C. Carraro and D. Siniscalco. Strategies for the international protection of the environment. *Journal of Public Economics*, 52(3):309–328, 1993.

P. Chander and H. Tulkens. Theoretical foundations of negotiations and cost sharing in transfrontier pollution problems. *European Economic Review*, 36:388–398, 1992.

I. Falk and R. Mendelsohn. The economics of controlling stock pollutants: An efficient strategy for greenhouse gases. *Journal of Environmental Economics and Management*, 25:76–88, 1993.

M. Finus. *Game Theory and International Environmental Cooperation.* Edward Elgar, Cheltenham UK, 2001.

M. Finus and B. Rundshagen. Membership rules and stability of coalition structures in positive externality games. *Social Choice and Welfare*, 32:389–406, 2009.

H. Gersbach. The global refunding system and climate change. mimeo, University of Heidelberg, 2005.

H. Gersbach and R. Winkler. On the design of global refunding and climate change. Discussion Paper DP6379, Center of Economic Policy Research, 2007.

L. H. Goulder and K. Mathai. Optimal CO_2 abatement in the presence of induced technological change. *Journal of Environmental Economics and Management*, 39:1–38, 2000.

M. Hoel. International environmental conventions: the case of uniform reductions of emissions. *Environmental and Resource Economics*, 2:141–159, 1992.

A. Lange. Providing public goods in two steps. *Economics Letters*, 91(2):173–178, 2006.

B. Lockwood. Uniqueness of Markov-perfect equilibrium in infinite-time affine-quadratic differential games. *Journal of Economic Dynamics and Control*, 20:751–65, 1996.

W. D. Nordhaus. *Managing the Global Commons: The Economics of the Greenhouse Effect.* MIT, Cambridge, MA, 1994.

C. Rowat. Non-linear strategies in a linear quadratic differential game. *Journal of Economic Dynamics and Control*, 31:3179–202, 2007.

S. Tsutsui and K. Mino. Nonlinear strategies in dynamic duopolistic competition with sticky prices. *Journal of Economic Theory*, 52:136–61, 1990.

H. Tulkens. *in: Communication and Control in Society*. Gordon and Breach, New York, 1979.

H. R. Varian. Sequential provision of public goods. Public economics, EconWPA, 1994.

S-S. Yi. Stable coalition structures with externalities. *Games and Economic Behavior*, 20:201–237, 1997.