Working Paper 9/2012

The Trade-off between Intra- and Intergenerational Equity in Climate Policy: Can Carbon Leakage be Justified?

Snorre Kverndokk, Eric Nævdal and Linda Nøstbakken

The CREE Centre acknowledges financial support from The Research Council of Norway, University of Oslo and user partners.
ISBN: 978-82-7988-122-3
ISSN: 1892-9680
http://cree.uio.no

August 2012

The Trade-off between Intra- and Intergenerational Equity in Climate Policy: Can Carbon Leakage be Justified?[1]

by

Snorre Kverndokk[2], Eric Nævdal[3] and Linda Nøstbakken[4]

Abstract

This paper focuses on two equity aspects of climate policy, intra- and intergenerational equity, and analyzes the implications of equity preferences on climate policy, and on the production and consumption patterns in rich and poor countries. We develop a dynamic two-region model, in which each region suffers from local pollution and global warming, but also has an inequality aversion over current consumption allocations. Inequality aversion over consumption lifts the consumption path of the poor region, while it lowers the consumption path of the rich region, which must take a greater share of the climate burden. Therefore, a high abatement in the rich region is met by more pollution in the poor region, thus justifying carbon leakages. Moreover, the poor region may even be allowed to increase emissions relative to business as usual under the optimal climate policy. These effects are reinforced when introducing transfers between the regions. However, loans to poor countries to reduce inequality may result in a debt crisis, and debt remittance may be part of the optimal climate policy.

JEL codes: C63, D31, D63, Q54.

Keywords: Intragenerational equity; intergenerational equity; inequality aversion; climate policy; economic development; carbon leakage; international transfers; debt crisis

---

[1] This paper is funded by the MILJØ2015 program at the Research Council of Norway. We have benefitted from discussions with Geir Asheim, Reyer Gerlagh, Johan Eyckmans, Samuel Fankhauser, Haifang Huang, Itziar Lazkano and Emilson C.D. Silva, in addition to comments from participants at the SURED conference in June 2012. The authors are associated with CREE - the Oslo Centre for Research on Environmentally Friendly Energy - which is supported by the Research Council of Norway.

[2] Corresponding author. The Ragnar Frisch Centre for Economic Research, Gaustadallèen 21, 0349 Oslo, Norway. E-mail: snorre.kverndokk@frisch.uio.no.

[3] The Ragnar Frisch Centre for Economic Research, Gaustadallèen 21, 0349 Oslo, Norway. E-mail: eric.navdal@frisch.uio.no.

[4] Alberta School of Business, University of Alberta, 2-32D Business, Edmonton, Alberta, Canada T6G 2R6. E-mail: linda.nostbakken@ualberta.ca.

1. Introduction

While climate change has been recognized as a threat to the future by most scientists and politicians for many years, there is still an ongoing debate as to what to do about it. Researchers may not agree on the optimal emission reductions, even if they agree on the natural science background, the impacts and the costs of abating greenhouse gas emissions. To a large extent, the reason for this is that optimal emission reductions depend on equity issues, and how we discount the future climate impacts is particularly important.[5] However, ethical issues have not been fully explored in economic analyses, as greenhouse gas abatement not only affects the welfare distribution between present and future generations, but also the distribution within a generation such as that between rich and poor countries. These two equity aspects are important when studying optimal emissions reductions, and as we explain below, they may work in different directions. The purpose of the current study is to analyze the trade-off between the two dimensions of equity in climate policy.

The two dimensions of equity in climate change polices can be referred to as *intra-* and *intergenerational*. The first is primarily about how we should distribute the burdens within a generation, either within the generation living today or in the future, see Kverndokk and Rose (2008). Two examples of this can be: who would suffer from climate change (inaction), and how should the burdens of mitigation (action) be distributed? In the years to come, the world may face large climatic changes such as increased temperatures, sea level rise, changed wind and precipitation patterns, more extreme weather, etc. (IPCC, 2007a). Nevertheless, the damage associated with climate change will not be evenly distributed among countries or within a given country. Studies such as those by Tol et al. (2000), Tol (2002a,b) and Yohe et al. (2007) show that some sectors will lose from climate change while others will benefit, with the poorer countries likely to face relatively stronger negative impacts than the richer countries. Several economic studies also reveal that the costs of action will vary among countries and sectors, and that it is generally more expensive to abate the more energy efficient the economy is (IPCC, 2007b). Policy instruments implemented to reduce greenhouse gas emissions will impose different burdens on people, and economic instruments such as carbon taxes will often be regressive, i.e. the burden will be the highest for the poorest (see, e.g. Bye et al., 2002).

[5] See Dasgupta (2008) for an overview and Anthoff and Tol (2009) for illustrations.

While intragenerational equity is important, most of the equity debate on climate change issues in the economic literature has been on intergenerational equity issues, i.e. distribution across generations, focusing on how large emissions reductions we should be aiming for, or how large the atmospheric greenhouse gas concentration or global mean temperature ceiling should be. This affects the distribution of burdens between the current generation and future generations, as the burdens of mitigation - the costs - have to be taken by the present generation, while the benefits of mitigation will be felt by future generations.

There are several reasons for extensive mitigation today such as attitudes toward risk as well as concerns about catastrophic events (Weitzman, 2007a), although most of the discussions have been about the appropriate discount rate to use for climate change policy decisions, as the optimal level of abatement is very sensitive to the choice of discount rate. Discount rates are weights put on the future benefits of climate change policies in order to compare them to the present costs. If we measure the costs and benefits in consumption units, the main reasons for discounting are that we may treat different generations differently (the pure rate of time preferences), and that the benefit of a consumption unit differs depending on the consumption level. The second argument is that a high level of consumption gives a low marginal utility of an additional unit, represented by the elasticity of the marginal utility of consumption. Thus, a high consumption level for future generations may be an argument for paying less attention to these generations. The consumption discount rate used in economic analyses combines these two arguments, which both represent ethical choices.

The choice of the appropriate discount rate has been a controversial issue for many years. For instance, the Stern Review (Stern, 2007) used a quite low consumption discount rate, and therefore found a high level of optimal abatement compared to other studies such as those of Nordhaus (1993) and Nordhaus and Boyer (2000). This created a lively debate in which, e.g. Nordhaus (2007) and Weitzman (2007b) argued against the choice of the low discount rate based on the observed values of the long-run return to capital. However, Heal (2009) shows that equality between the consumption discount rate and the return to capital only holds under idealized conditions such as a one-good model and perfect market assumptions including no external effects, which obviously do not hold in the presence of climate change. The one-good model has also been criticized by Hoel and Sterner (2007) and Sterner and Persson (2008), who pointed out that if there is consumption of an environmental good in addition to a produced consumption good, the environmental good will become more scarce over time relative to the produced good as we experience more climate change, which could be an argument for using a low and even negative consumption discount rate for environmental services. Lastly, Nævdal and Vislie (2008) argue that the discount rate has a minor impact on the stabilization target to avoid climate catastrophes, and therefore, a large impact on future generations.

Most studies focus on either an intra- or intergenerational distributive justice problem, thereby implicitly assuming that these problems are autonomous, i.e. that they can be treated separately. However, choices that affect intergenerational distribution also impact on the intragenerational distribution between rich and poor countries. As Heal (2009) points out, there are two ways that preferences for equality affect the choice of climate action. First, a high elasticity of the marginal utility of consumption will lead to less aggressive action if we believe that consumption is growing over time. The reason for is that this makes future generations richer, and if we care about inequality between the present and future generations, we place a lower value on the richer future generations (intergenerational equity), but there is also another effect. The rich countries are primarily responsible for the aggregate level of greenhouse gases in the atmosphere, while as mentioned above, the poor countries are likely to suffer the most from climate change. Hence, if we put a low weight on future outcomes, climate change is more likely to occur, and poor countries may suffer more (intragenerational equity). Consequently, the gap between the welfare levels of the rich and the poor may be higher, and based on the latter effect, stronger preferences for equality should go in the direction of more action to help prevent climate change.

These two effects of inequality aversion go in different directions, and the impacts of stronger preferences for equity on the level of greenhouse gas abatement are ambiguous. The problem is that the global models used to determine the optimal level of greenhouse gas emissions only take the first effect into account (intergenerational), meaning that stronger preferences for equality in these models actually induce low abatement as found in, e.g. Nordhaus and Boyer (2000).

As previously mentioned, economic analyses have mainly concentrated on separately studying the two aforementioned equity aspects. One exception to this is Baumgärtner et al. (2012), who provide a general discussion about economic analysis and the trade-offs between inter- and intragenerational justice, while Glotzbach and Baumgärtner (2012) give an analysis of the relationship between these two justice aspects when it comes to ecosystem management. Nonetheless, we are not aware of any studies that take both inequality aversions into account when finding the optimal climate policy and this paper closes that gap.

We ask the following question: How should climate policies be designed when policy makers have preferences for both intra- and intergenerational equity? A simple model of two regions, one rich and one poor, is set up to explicitly take into account preferences for the two aspects of equality before finding the optimal climate policy. The intergenerational aspect is represented by the trade-offs between welfare in the present and future generations due to the impact of global warming, while the intragenerational equity concern in this setup is purely a developmental issue, i.e. the consumption level of the poor compared to the rich. For this reason, we do not study the vulnerability to climate damage between the rich and poor countries. Our main finding is that the preferences for intragenerational equality will go in the direction of reduced consumption in the rich region and increased consumption in the poor region. The optimal climate policy subsequently gives the highest abatement burden to the rich world, which again has impacts for pollution as more dirty production is moved to the poor world, therefore justifying carbon leakage.[6]

This paper is organized in the following way. In the next section, we study the optimal climate policy contract in the presence of preferences for both intra- and intergenerational equity, while Section 3 compares this outcome with a situation when a social contract is not possible. Interactions between the regions such as transfers and international trade are introduced in Section 4 to see how this affects the social contract, and our results are illustrated with some numerical simulations in Section 5. The final section concludes.

2. Deciding on a Social Contract: A Model of Inequality Aversion

We start by considering the optimal global climate contract. One way to think about this is to follow Rawls (1971) and to assume a *veil of ignorance* situation, in which representative individuals from prospective countries around the world meet to design the future organization of the society, the original position. All individuals know that some countries are going to be rich and some are going to be poor, and that economic activity will result in

[6] The term carbon leakage is often used when there is a limited participation in climate treaties (see e.g. Felder and Rutherford, 1993). In this paper, we use the term when emission reductions in one region are met by emission increases in another region also within a climate treaty. Alternatively, the term pollution haven can be used in this case.

greenhouse gas emissions that will have a negative impact on them in terms of climate change. It is also known that mitigation today may affect developing countries more negatively than industrialized countries as it may harm economic growth and their possibility to develop. The representative individuals are concerned about consumption inequality, but they also have preferences over the welfare of future generations compared to that of the present generation. However, they do not know if they are going to live in a rich or poor country, or if they are going to live in a time with a high or low global environmental quality produced by the concentration of greenhouse gases in the atmosphere.

Behind the veil of ignorance, the representative individuals agree on the social planner's decision problem. Since climate change is a global problem, the representative individuals or countries would not have an incentive to pursue the global social optimum solution without a contract. Thus, they commit to follow the policy that comes out of the social planner's optimization problem, as this will work as insurance for all.

We take a consequentialist standpoint and consider the welfare of the individuals as the final aim of its actions. Based on this, a good society is one that maximizes a social welfare function.

2.1 The Basics of the Model

To formalize this, consider two regions *n* and *s*, where n denotes the developed region (North) and s the developing region (South). The welfare of a representative consumer/country in region $r = n,s$ at time t is:

$$U_{r,t} = u(c_{r,t}, D_{r,t}, S_t) - \alpha_r \max\left(c_{i,t} - c_{r,t}, 0\right) - \beta_r \max\left(c_{r,t} - c_{i,t}, 0\right),\ r, i = n, s,\ r \neq i. \quad (1)$$

c is consumption, D is local pollution and S is the state of the global environment, whereas i represents the other representative consumer/region. $u(c, D, S)$ is a standard utility function, which is increasing in *c* and *S*, but falling in *D*, and with: $\frac{\partial u(c,D,S)}{\partial c} \to \infty \ for\ c \to 0$, $\frac{\partial^2 u(c,D,S)}{\partial c\ \partial D} \leq 0$ and $\frac{\partial^2 u(c,D,S)}{\partial D^2} \leq 0$. The latter condition implies that local pollution becomes increasingly damaging as the pollution level rises. Note that in this simple model, we do not consider different degrees of vulnerability to climate change within the two regions.

Preferences for equality are modeled as inequality aversion following Fehr and Schmidt (1999), and based on this, consumers have an inequality aversion in such a way that they dislike having higher consumption than others, although they dislike even more having a lower consumption than other consumers that they compare themselves to. This streamlines the economic development perspective as the intragenerational aspect, while, in contrast, the climate change perspective is the intergenerational aspect in this model. The Fehr and Schmidt framework has primarily been used in describing preferences for income equality among individuals, but may also be useful as a description of the social preferences of policy makers in different regions, so long as the transfers between regions are not assumed to be due to strategic reasons only.[7]

Following this, let α be a parameter representing the negative feeling of being worse off than others, while β is the parameter representing the negative feeling of being better off. We then have that $\alpha \geq \beta$. We ignore strategic interactions by assuming that each region, both North and South, consists of many identical countries that do not have any market power and cannot individually affect the overall global environmental quality.[8]

Without a loss of generality, let us assume that the population sizes of the two regions are equal and normalized to unity. Therefore, c_r is also per capita consumption in region *r*.

Furthermore, we model each representative country's production of an aggregate good, $Y_{r,t}$, with two types of inputs, clean and dirty, both of which are perfect substitutes:[9]

$$Y_{r,t} = Y_{r,c,t} + Y_{r,d,t} \tag{2}$$

[7] Other alternative social preferences could be used, but this is not crucial to our conclusions. One example is Charness-Rabin preferences applied by Kolstad (2011) to study coalitions in public goods provision.
[8] One could also argue that the rich world not only has preferences for the equality of consumption, but also for the total welfare of the poor world such as how climate change affects poor societies. One example of this can be that the rich world feels responsible for a large part of historical greenhouse gas emissions, which also has a negative effect on the welfare of the poor. If the utility function was region specific, this could be thought of as being embodied in the utility function. However, as mentioned above, we have chosen to focus on the economic development perspective of intragenerational equity.
[9] One example can be electricity.

$$Y_{r,c,t} = A_{c,t} K_{r,c,t}^{\gamma_c} \tag{3}$$

$$Y_{r,d,t} = A_{d,t} K_{r,d,t}^{\gamma_d} \tag{4}$$

where subscripts c and d denote clean and dirty. Hence, the clean input is produced using clean (green) capital, while the dirty input is produced using dirty (brown) capital. We assume that $\gamma_j \in \langle 0,1]$ for $j = c,d$, which implies diminishing marginal returns. The productivity of dirty capital is higher than for clean capital, i.e. $A_{d,t} > A_{c,t}$ for all t. For simplicity, we assume these total factor productivities to be constant. However, within this framework, we can easily introduce technological development as a positive trend in the total factor productivities.

We assume that the capital stock of dirty capital is initially higher in the North than in the South, i.e. $K_{n,d,1} > K_{s,d,1}$. Even so, a country can invest in clean and dirty capital, with capital dynamics given by:

$$K_{r,j,t+1} = \left(1-\delta_j\right) K_{r,j,t} + I_{r,j,t}, \ \ r = n,s, \ j = c,d, \tag{5}$$

where δ_j is the capital depreciation rate.

The resource constraints are:[10]

$$Y_{r,t} = c_{r,t} + I_{r,c,t} + I_{r,d,t}, \quad r = n,s \tag{6}$$

By substituting for investment, $I_{r,c,t}$ and $I_{r,d,t}$, using equation (5), the resource constraint for a country in region r can be rewritten as:

$$Y_{r,t} = \left[K_{r,c,t+1} - \left(1-\delta_c\right) K_{r,c,t}\right] + \left[K_{r,d,t+1} - \left(1-\delta_d\right) K_{r,d,t}\right] + c_{r,t}\,. \tag{7}$$

[10] Note that we have not explicitly modeled markets here, which can be justified by a sequence of spot markets that are renewed across generations.

Similarly to Silva and Zhu (2009), we assume that there is a co-production of local and global pollution from the use of dirty capital.[11] As a result, these pollutants are correlated. Assume that κ_l is a coefficient that reflects the local pollution per unit of dirty capital used, while κ_g reflects the global pollution (greenhouse gas emissions) per unit of dirty capital. For local pollution, we then have:

$$D_{r,t} = \kappa_l K_{r,d,t}, \;\; r = n, s. \tag{8}$$

The global environment is modeled as a stock variable that deteriorates with global pollution following from the aggregate use of the dirty capital over the two regions, and regenerates naturally at a rate σ:

$$S_{t+1} = \sigma \overline{S} + (1-\sigma) S_t - \kappa_g \sum_r K_{r,d,t} \tag{9}$$

The global environmental quality is constrained as follows: $S_t \in \left[0, \overline{S}\right]$, where $\overline{S}$ is the quality in absence of pollution. Note that in this case, *S* will converge asymptotically to $\overline{S}$. We therefore treat climate change as a reversible process in the very long run.

Lastly, we assume that consumption is initially larger in the North than in the South, i.e. $c_{n,1} > c_{s,1}$, since the North is endowed with more initial productive capital than the South. Because of this, we can write the welfare functions of the two regions at time *t = 1* as:

$$U_{n,1} = u\left(c_{n,1}, D_{n,1}, S_1\right) - \beta_n \left(c_{n,1} - c_{s,1}\right) \tag{10}$$

$$U_{s,1} = u\left(c_{s,1}, D_{s,1}, S_1\right) - \alpha_s \left(c_{n,1} - c_{s,1}\right) \tag{11}$$

[11] Silva and Zhu (2009) assume that pollution follows from the production of the dirty good. In our model, it follows from the use of dirty capital. In the long run (steady state), however, there is a constant relationship between production and the capital stock.

2.2 The Social Contract

Within this modeling framework, and with a social planner who determines production, consumption and capital investment in every country (region), consumption in the South will never exceed consumption in the North (per capita). Hence, we use the simplified welfare functions (10) and (11) also for *t > 1*.

The social planner seeks to maximize the sum of discounted welfare across regions, where welfare in period t is defined as:

$$W_t = u(c_{n,t}, D_{n,t}, S_t) + u(c_{s,t}, D_{s,t}, S_t) - (\beta_n + \alpha_s)(c_{n,t} - c_{s,t}) \tag{12}$$

The maximization problem can then be expressed as:

$$\max_{\{c_{nt}, c_{st}, K_{nct+1}, K_{sct+1}, K_{ndt+1}, K_{sdt+1}, S_{t+1}\}} \sum_{t=0}^{\infty} \rho^t \left(\sum_r u(c_{r,t}, D_{r,t}, S_t) - \phi(c_{n,t} - c_{s,t}) \right), \tag{13}$$

subject to (2)-(4) and (7)-(9). ρ is the time preference rate and $\phi \equiv (\beta_n + \alpha_s)$, which is a constant.

The Lagrangian of the maximization problem can now be expressed as follows:

$$\begin{aligned} L_{so} = \sum_{t=1}^{\infty} \rho^{t-1} \Bigg\{ & \sum_r u(c_{r,t}, D_{r,t}, S_t) - \phi(c_{n,t} - c_{s,t}) \\ & + \sum_r \lambda_{r,t} \left[Y_{r,t}(K_{r,c,t}, K_{r,d,t}) - K_{r,c,t+1} + (1-\delta_c) K_{r,c,t} - K_{r,d,t+1} + (1-\delta_d) K_{r,d,t} - c_{r,t} \right] \\ & + \mu_t \left[S_{t+1} - \sigma \bar{S} - (1-\sigma) S_t + \kappa_g \sum_r K_{r,d,t} \right] \Bigg\} \end{aligned} \tag{14}$$

where λ is the shadow price on capital, while μ is the shadow price on the pollution stock.

First order conditions include:

$$[c_{n,t}]: \qquad \frac{\partial u(c_{n,t}, D_{n,t}, S_t)}{\partial c_{n,t}} - \phi = \lambda_{n,t} \qquad (15)$$

$$[c_{s,t}]: \qquad \frac{\partial u(c_{s,t}, D_{s,t}, S_t)}{\partial c_{s,t}} + \phi = \lambda_{s,t} \qquad (16)$$

$$[K_{r,d,t+1}]: \kappa_l \frac{\partial u(c_{r,t+1}, D_{r,t+1}, S_{t+1})}{\partial D_{r,t+1}} + \lambda_{r,t+1}\left[\frac{\partial Y_{r,t+1}}{\partial K_{r,d,t+1}} + 1 - \delta_d\right] + \mu_{t+1}\kappa_g = \rho^{-1}\lambda_{r,t}, \ r = n, s \qquad (17)$$

$$[K_{r,c,t+1}]: \qquad \lambda_{r,t+1}\left[\frac{\partial Y_{r,t+1}}{\partial K_{r,c,t+1}} + 1 - \delta_c\right] = \rho^{-1}\lambda_{r,t}, \ r = n, s \qquad (18)$$

$$[S_{t+1}]: \qquad \sum_r \frac{\partial u(c_{r,t+1}, D_{r,t+1}, S_{t+1})}{\partial S_{t+1}} = \mu_{t+1}(1-\sigma) - \rho^{-1}\mu_t \qquad (19)$$

Clearly, an important question when analyzing the optimal solution is whether consumption in the two regions will converge to the same level in the long run. In Appendix 1, we show that this will be the case, which gives us the following Lemma:

Lemma 1: *In the long run, the socially optimal consumption levels of North and South will converge.*

Note that this result is independent of inequality aversion in consumption, but that inequality aversion accelerates the convergence process, see discussion below.

2.3 Optimal Policy

Let us now consider how inequality aversion affects the optimal consumption and capital paths of the two regions. We start by looking at optimal consumption. From the first order conditions (equations (15) and (16)), we know that the marginal utility from goods consumption and inequality (ϕ) must equal the shadow value of capital. For the North, this implies that an additional unit of consumption today yields higher utility from consumption, but disutility from more inequality. In the South, we have the opposite effect of inequality aversion: An additional unit of consumption today increases utility, both from consumption and from reduced inequality. Consequently, introducing inequality aversion shifts the consumption path of the North downwards, while the consumption path of the South shifts upwards. We summarize this result in Proposition 1:

Proposition 1: Compared to the case without inequality aversion, inequality aversion within one generation will lead to a reduction in consumption in the North and an increase in consumption in the South under the social contract.

For the North to reduce consumption and the South to increase consumption relative to the case without inequality aversion, the North invests relatively more in clean capital while the South invests relatively more in dirty capital. It follows that the North is taking a bigger hit to improve the environment, which can be thought of as an indirect transfer of welfare from North to South. This also implies that a stronger concern for equality causes the two regions' consumption levels to converge faster. We can then write Proposition 2:

Proposition 2: Inequality aversion within one generation will increase investments in clean capital and reduce investments in dirty capital in the North in the social optimum. For the South, we find the opposite result.

This proposition shows that inequality aversion within one generation may justify carbon leakage in which pollution is transferred from the North to the South. While carbon leakage is usually considered as a bad (see e.g. Felder and Rutherford, 1993; Böhringer et al., 2010), this Proposition opens the possibility that it may be the solution to balancing the trade-off between intra- and intergenerational equity.

To see how the transition works, consider the optimal capital paths. To simplify the notation, we define the following: $MP_{r,j} \equiv \frac{\partial Y_r}{\partial K_{r,j}}$ is the marginal productivity of capital j in region r, and $MU_r^x \equiv \frac{\partial u(c_r, D_r, S)}{\partial c_r}$, with $x = (c_r, D_r, S)$, is the marginal utility of consumption, local pollution, and global environmental services, respectively, in region r. By using this, eliminating common terms from the two regions' optimality conditions for clean capital (18) and rearranging, we obtain the following condition:

$$MP_{s,c,t} - MP_{n,c,t} = \rho^{-1}\left(\frac{\lambda_{s,t}}{\lambda_{s,t+1}} - \frac{\lambda_{n,t}}{\lambda_{n,t+1}}\right). \tag{20}$$

In the short run, before the two regions converge, we know that the North is richer and therefore has a higher utility than the South, although the difference between the two regions is decreasing over time. Regardless of whether the optimal steady-state policy involves a reduction or an increase in the North's utility relative to the initial state, it must be the case that $MU_{n,t}^{c} \big/ MU_{n,t+1}^{c} < MU_{s,t}^{c} \big/ MU_{s,t+1}^{c}$ until the South has caught up with the North. In addition, we know from (15) and (16) that $\lambda_{n,t} = MU_{n,t}^{c} - \phi$ and $\lambda_{s,t} = MU_{s,t}^{c} + \phi$. With this information, we can show that the following inequality must hold until the regions converge:

$$\frac{\lambda_{s,t}}{\lambda_{n,t}} > \frac{\lambda_{s,t+1}}{\lambda_{n,t+1}}. \tag{21}$$

Inequality (21), which shows that the difference in the shadow values of capital between the South and the North decreases over time, also implies that the right-hand side of (20) is positive. Hence, while the South catches up with the North, the marginal product of clean capital is higher in the South than in the North, which implies a higher stock of clean capital in the North than in the South, $K_{n,c,t} > K_{s,c,t}$.

Lastly, let us consider the optimal investment in clean and dirty capital in the long run. To do this, we rewrite the first order conditions (15)-(19) for the steady-state levels of the variables using the new notation:

$$MU_{n}^{c} - \phi = \lambda_{n} \tag{22}$$

$$MU_{s}^{c} + \phi = \lambda_{s} \tag{23}$$

$$\kappa_{l} MU_{r}^{D} + \lambda_{r}\left[MP_{r,d} + 1 - \delta_{d}\right] + \mu\kappa_{g} = \rho^{-1}\lambda_{r},\ r = n,s \tag{24}$$

$$\lambda_{r}\left[MP_{r,c} + 1 - \delta_{c}\right] = \rho^{-1}\lambda_{r},\ r = n,s \tag{25}$$

$$\sum_{r} MU_{r}^{S} = \mu\left(1 - \sigma - \rho^{-1}\right) \tag{26}$$

By substituting for the steady-state shadow prices of capital ((22)-(23)) and the environment (26) in the optimality conditions for clean and dirty capital ((24)-(25)), we obtain the following relationships that the marginal productivities must satisfy in the optimum:

$$MP_{r,d} = \delta_d + v - \kappa_l \frac{MU_r^D}{MU_r^C + \psi_r} + \left(\frac{\kappa_g}{\sigma + v}\right) \frac{MU_n^S + MU_s^S}{MU_r^C + \psi_r} \tag{27}$$

$$MP_{r,c} = \delta_c + v \tag{28}$$

where $(\psi_n, \psi_s) = (-\phi, \phi)$ captures the effect of inequality aversion for both North and South, and $v = \rho^{-1} - 1$ is the discount rate. The optimal level of clean capital requires its marginal productivity to equal the sum of the depreciation and discount rates. These terms also enter the condition for the marginal productivity of dirty capital, but when considering the optimal level of dirty capital we must also account for the welfare effects of local and global pollution. Consequently, the marginal productivity of dirty capital must equal the sum of depreciation and discount rates, the value of reduced local pollution and the value of improved global environmental quality. The two latter effects are measured in terms of increased consumption, with inequality considerations accounted for (ψ_r) as consumption levels converge. Once convergence occurs, the marginal productivity of dirty capital must be the same in all countries.

Note that every term in the optimality condition for dirty capital (27) is positive. Hence, even if the capital depreciation rates for clean and dirty capital are equal, we would require a higher marginal productivity from dirty- than clean capital to invest. As discussed above, this is because an investment in dirty capital must also compensate for the welfare effects of local and global environmental damage. Note that the inequality aversion parameters enter into the denominators of equation (27), thus partially decreasing the marginal productivity of dirty capital in the North while increasing it in the South. However, as consumption converges in the long run, and as the effect of global pollution enters equally in the utility functions, we find that the levels of dirty capital in the two countries converge.

How does inequality aversion across a generation affect greenhouse gas emissions? The steady-state condition for global environmental quality can be rearranged and expressed in terms of the shadow prices of the environment:

$$\mu = \frac{MU_n^S + MU_s^S}{1 - \sigma - \rho^{-1}} \tag{29}$$

It follows that the steady-state value of the shadow price of the global environment (μ) increases in the marginal utility of environmental quality given by the numerator in equation (29), increases in the replenishment rate of the environment and decreases in the discount factor. Equation (29) also reveals that the steady-state level of global environmental quality does not depend on the regions' preferences for equality (ϕ). Nevertheless, the path of environmental quality towards its steady-state level does depend on the regions' inequality preferences. The steady-state level of global environmental quality is given by:

$$S^* = \bar{S} - \frac{\kappa_g}{\sigma} \sum_r K_{r,d} \tag{30}$$

The environmental quality path towards the steady-state depends on the aggregate level of dirty capital in the two regions. As seen above, inequality aversion reduces emissions in the North, while increasing emissions in the South before the steady-state is reached. If the emissions reductions in the North exceed the increase in the South, the environmental quality path would shift upwards (better environment). Still, without specifying functional forms, it is not clear which of the two effects would dominate. Hence, the impact of inequality aversion on environmental quality is ambiguous, and we will return to this in the numerical analysis in Section 5.

3. But what if a Contract is Not Possible? The Business as Usual Case

The next question is what the actions of the two regions of the world would be if the social contract cannot be reached? As long as there is no enforcement mechanism in place, the regions will be better off by following their own interest and maximizing the welfare of a representative consumer. This may be called the Business as Usual problem (BAU), i.e. the optimization problem of the policy makers when there is no coordinated action or global environmental agreement.

As assumed above, consumption in the North is larger than in the South at the starting point (*t = 1*), i.e., $c_{n,1} > c_{s,1}$. Countries are identical in all respects other than consumption and capital stock levels. If the initial consumption levels follow from optimization in the two regions, the South will never have a higher consumption than the North. Therefore, we can set up the BAU optimization problem for the original model as consumers maximizing:

$$\sum_{t=1}^{t=\infty} \rho^{t-1} U_{n,t} = \sum_{t=1}^{t=\infty} \rho^{t-1} \left[u\left(c_{n,t}, D_{n,t}, S_t\right) - \beta_n \left(c_{n,t} - c_{s,t}\right) \right] \tag{31}$$

$$\sum_{t=1}^{t=\infty} \rho^{t-1} U_{s,t} = \sum_{t=1}^{t=\infty} \rho^{t-1} \left[u\left(c_{s,t}, D_{s,t}, S_t\right) - \alpha_s \left(c_{n,t} - c_{s,t}\right) \right] \tag{32}$$

respectively, given the technology (2)-(4), the resource constraint (7) and the local environmental constraints (8). The time preference rate, ρ, is set equal for both regions in order to not have the effects of inequality aversion confounded by the effects of discounting.

To find the non-cooperative Nash equilibria for the two regions, we first define the Lagrangians. For the North, the Lagrangian is:

$$\begin{aligned} \mathrm{L}_{BAU} = \sum_{t=1}^{\infty} \rho^{t-1} \Big\{ & \left[u(c_{n,t}, D_{n,t}, S_t) - \beta_n \left(c_{n,t} - c_{s,t}\right) \right] \\ & + \lambda_{n,t} \left[Y_{n,t}(K_{n,c,t}, K_{n,d,t}) - K_{n,c,t+1} + \left(1-\delta_c\right) K_{n,c,t} - K_{n,d,t+1} + \left(1-\delta_d\right) K_{n,d,t} - c_{n,t} \right] \Big\} \end{aligned} \tag{33}$$

where $\lambda_{n,t} > 0$ is the shadow price on capital. Recall that we made the assumption of each representative country being so small that its impact on the dynamics of global environmental quality is approximately zero. As a result, (33) is maximized over consumption and dirty and clean capital stocks, in which the local environment is taken into account, though global environmental quality is taken as given.

The first order conditions can be written as:

$$[c_{n,t}]: \qquad MU_n^c - \beta_n = \lambda_{n,t} \tag{34}$$

$$[K_{r,c,t+1}]: \qquad \lambda_{n,t} = \rho\lambda_{n,t+1}\left[MP_{n,c,t+1} + 1 - \delta_c\right] \tag{35}$$

$$[K_{r,d,t+1}]: \qquad \lambda_{n,t} = \rho\left[MU^D_{n,t+1}\kappa_l + \lambda_{n,t+1}\left(MP_{n,d,t+1} + 1 - \delta_d\right)\right] \tag{36}$$

The first order conditions for the South are similar apart from the determination of consumption, which is found from:

$$[c_{s,t}]: \qquad MU^c_{s,t} + \alpha_s = \lambda_{s,t} \tag{37}$$

Equations (34) and (37) show that in the optimum, the present marginal benefit from increased consumption should equal the opportunity cost of increasing the consumption, i.e. the shadow price of capital. Compared to standard preferences, i.e. without any inequality aversion ($\alpha_r = \beta_r = 0$, $r = n,s$), the marginal benefits of consumption is lower in the North, while being higher in the South, thereby indicating a lower level of consumption in the North and a higher level of consumption in the South. This gives us Proposition 3:

Proposition 3: *In business-as-usual, inequality aversion within one generation will lead to reduced consumption in the North and higher consumption in the South compared to standard preferences.*

Let us now compare the optimal consumption levels under BAU to the socially optimal levels by comparing equations (15) and (16) to equations (34) and (37). The difference in the optimality conditions for consumption is that each country under the social contract takes into account both countries' disutility from inequality, and not just its own disutility, i.e. it takes into account the consumption externality that it imposes on the other region. Hence, the North reduces- and the South increases consumption more under the social contract than in the business-as-usual case, with all else being equal. In addition to this, the regions do not take into account their impact on the global environment, which goes in the direction of lower consumption for both regions in the social optimum compared to BAU. For the North, both incentives mean a lower consumption level under the social contract compared to BAU, but for the South we have opposing incentives. We summarize this result in Proposition 4:

<u>Proposition 4:</u> *Inequality aversion within one generation will lead to a larger reduction in consumption in the North under the social contract than in business-as-usual, while for consumption in the South, the result is ambiguous.*

Equations (35) and (36) determine the investments in clean and dirty capital, respectively. The interpretations of these equations are that the benefits of increased capital in period *t* should equal the social costs of the capital investments. For both types of capital, the cost of investment is lower consumption in the current period, while for the dirty capital the cost also includes lower utility due to a worsening of the local environment.

Using equations (35) and (36), we can eliminate $\lambda_{n,t+1}$ and ρ ,which yields the following relationship between clean and dirty capital in region $r = n, s$:

$$MP_{r,d,t} - MP_{r,c,t} = \delta_d - \delta_c - \lambda_{r,t}^{-1} MU_r^D \kappa_l . \quad (38)$$

Equation (38) states that the country should adjust its stock of clean and dirty capital so that the difference in the marginal productivities between them equals the difference in capital depreciation rates and the extra social cost of using dirty capital due to the deterioration of the local environment. Note that the last term in equation (38) is positive. Consequently, if clean- and dirty capital depreciate at the same rate, the country should adjust its capital stocks so that the marginal productivity of dirty capital is higher than that of clean capital.

How would inequality aversion affect investments in clean and dirty capital in the two regions, respectively? Consider the North first, as in this region inequality aversion leads to lower consumption, which will further lead to lower investments in dirty capital and higher investments in clean capital. The intuition behind this is that a lower consumption would not require an investment in the most effective capital (dirty), while an investment in clean capital will reduce the emissions of local pollutants, thereby improving the environment. A better local environment will increase utility and therefore work as a substitute to lowering consumption in the utility function in the North. In the South, the intuition works the other way around and investments in dirty capital will increase, while investments in clean capital will be lower. Thus, inequality aversion still creates carbon leakage from North to South.

Let us next conduct the same comparison between the social contract and BAU for regional capital stocks as we did for consumption levels, which involves comparing optimality conditions (17) and (18) to conditions (35) and (36). The optimality condition for clean capital is the same in the two cases. However, the optimality condition for dirty capital under the social contract includes an additional term compared to the BAU case. This term represents the marginal effect of more dirty capital on global environmental quality, with the inclusion of this term in the social contract case implying a lower investment in dirty capital compared to BAU since the environmental externality is internalized.

An interesting conclusion from inequality aversion is that while the stock of dirty capital will be lower for the North under the social contract than in BAU, we cannot tell if the stock should be lower for the South under the social contract. For this reason, the implication is that poor countries should not necessarily have a pollution constraint under a global climate treaty. Under certain conditions it may actually be optimal that poor countries increase their emissions under a climate treaty. This gives us Proposition 5:

<u>Proposition 5:</u> *Inequality aversion within one generation will lead to a reduction in emissions in the North under the social contract compared to business-as-usual, while for the South the result is ambiguous, and emissions may actually be higher under the social contract.*

4. Interactions between the Regions

Let us now turn once again to the social contract. In the model above, the only interaction between the two regions came via the impact of pollution on the global environment. We now introduce international transfers (such as enhancing development in the poor world) and international trade to see how this changes the conclusions.

4.1 International Transfers

First consider the implications of allowing for unlimited international transfers between the two regions.

Transfers between the two regions can easily be introduced into the model presented above by adding the term τ_t, which represents transfers from North to South in period t to the regions' resource constraints. The two regions' modified resource constraints then become:

$$Y_{n,t}(K_{n,c,t},K_{n,d,t}) = c_{n,t} + \tau_t + K_{n,c,t+1} - (1-\delta_c)K_{n,c,t} + K_{n,d,t+1} - (1-\delta_d)K_{n,d,t} \tag{39}$$

$$Y_{s,t}(K_{s,c,t},K_{s,d,t}) = c_{s,t} - \tau_t + K_{s,c,t+1} - (1-\delta_c)K_{s,c,t} + K_{s,d,t+1} - (1-\delta_d)K_{s,d,t} \tag{40}$$

In addition, we introduce a constraint on North-South transfers such that $\tau_t \leq M$ in every time period. The social planner must now also determine the size of the transfers τ_t; therefore, we obtain an additional first order condition for this variable. The first order conditions for the modified optimization problem include the optimality conditions stated above for the case without transfers ((15)-(19)), along with the following condition for the North-South transfer:

$$\lambda_{n,t} + \omega_t = \lambda_{s,t}, \tag{41}$$

where $\omega_t \geq 0$ is the shadow price on the transfer constraint ($\tau_t \leq M$).

The optimality conditions show that the optimal transfer policy is a most rapid approach path towards equality. In the case of unlimited transfers, which means that $\omega_t = 0$ for all t in (41), there should be a transfer from North to South in the first period that completely eliminates inequality. From the second period onwards, all countries are equal and there is no disutility from inequality. If the transfer constraint is binding, condition (41) shows that the shadow price in the North is higher than the shadow price in the South. In this limited transfer case, the optimal policy is to transfer as much as possible from North to South ($\tau_t = M$) until the consumption paths converge.

The most rapid approach path to equality is the optimal policy regardless of whether we account for inequality aversion. The result that consumption levels should be equalized as soon as possible is also evident if we combine condition (41) with the optimality conditions for consumption in both the North and South.

Allowing for transfers from North to South enables the social planner to eliminate inequality sooner. Furthermore, the social planner seeks to eliminate inequality between the two regions regardless of whether the countries have inequality preferences, since the best use of available resources requires that marginal productivities are equal across all countries at all times. Hence, it is desirable to have the same levels of clean and dirty capital in both regions.

This result also supports the conclusion of higher consumption and more dirty production in the South. Nonetheless, note that in this case the carbon leakage from North to South will follow independently of the inequality aversion.

We summarize the results on transfers in Proposition 6:

Proposition 6: *Transfers between the regions will enhance the convergence process under the social contract. Transfers from North to South should follow the most rapid approach path to equality between the regions. As a result, transfers increase pollution in the South and decrease pollution in the North.*

Figure 1 illustrates the movements in the dirty capital stock. The capital levels will converge and the dirty capital stock will fall in the North, while increasing in the South. Note, however, that with technological growth, i.e. a positive trend in the total factor productivities, we cannot rule out the possibility of an increasing dirty capital stock in the North, but transfers will unambiguously increase the dirty capital stock of the South, again justifying carbon leakage.

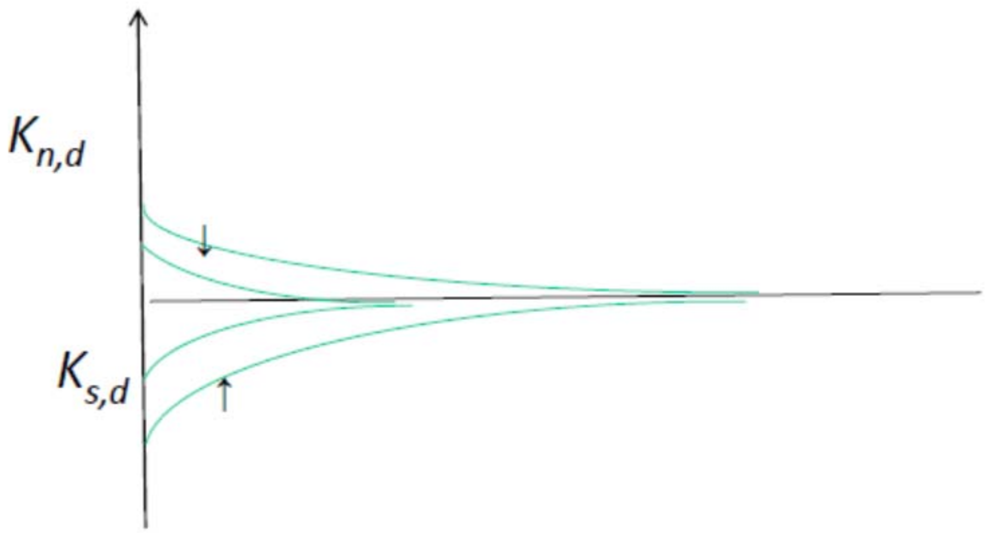

Figure 1: The paths of dirty capital with limited transfers compared to no transfers

4.2 International Trade

Thus far, we have discussed pure transfers from the North to South. Let us now see how the results change if we allow for international trade in the input factors. Define $E_{d,t}$ as export from North to South of the dirty input factor, while $E_{c,t}$ is export from North to South of the clean input factor, meaning that the South's export is equal to $-E_{d,t}$ and $-E_{c,t}$ respectively.

The available amounts of input are then:

$$Y_{n,j,t} = Y_{n,j,t}(K_{n,j,t}) - E_{j,t}, \quad j = c,d \tag{42}$$

$$Y_{s,j,t} = Y_{s,j,t}(K_{s,j,t}) + E_{j,t}, \quad j = c,d \tag{43}$$

In this simple setup, we impose market clearing without considering prices. This can be justified in the social optimum, as we do not consider any strategic interactions that may affect prices.

The temporal budget constraints are:

$$c_{n,t} = Y_{n,d,t}(K_{n,d,t}) - E_{d,t} + Y_{n,c,t}(K_{n,c,t}) - E_{c,t} - K_{n,c,t+1} + \left(1-\delta_c\right)K_{n,c,t} - K_{n,d,t+1} + \left(1-\delta_d\right)K_{n,d,t} \tag{44}$$

$$c_{s,t} = Y_{s,d,t}(K_{s,d,t}) + E_{d,t} + Y_{s,c,t}(K_{s,c,t}) + E_{c,t} - K_{s,c,t+1} + \left(1-\delta_c\right)K_{s,c,t} - K_{s,d,t+1} + \left(1-\delta_d\right)K_{s,d,t} \tag{45}$$

Lastly, we have the intertemporal trade balance:

$$\sum_{t=1}^{\infty}(1+v)^t \left(E_{d,t} + E_{c,t}\right) = 0 \tag{46}$$

where *v* is the interest rate. Note, however, that in this model the interest rate equals the time preference rate, see equation (18). This means that the international trade balance can be written as:

$$\sum_{t=1}^{\infty}\rho^{1-t}\left(E_{d,t} + E_{c,t}\right) = 0 \tag{47}$$

In a social optimum, trade is determined directly by the social planner; as a result, trade can be viewed as transfers that have to be paid back with interest.

Introducing this into our model yields the following first order condition where $\varphi > 0$ is the shadow price on the intertemporal trade balance:

$$\lambda_{s,t} - \lambda_{n,t} = \rho^{2(1-t)}\varphi \tag{48}$$

Note that without the trade balance constraint, we find that $\lambda_{n,t} = \lambda_{s,t}$, i.e. the result with unlimited transfers.

The development of the consumption paths are as follows. At time $t = 1$, we have $\lambda_{s,t} - \lambda_{n,t} = \varphi$, thus the difference between the shadow prices of capital is at its minimum and is equal to the shadow price of the trade balance, which represents the payback from South to North. However, note that with a positive discounting, the right-hand side of (48) increases exponentially and converges to ∞ as $t \to \infty$. Therefore, we obtain a *"Greek disease,"* in which it is better to equalize consumption as far as possible now and to take the costs of paying back in the future, though these costs will explode over time. Yet once again, this result reinforces our previous results with carbon leakage; an increase in dirty capital and pollution in the South that allow the South to increase its consumption level and pay back its debt. This gives us the final proposition:

<u>Proposition 7:</u> *International trade between regions in under the social contract will reduce the difference in consumption between them in the near term, but with a positive discount rate, the gap in consumption will increase over time until it converges.*

To help see this, let us start with the left-hand side of (48), which must also converge to infinity as $t \to \infty$. This means that the difference in the shadow prices of capital between the South and the North must approach infinity. We know that in the long run, the South must export goods to the North for the trade balance constraint to hold. As a consequence of this, the South's consumption level declines, while the North's consumption increases over time.

From the first order conditions for consumption (equations (15) and (16)) and the assumption $\lim_{c \to 0} MU_{r,t}^{c} = \infty$, it is clear that for the term $\left(\lambda_{s,t} - \lambda_{n,t}\right)$ to approach infinity as $t \to \infty$, consumption in the South must approach zero. Hence, in the long run, consumption approaches zero in the South, while it increases towards a finite upper bound in the North, which is illustrated in Figure 2:

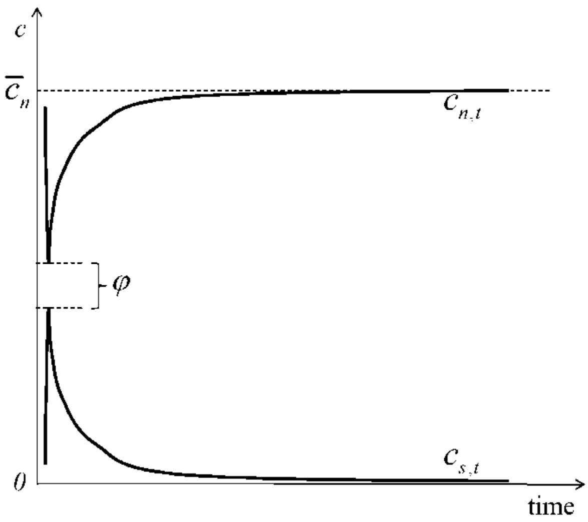

Figure 2: Greek disease (positive discount rate)

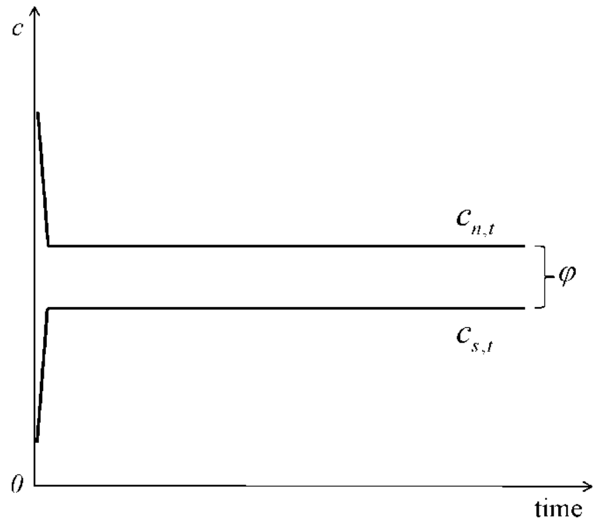

Figure 3: Greek disease (no discounting)

In the special case without discounting ($\rho = 1$), the optimality condition for trade simplifies to $\lambda_{s,t} - \lambda_{n,t} = \varphi$. As with discounting, there should be trade from North to South in the first period so that the difference between the two shadow prices of capital equals the shadow price of balanced trade. To eliminate the two regions' trade balance surpluses and deficits in the long run, the South must export goods to the North from period 2 and onward. Because of this, the two regions will never converge completely. Instead, trade allows the two regions to immediately adjust to the long-run level with some inequality, while that inequality is maintained over time to achieve balanced trade, which is illustrated in Figure 3. Note that the less trade that is initially needed to eliminate inequality, the less constraining the trade balance constraint will be, and the smaller the difference between the countries in the long run (i.e. smaller φ).

As is evident from Figures 2 and 3, the discount factor is a critical determinant of the dynamics of the two regions. Regardless of the discount factor, trade from North to South should be used to reduce inequality today. The gain from reduced inequality in the short run is balanced against the loss from inequality in the long run. Without discounting, the South consumes less than the North because of the exports needed to achieve balanced trade in the long run. As we introduce discounting, the South's consumption level will not just be lower than that of the North, it will approach zero in the long run. A higher discount rate means that we care less about future consumption and inequality, but it also implies that the South must reduce consumption more when it pays back its trade debt in the future. Thus, it is socially optimal to help the South today to eliminate inequality between the regions, even though the South will collapse in the very long run. It follows that even when people care about equality and such preferences are accounted for by a welfare maximizing social planner, trade does not ensure equality in the long run.

Clearly, the introduction of an additional constraint in the optimization problem of the social planner can only worsen the outcome. Unlimited transfers without a (binding) trade balance constraint would therefore yield higher welfare than the trade solution discussed above, which has several policy implications. First, this supports debt remittance to developing countries to help reduce inequality. The developed world can then transfer wealth to the developing world, both by remitting debt and by using more clean technology, thereby allowing poorer countries to use more productive though dirtier technology to develop faster. Eventually, however, as the inequality between different regions diminishes, the poor countries must also use more clean technologies to reduce aggregate pollution. Second, our findings imply that wealth transfers from rich to poor countries, such as development aid and foreign direct investment, can be welfare improving, particularly in the short run to help speed up the process of eliminating inequality between countries.

5. Numerical Simulations

In order to illustrate our results, a number of simulations have been performed where we examine how inequality aversion affects the social contract. In the simulations we have assumed that the North and the South have the same utility function given by:

$$U\left(c_t, D_t, S_t\right) = c_t^a \frac{1}{\left(1+D\right)^b} S_t^c \tag{49}$$

The social contract has been simulated for three values of aggregate inequality aversion $\phi = (0,1,1.5)$ over a time horizon of 150 years. As these simulations are for illustrations only, the parameter values chosen are based on best-guess, apart from the total factor productivities in the production functions (3) and (4) that are taken from Acemoglu et al. (2012). See Appendix 2 for parameter values.[12]

Consumption levels in the North and the South have been plotted in Figure 4. As expected, the consumption patterns change as inequality aversion increases. There are several things to notice about these changes. The simulations confirm that consumption levels in the North and the South converge over time, and the point of time when inequality is completely eliminated decreases in inequality aversion, ϕ. An interesting aspect of the consumption dynamics is that there is a dip in North's consumption just after $t = 1$. This dip becomes more pronounced as ϕ increases. Note that given our parametric choices, both regions choose to build down their stocks of dirty capital.[13] When $\phi = 0$, the initial dip in consumption is caused by depreciation of dirty capital before the stock of clean capital has accumulated sufficiently to compensate. This can be seen in Figure 5. As ϕ increases, it becomes more and more important to reduce the difference between North and South's consumption. This confirms Proposition 1. When $\phi = 1.5$, this implies rapid convergence between consumption levels. However, in this particular case, after an initial period of convergence consumption patterns diverge before finally converging at a later point in time. This divergence may appear puzzling, but is actually an end term effect as the simulations are run within a given time frame. As seen from Figure 5, the higher the inequality aversion, the more the North invests in clean capital. This is not because the capital accumulation is directly economically beneficial, but more as a device to reduce consumption to make it more in line with consumption in the South. When consumption patterns later diverge, it is because the North has accumulated excess capital and needs to consume it at some point in time. This consumption is postponed as discounting makes the difference in consumption levels less costly when postponed.

Figure 5 also confirms the theoretical result that inequality aversion leads to higher levels of dirty capital in the South. Although dirty capital initially depreciates at the same rate

[12] The Matlab code used to run the simulations are available from the authors upon request.
[13] In our simulations, the exponents in the production functions for dirty and clean inputs are set equal (0.5), see Appendix 2. An increase in the relative productivity of dirty capital would result in more investments in this type of capital.

regardless of region and inequality aversion, carbon leakage later becomes an issue and more so the higher the inequality aversion. Finally, the figure confirms our theoretical result of long run convergence in the clean and dirty capital stocks for the regions.

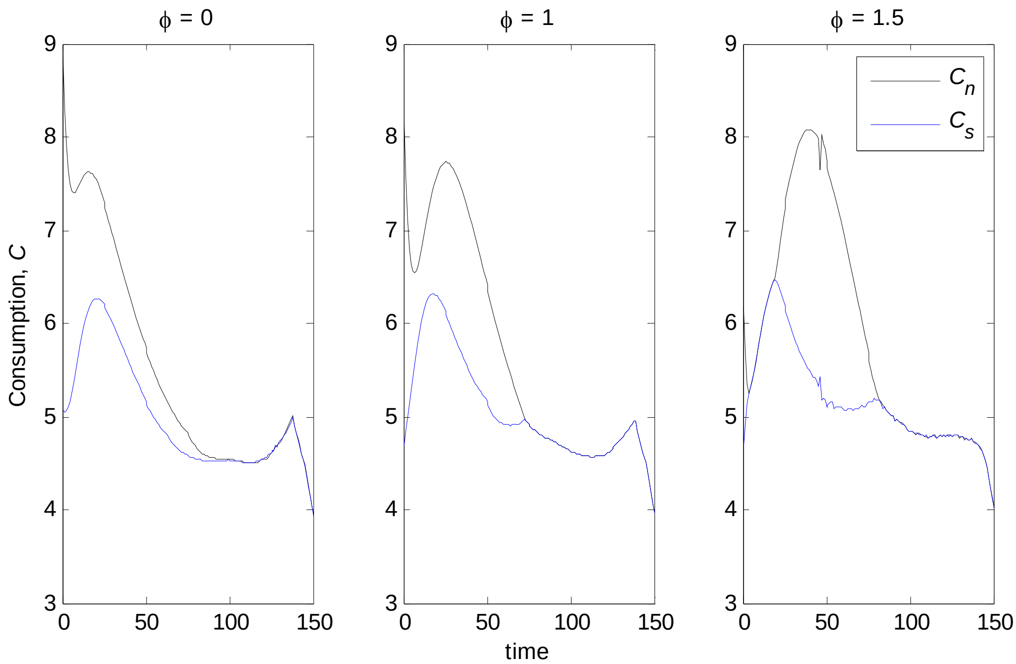

Figure 4: Consumption in the North (black) and the South (blue) for different levels of inequality aversion

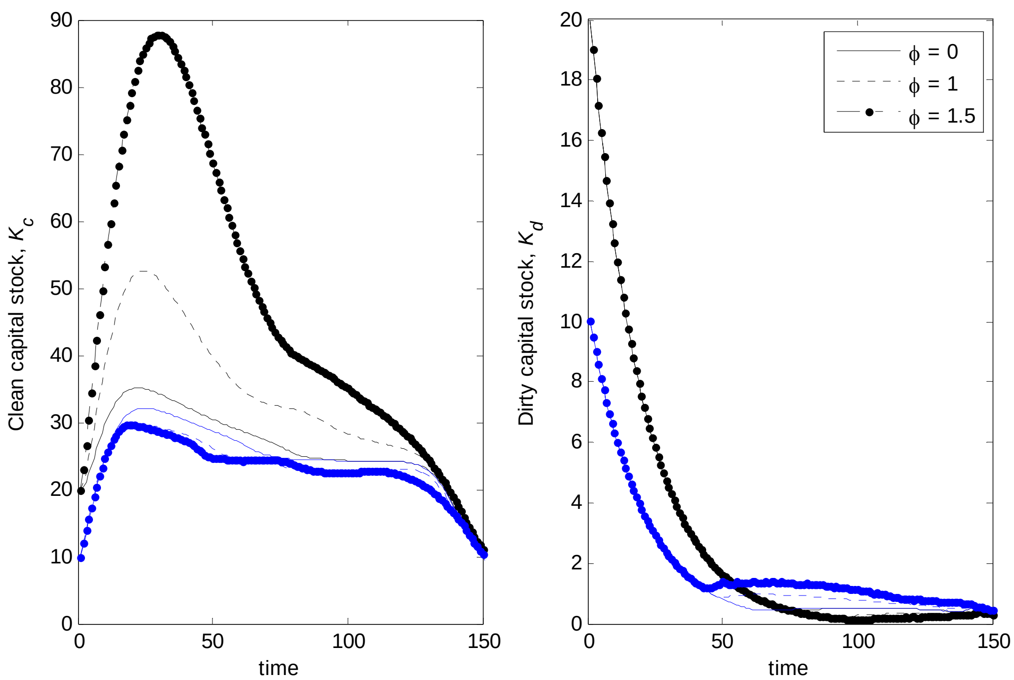

Figure 5: Accumulation of clean and dirty capital under different degrees of inequality aversion: North (black) and South (blue)

The effect of inequality aversion on global environmental services can be seen in Figure 6. Because of the carbon leakage to the South, as seen in Figure 5, the quality of the global environment is lower when inequality aversion is high. This effect is however transitory. In the long run we see that the environmental quality is independent of inequality aversion, which also confirms our previous theoretical result.

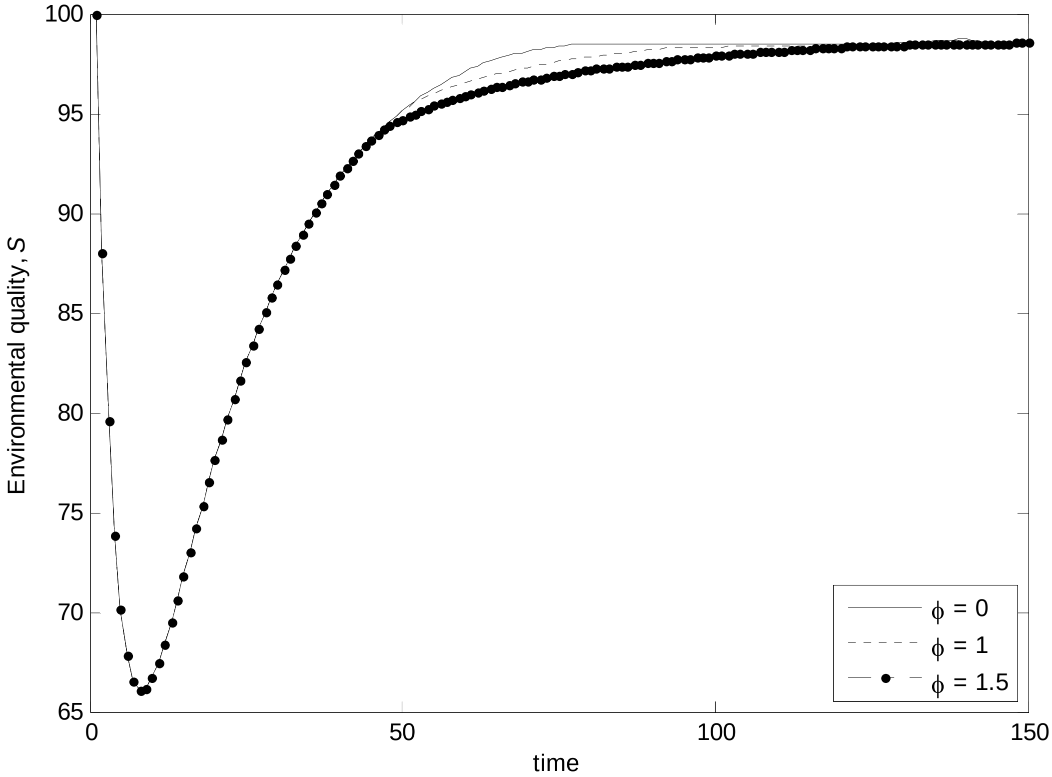

Figure 6: Global environmental quality

6. Conclusions

This paper studies the trade-off between intra- and intergenerational equity as represented by preferences to reduce future climate damage and to increase economic development in the poor world today. We find that inequality aversion within a generation will reduce consumption and greenhouse gas emissions in the rich region of the world, while consumption and emissions will increase in the poor region. This follows from a transfer from dirty to clean capital in the rich region, while the poor region predominantly invests in more productive dirty capital to catch up with the rich North. This result justifies carbon leakage from North to South as a result of promoting economic development. Introducing transfers between the regions reinforces this result. Whether total greenhouse gas emissions will increase as a result of income inequality aversion within one generation depends on the relative size of the emissions reductions in the North compared to increases in the South.

If we consider the Kyoto Protocol in light of our findings, the division between Annex I- and non-Annex I countries is justified. The first group consists of rich countries that have committed to reducing their greenhouse gas emissions, while the latter group consists of poorer developing countries that did not have to undertake emissions reductions. Hence, the division between Annex I- and non-Annex I countries was a way to transfer wealth from North to South by imposing the use of cleaner capital in the richest region, while poor countries could use dirty though more productive capital to speed up their development. The Kyoto Protocol also defined mechanisms that support capital transfers from rich to poor countries, such as the clean development mechanism, but did not consider debt remittance or similar direct transfers as ways to reduce the inequality between regions. Finally, our results suggest that poor countries only should be allowed to pollute more in the short run while they catch up with the North. In the long run, however, all countries must contribute to improved environmental quality by restricting their emissions. There was no mechanism in the Kyoto Protocol to commit non-Annex I countries to reduce emissions as they become richer.

Our work suggests that future climate agreements should contain mechanisms for wealth transfers to developing countries to reduce inequality, both in terms of lower emissions reductions and direct transfers in the form of debt remittance, development aid or the like. There should also be explicit mechanisms that limit the emissions of poorer countries as they become richer. Moreover, the results show that in certain situations, climate agreements should impose upon poor countries to increase their use of dirty capital in the short run to speed up their development, even if this yields higher levels of both local and global pollutants. This result is driven by an inequality aversion externality, as poor countries do not take into account the welfare loss of rich countries from inequality, and may therefore pollute too little and develop too slowly without emissions restrictions.

There are many possibilities for extending the current work. Our work abstracts from strategic considerations since all countries are assumed to be sufficiently small to take both prices and global environmental quality as given. In recent climate agreement discussions, it is quite clear that there are several big players in the game that exert a significant impact, both in relevant markets and on the climate. These include large countries such as the United States and China, as well as groups of countries that coordinate their actions (e.g. the European Union). Extending our analysis to allow for strategic interaction between countries and regions therefore seems highly relevant. Another possibility for future work is to investigate the implications of limited substitutability between clean and dirty capital. While we have assumed perfect substitution between clean and dirty capital, restrictions on the substitutability may help modify the results.

Appendix 1: Will consumption levels converge in the long run?

In the long run, the economy will enter a steady state in which there will be no growth in any of the model variables if there is no technological progress. We can therefore rewrite the first order conditions (15)-(19) for the steady-state levels of the variables:

$$\frac{\partial u(c_n, D_n, S)}{\partial c_n} - \phi = \lambda_n \tag{50}$$

$$\frac{\partial u(c_s, D_s, S)}{\partial c_s} + \phi = \lambda_s \tag{51}$$

$$\kappa_l \frac{\partial u(c_r, D_r, S)}{\partial D_r} + \lambda_r \left[\frac{\partial Y_r}{\partial K_{r,d}} + 1 - \delta_d \right] + \mu \kappa_g = \rho^{-1} \lambda_r, \; r = n, s \tag{52}$$

$$\lambda_r \left[\frac{\partial Y_r}{\partial K_{r,c}} + 1 - \delta_c \right] = \rho^{-1} \lambda_r, \; r = n, s \tag{53}$$

$$\sum_r \frac{\partial u(c_r, D_r, S)}{\partial S} = \mu \left(1 - \sigma - \rho^{-1} \right) \tag{54}$$

From equation (53), it is clear that the long-run clean capital level must be the same in the two regions. Since ρ and δ_c are the same for both countries, the marginal product of clean capital must be the same in the two regions, and with a production function that is increasing monotonically in clean capital, this implies that $K_{n,c} = K_{s,c} \equiv K_c^*$ in equilibrium.

Let us next consider the long-run levels of dirty capital given by equation (52), which can be rewritten as follows:

$$\lambda_r \left(\frac{\partial Y_r}{\partial K_{r,d}} + 1 - \delta_d - \frac{1}{\rho} \right) + \kappa_l \frac{\partial u(c_r, D_r, S)}{\partial D_r} = -\mu \kappa_g, \; r = n, s \tag{55}$$

Note that the right-hand side of this expression is the same for both the North and South. Consequently, we can write:

$$\lambda_n \left(MP_{n,d} + 1 - \delta_d - \rho^{-1} \right) - \lambda_s \left(MP_{s,d} + 1 - \delta_d - \rho^{-1} \right) = \kappa_l \left(MU_s^D - MU_n^D \right), \tag{56}$$

where $MP_{r,j} \equiv \frac{\partial Y_r}{\partial K_{r,j}}$ is the marginal productivity of capital j in region r, and

$MU_r^D \equiv \frac{\partial u(c_r, D_r, S)}{\partial D_r}$ is the marginal utility of local pollution in region r.

Let us first assume that the South never catches up with the North, so that $c_n > c_s$ in a steady state. However, introducing this into condition (56) shows that the condition cannot hold for $c_n > c_s$. To see this, let us rewrite equation (56) as follows:

$$(\lambda_s - \lambda_n)(1 - \delta_d - \rho^{-1}) + (\lambda_s MP_{s,d} - \lambda_n MP_{n,d}) + \kappa_l (MU_s^D - MU_n^D) = 0. \quad (57)$$

We know from (50) and (51) that the shadow price of capital must be lower in the North than in the South ($\lambda_n < \lambda_s$) since the North has a higher consumption level, and therefore a lower marginal utility from goods consumption, than the South. Thus, the first term of (57) must be positive for $c_n > c_s$. If $c_n > c_s$ it must also be the case that $K_{n,d} > K_{s,d}$, as we have already established that the level of clean capital is the same in the two regions in the long run. Since the marginal productivity is decreasing in the level of dirty capital, this implies that $MP_{n,d} < MP_{s,d}$. For this reason, the second term of (57) is also positive. As a result, condition (57) can only hold if its third term is negative. To evaluate the sign of the third term, note first that from equation (8) we have that $D_n > D_s$ if $K_{n,d} > K_{s,d}$. Next, from the assumptions $u''_{cD} \leq 0$ and $u''_{DD} \leq 0$, we have that MU_r^D is declining, both in consumption and local pollution. Since the North has higher levels of consumption and local pollution, it must be the case that $MU_s^D > MU_n^D$, which means that the third term of (57) is positive when $c_n > c_s$. Hence, condition (56) cannot hold unless $c_n = c_s$, which means that the consumption level of the two regions must converge in the long run.

Appendix 2: Parameters used in the simulation

a	Exponent in utility function (consumption)	0.5
b	Exponent in utility function (local environment)	0.2
c	Exponent in utility function (global environment)	0.4
ϕ	Aggregate inequality aversion	0, 1, 1.5
A_c	Total factor productivity in clean input production	0.9848
A_d	Total factor productivity in dirty input production	1.2933
Υ_c	Exponent in clean input production function	0.5
Υ_d	Exponent in dirty input production function	0.5
δ_c	Depreciation rate of clean capital	0.05
δ_d	Depreciation rate of dirty capital	0.05
$K_{n,c,0}$	Initial capital stock of clean capital in the North	20
$K_{n,d,0}$	Initial capital stock of dirty capital in the North	20
$K_{s,c,0}$	Initial capital stock of clean capital in the South	10
$K_{s,d,0}$	Initial capital stock of dirty capital in the South	10
κ_l	Local pollution per unit of dirty capital	0.5
κ_g	Global pollution per unit of dirty capital	0.4
σ	Regeneration rate of global environmental quality	0.25
$\bar{S}$	Initial environmental quality	100
v	Time preference rate	0.05
T	Time horizon	150

References

Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous (2012): The Environment and Directed Technical Change, The American Economic Review, 102(1): 131–66.

Anthoff, D. and R. Tol (2009): The Impact of Climate Change on the Balanced Growth Equivalent: An Application of FUND, *Environment and Resource Economics*, 43: 351-367.

Baumgärtner, S., S. Glotzbach, N. Hoberg, M. F. Quaas and K. H. Stumpf (2012): Economic Analysis of Trade-offs between Justices, *Intergenerational Justice Review* (1): 4-9.

Bye, B., S. Kverndokk and K. E. Rosendahl (2002): Mitigation costs, distributional effects, and ancillary benefits of carbon policies in the Nordic countries, the UK, and Ireland, *Mitigation and Adaptation Strategies for Global Change*, 7(4): 339-366.

Böhringer, C., C. Fischer and K. E. Rosendahl (2010): The Global Effects of Subglobal Climate Policies, *The BE Journal of Economic Analysis & Policy*, 10(2) (Symposium), Article 13.

Dasgupta, P. (2008): Discounting Climate change, *Journal of Risk and Uncertainty*, 37: 141-169.

Fehr, E. and K. Schmidt (1999): A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3): 817-868.

Felder, S. and Rutherford, T. F. (1993): Unilateral CO_2 Reductions and Carbon Leakage: The Consequences of International Trade in Oil and Basic Materials. *Journal of Environmental Economics and Management* 25(2): 162–76.

Glotzbach, S. and S. Baumgärtner (2012): The relationship between intragenerational and intergenerational ecological justice, *Environmental Values*, forthcoming.

Heal, G. (2009): Climate Economics: A Meta-Review and Some Suggestions for Future Research, *Review of Environmental Economics and Policy*, 3(1): 4-21.

Hoel, M. and T. Sterner (2007): Discounting and relative prices, *Climatic Change* 84: 265-280.

Intergovernmental Panel on Climate Change (IPCC) (2007a): *Climate Change 2007: The Physical Science Basis*, *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, UK.

Intergovernmental Panel on Climate Change (IPCC) (2007b): *Climate Change 2007: Mitigation of Climate Change, Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, UK.

Kolstad, C. (2011): Public Goods Agreement with Other-Regarding Preferences, NBER Working Paper No. 17017, The National Bureau of Economic Research.

Kverndokk, S. and A. Rose (2008): Equity and justice in global warming policy, *International Review of Environmental and Resource Economics*, 2(2): 135-176.

Nordhaus, W. D. (1993): Rolling the DICE: An optimal transition for controlling the emission of greenhouse gases, *Resource and Energy Economics*, 15: 27-50.

Nordhaus, W. D. and J. G. Boyer (2000): *Warming the World: The Economics of the Greenhouse Effect*, Cambridge, MA: MIT Press.

Nordhaus, W. D. (2007): A Review of the Stern Review on the Economics of Climate Change, *Journal of Economic Literature*, 45(3): 686–702.

Nævdal, E. and J. Vislie (2008): Climate Change, Catastrophic Risk and the Relative Unimportance of Discounting, Memorandum 28/2008, Department of Economics, University of Oslo.

Rawls, J. (1971): *A Theory of Justice*. Oxford: Oxford University Press.

Silva, E. C. D. and X. Zhu (2009): Emissions trading of global and local pollutants, pollution havens and free riding, *Journal of Environmental Economics and Management*, 58: 169-182.

Stern, N. (2007): *The Economics of Climate Change: The Stern Review*, Cambridge University Press, Cambridge, UK.

Sterner, T. and M. Persson (2008): An even Sterner review: Introducing relative prices into the discounting debate, *Review of Environmental Economics and Policy*, 2: 77-93.

Tol, R., S. Fankhauser, R. Richels and J. Smith (2000): How Much Damage Will Climate Change Do? Recent Estimates. *World Economics*, 1(4): 179-206.

Tol, R. (2002a): Estimates of the Damage Costs of Climate Change. Part 1: Benchmark Estimates, *Environmental and Resource Economics*, 21: 47-73.

Tol, R. (2002b): Estimates of the Damage Costs of Climate Change. Part II: Dynamic Estimates, *Environmental and Resource Economics*, 21: 135-160.

Weitzman, M. (2007a): Structural uncertainty and the value of statistical life in the economics of catastrophic climate change, NBER Working Paper No. 13490, the National Bureau of Economic Research, Cambridge, MA.

Weitzman, M. (2007b): A Review of the Stern Review on the Economics of Climate Change, *Journal of Economic Literature*, 45(3): 703–724.

Yohe, G. W., R. D. Lasco, Q. K. Ahmad, N. W. Arnell, S. J. Cohen, C. Hope, A. C. Janetos and R. T. Perez (2007): Perspectives on climate change and sustainability, Chapter 20 in *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, UK, 811-841.