Passing the Buck(s): Welfare-Improving International Emission-Reduction Subsidies

Nicholas Lawson*
Dalhousie University

January 22, 2025
PRELIMINARY - PLEASE DO NOT CITE

Abstract

Climate change is the most important environmental issue facing the world today, and yet it has proven difficult to achieve international cooperation in solving this massive global externality. My paper proposes an alternative approach based on a unilateral emissions-reduction subsidy offered to poor countries by a rich country. I start with a simple two-country model in which a rich country and a poor country choose their emissions and experience the damages caused by the resulting temperature increase. Assuming that the rich country is the Stackelberg leader, I calibrate the model to data from a variety of sources, where I model the OECD as the rich country and non-OECD countries as the poor country. I show that a selfish rich country would choose to offer a large subsidy per unit of total emissions reduction by the poor country, amounting to more than $1 trillion dollars per year. Such a subsidy policy would significantly reduce emissions by the poor country, dramatically lowering the global temperature increase, and could even raise global welfare above that of any climate treaty that excludes international transfers, because it reduces both emissions and poverty. These results are largely robust to extensions of the model to settings with multiple poor and rich countries.

Keywords: climate change, greenhouse gas abatement, international environmental agreements, emissions-reduction subsidy, technology transfer

*nicholas.lawson@dal.ca. I thank Dean Spears and Yan Lau for their helpful comments and suggestions, as well as participants of a seminar at Dalhousie University, as well as presentations at the CIRANO Applied Microeconomics Day 2024, the 23rd Journées Louis-André Gérard-Varet, and the 17th Meeting of the Society for Social Choice and Welfare. Any errors or omissions are the responsibility of the author.

1 Introduction

Climate change has been described as the "biggest threat to security that modern humans have ever faced", and as "the defining issue of our time" (United Nations Security Council, 2021). However, achieving sustained and effective international cooperation in solving this massive global externality has proved to be very difficult. The standard approach has centred around international environmental agreements (IEAs) such as the Kyoto Protocol and the Paris Agreement, but there is a lot of concern among environmental economists that such agreements are destined to be ineffective (Barrett, 1994, 2003, 2013; Maggi and Staiger, 2022; de Zeeuw, 2015). As a result, a significant literature has developed to study methods of encouraging cooperation, such as conditional transfers (Goussebaïle et al., 2023; Barrett, 2001; Carraro et al., 2006; Li and Fujita, 2021).

In this paper, I propose an alternative approach to international policy action on climate change. Rather than considering IEAs and the possible ways to provide incentives for countries to participate in such agreements, I consider the way in which a single large country could act unilaterally to incentivize other countries to reduce emissions.

I start by presenting a simple two-country model of climate change, in which both a rich country and a poor country choose their emissions subject to an environmental constraint, in the form of damages caused by the resulting temperature increase which reduce production possibilities, as is standard in integrated assessment models.[1] I assume a Stackelberg setting in which the rich country moves first, and I calibrate the model to data from a variety of sources, where I model the OECD together as the rich country and non-OECD countries as the poor country.

I then consider a variety of scenarios. I start with the Stackelberg equilibrium, in which emissions are inefficiently high. I can also evaluate the global emissions optimum, where emissions in each country are chosen to maximize the sum of rich and poor country utility, which is my measure of global welfare. This scenario is a useful comparison as it would

[1] In an alternative version of my model, I also considered a setting in which countries can choose both emissions per person *and* population, since about three-quarters of the increase in emissions since 1970 are due to the growth of population rather than emissions per person (Gerlagh et al., 2023). However, it was difficult to model a choice over population in an intuitive way, and population did not vary much in response to policies, so here I make population exogenous, as is standard in integrated assessment models (Nordhaus, 2018).

presumably be the optimal outcome of a negotiated climate treaty if one ruled out transfers or subsidies, but recent experience suggests that it may be politically infeasible. The next scenario would likely be even less feasible, as it supposes that the rich country wants to maximize global welfare rather than just their own utility, accounting for the poor country's response, generating an 85% reduction in OECD emissions per person to make space for non-OECD countries.

My main results come from two scenarios in which a selfish rich country can offer the poor country a subsidy per unit of total emissions reduction from the Stackelberg baseline. I consider two possible forms that the subsidy could take: a simple monetary subsidy, and a conditional technology transfer. The results are remarkable: to reduce the damages caused by excessive temperature increase, the OECD would selfishly want to offer a large emission-reduction subsidy to the poor country, leading to large reductions in emissions per person in the latter. When I consider a monetary subsidy, the subsidy amounts to about $1.11 trillion dollars per year, or 1.8% of OECD GDP at the optimum, in the form of a subsidy of $37 per tonne of emission reductions, and would be estimated to reduce the global temperature increase above the level from the year 1900 by nearly a degree and a half. The global welfare gain from this monetary subsidy would be equivalent to a 1.8% increase in consumption in both countries. If a conditional technology transfer is feasible, it can also improve welfare, though it appears to be dominated by the monetary subsidy at the very approximate values of the cost of the transfer that I use.

The logic of this result is that rich countries have better technologies, which can produce more output per unit of emissions, so one efficient solution is for them to pay poor countries to produce less. Because poor countries have low incomes, and thus high marginal utilities of income, a relatively small subsidy from the perspective of the rich country can have a significant impact on the incentives facing poor countries. Remarkably, such a unilateral subsidy policy could even make the rich countries – and the world as a whole – better off than at the aforementioned global "emissions optimum", or the best possible negotiated climate treaty that excludes transfers between countries. The subsidy policy kills two birds with one stone: redistribution and emissions abatement at the same time!

Finally, I extend the model beyond its two-country setting, to consider settings with

many poor countries and either one or two rich countries. The results of these extensions are generally similar to those from the baseline model, in that the optimal policy still involves a large monetary emission-reduction subsidy to poor countries. The welfare gain from such a policy is in fact much larger in the multi-country model, because of a greater unresolved externality between poor countries. However, the welfare gain from the monetary subsidy is now slightly smaller than the gain that could be obtained from the global emissions optimum, though the main conclusion of the paper – that it is in the rich countries' selfish interest to implement a large emissions-reduction subsidy, which would significantly increase global welfare – remains robust in this setting.

This paper contributes to the extensive literature on international environmental agreements, as well as the related literatures on transfers and other mechanisms to encourage cooperation. My results at least partly contradict the argument of Hovi et al. (2015) that "unilateral climate policies – although certainly worthwhile as a supplement to international cooperation – cannot be expected to cut global emissions to the extent required to avoid dangerous anthropogenic climate change". Indeed, my paper argues that unilateral policy can make a major contribution to cutting global emissions, if we include emission-reduction subsidies that one rich country offers to other countries within the category of unilateral action. In that sense, my results are closer to the more optimistic view of Goussebaïle et al. (2023): "Given the practical difficulties of negotiating with a large number of countries, it is worth highlighting that an optimal outcome can be attained by giving the right incentives for the provision of global public goods to countries outside a formal coalition." My paper shows how this can be done in the context of climate change, and my paper also advances the objectives of Moellendorf (2014) of working towards both solving climate change and eradicating global poverty, and supports the claims in Sheeran (2006) and Rive and Rübbelke (2010) that transfers may help people in poor countries, improving global equality.

It is important to note that my analysis uses a simple utilitarian social welfare function that adds the welfare from each country when analyzing global welfare; unlike many regionally-disaggregated integrated assessment models like RICE (Regional Integrated model of Climate and the Economy) (Nordhaus, 2017), I do not use Negishi welfare weights, which weight a country's welfare according to the inverse of their marginal utility of income. Negishi

weights imply that utility in rich countries has a greater weight in global welfare than that in poor countries; see Stanton (2011) for further discussion of their use and consequences. Nonetheless, my results are not driven by this assumption, because my main results consider the selfishly-optimal policy from the perspective of the rich country rather than the global-welfare-maximizing outcome, and I find that a large emissions-reduction subsidy offered by the rich country to the poor country is actually a Pareto improvement with respect to the baseline scenario.

Of course, it is also important to acknowledge some important caveats. In particular, the political feasibility of the approach proposed by this paper is uncertain. Several papers argue that citizens of rich countries like the United States are unlikely to support arrangements that involve transfers to support mitigation in foreign countries (Underdal et al., 2012; Buntaine and Prather, 2018); Bernauer and Gampfer (2015) argues that there is robust support for unilateral action in both the United States and India, but the type of unilateral action considered does not include the type of subsidy proposed in this paper. However, Carlsson et al. (2022) and Gaikwad et al. (2023) identify conditions under which conditional climate leadership and climate transfers are more likely to be accepted; in particular, the latter finds that transfers should involve domestic agents and compensate the most affected communities in order to be acceptable to the citizens of both giving and receiving countries.

The rest of the paper proceeds as follows. Section 2 presents an overview of the relevant literature, and section 3 presents the baseline two-country model. Section 4 describes the calibration of that model, and section 5 presents the policy results, while section 6 presents results from multi-country extensions of the baseline model. Finally, section 7 concludes the paper.

2 Literature

A large economic literature uses game theory to study international environmental agreements. The conclusions of this literature are mostly fairly pessimistic, including Barrett (1994), Barrett (2003), Barrett (2013), and Maggi and Staiger (2022); the overall literature has been surveyed by de Zeeuw (2015), as well as by Hovi et al. (2015) who also survey some possible solutions. Among such possible solutions, McGinty (2007) argues that trade

in pollution permits may permit improvements with asymmetric nations, while Eichner and Pethig (2015) argue that the first-best IEA may be self-enforcing if a tax is used instead of cap-and-trade, and Heitzig et al. (2011) propose a dynamic system of linear compensation. DeCanio and Fremstad (2013) study a variety of possible game structures and argue that the science supports the use of a coordination game rather than the prisoners' dilemma, but Madani (2013) argues that such simple models are far too simplified to provide useful policy advice. While much of the literature has been pessimistic about IEAs, Michaelowa (2015) is also skeptical of the value of bottom-up voluntary approaches rather than top-down agreements.

Given the challenges of constructing sizable coalitions of countries that cooperate in an IEA, a large and growing literature considers the use of transfers to encourage cooperation. Buchholz and Konrad (1995) highlights the incentive that less-productive agents may have to make large unconditional transfers to more productive agents in the private provision of public goods, which could generate a Pareto improvement, and Buchholz et al. (1997) find that the follower may have an incentive to make a lump-sum monetary transfer to the leader in a Stackelberg game even if the two are equally productive. Goussebaïle et al. (2023) find that altruistic developed countries could provide optimal abatement incentives to developing countries through development aid without a global coalition.

In terms of more conditional transfers, Barrett (2001) argues that, if there is strong asymmetry among countries, buying cooperation through side payments can be beneficial, and so do Carraro et al. (2006) and Li and Fujita (2021), as well as Fuentes-Albero and Rubio (2010) as long as the asymmetry is in damages rather than abatement costs. Sælen (2016) also suggests the use of side payments as a means of expanding "climate clubs".[2] Kerr et al. (2019), meanwhile, suggest that the desirability of such transfers may depend on the timing of the transfers and if they are "results-based" or not; and Bayramoglu and Jacques (2015) seems to be the only paper that considers monetary transfers that are costly in terms of implementation, leakages, or administrative and political costs. The potential of conditional transfers to improve global equality is noted by Sheeran (2006) and Rive and

[2]Landis and Bernauer (2012) argue that differences between desired carbon prices could justify transfers, but do not consider incentive-compatibility.

Rübbelke (2010), since the countries with lower mitigation costs tend to be poorer.

However, this literature generally focusses on transfers as a mechanism to encourage countries to participate in an international environmental agreement, where the transfers are side payments that act as a mechanism for distributing the global efficiency gain from participation, rather than for providing marginal incentives for abatement. Unlike most of the existing literature, the latter mechanism is what my paper will focus on.

Technology transfers as a way to address greenhouse gas emissions have also been studied by several papers, with mixed findings. Yang (1999) finds positive effects if such technology can only reduce externalities, but Aronsson et al. (2010) argue that the attractiveness of technology transfers from the perspective of richer countries depends on the nature of the equilibrium, and Stephan and Müller-Fürstenberger (2015) express concern about rebound and leakage effects. Sarr and Swanson (2017) also reach a pessimistic conclusion if the transferred technology is sufficiently general. These concerns are part of the reason why I consider conditional technology transfers in this paper.

My paper proposes a unilateral subsidy policy by a rich country, and is thus related to the literature that considers a variety of unilateral or leadership-based policies, including papers related to technology transfers (Yang, 1999), unilateral contributions with matching (Arce M., 2001), and unilateral contributions with endogenous innovation (Chatterji and Ghosal, 2009). However, the first two do not expect policy to be selfishly beneficial to the rich country, and the third relies on a process of endogenous innovation that is not present in my model. A somewhat-related – but not unilateral – matching-commitment game is studied by Molina et al. (2020).

To my knowledge, there has been no research that has considered the use of conditional transfers or subsidies outside of the setting of an IEA or a "climate club". My main contribution will be to focus on unilateral policy of a new form and show that an IEA may not be necessary to achieve a significant reduction in greenhouse gas emissions.

3 Two-Country Model of Emissions and Climate Change

In this section, I will present the main model used in the paper, starting with the baseline version and then expanding to allow for subsidies. The basic structure of the model corre-

sponds to a simplified reduced form of the RICE model used by papers such as Budolfson et al. (2021). I assume a world with two countries represented by $i \in \{1, 2\}$; larger numbers of countries will be considered in the extensions in section 6. For simplicity, I model a one-shot game in which each country's exogenously-fixed population[3] N_i represents the expected population of the country in 2050;[4] this is related to assumptions in many models of climate change that focus on outcomes up to 2100 (Xu and Ramanathan, 2017; Nordhaus, 2018), on the grounds that that time period is crucial for addressing climate change. It is also useful to focus on optimal policies in a single period, whereas the dynamics of those policies across generations are both complicated and second-order to the main contribution of my paper.

Each country chooses their emissions per person e_i subject to production technology T_i, where I assume that $T_1 > T_2$, so that country 1 is the rich country and country 2 is the poor country. I will not focus on how to implement e_i within each country, including any particular domestic policies that are needed to achieve a given value of e_i, and I assume away heterogeneity within countries as is standard in the literature. Per-person gross output in country i is then equal to $T_i \left(\delta + (1-\delta)\left(\frac{e_i}{E}\right)^{\gamma}\right)$ for $\{\delta, \gamma\} \in (0, 1)$, where E will in practice be set to the calibration target for e_1. This concave production function is chosen to correspond to the types of abatement cost functions used in DICE models such as Nordhaus (2018), though I cannot use their standard functional form for abatement costs because I may need to allow for an increase in production when emissions increase above the initial level, which the standard form does not permit.[5]

As in most integrated assessment models, net output is equal to gross output multiplied

[3]I have also considered an alternative version of the model which allows for endogenous fertility, and which renders the final population endogenous to policies; however, the results – available upon request – are qualitatively similar to what I present here despite the greater complexity of the alternative model.

[4]It is important to account for the fact that the countries may be of very different sizes, and in particular that the poor country may have a much larger population than the rich country. It is difficult, and unusual in the literature, to model two players of different inherent "sizes" in a game such as mine; Boadway and Hayashi (1999) is one of the few papers to explicitly model players of different sizes. 2050 is used as an approximate midpoint of an approximately 80-year period, using a point before the actual midpoint as population growth is slowing over time.

[5]In Nordhaus (2018), the abatement cost as a percentage of lost output is $\Lambda(t) = \theta_1(t)\mu(t)^{\theta_2}$, for a percentage emissions reduction of $\mu(t)$, where $\theta_1(t)$ and θ_2 are technological parameters: $\theta_1(t)$ declines with time t and $\theta_2 = 2.6$. I will use this abatement cost function to calibrate the parameters of my production function later, but we can already see that it motivates a setting in which there is a zero-emissions technology that can produce $1 - \theta_1(t)$ percent of baseline output. However, negative abatement – or an increase in emissions above the baseline level – is not possible in this functional form.

by a damage function $1-D_i(t)$, and as in Nordhaus (2018) I use a quadratic damage function with no linear term: $D_i(t) = \psi_i t^2$, where $\psi_i > 0$ is the country-specific damage parameter and t is the global temperature increase in 2100 above the level in the year 1900.[6] Per-person consumption is then equal to:

$$c_i = \left(1 - \psi_i t^2\right) T_i \left(\delta + (1-\delta)\left(\frac{e_i}{E}\right)^{\gamma}\right).$$

Per-person utility from consumption is given by the constant relative risk-aversion function $U(c_i) = \frac{c_i^{1-\alpha}}{1-\alpha}$ for $\alpha > 0$, and so welfare in country i is given by:

$$V_i = \frac{N_i}{1-\alpha}\left[\left(1 - \psi_i t^2\right) T_i \left(\delta + (1-\delta)\left(\frac{e_i}{E}\right)^{\gamma}\right)\right]^{1-\alpha}.$$

To specify the function linking emissions to the temperature increase t, I use the "baseline-fast" scenario in Xu and Ramanathan (2017), which corresponds to about 4 trillion tonnes of additional projected carbon-dioxide-equivalent emissions by 2100. Box 1 Figure in Xu and Ramanathan (2017) then indicates that the temperature increase caused by the emissions up to 2100 would be about 4 degrees, whereas the baseline concentrations in 2020 would already ensure 1 degree of warming, indicating a slope of 0.75 degrees per trillion tonnes. Therefore, I use the function $t = 1 + 0.75 e_T$, where $e_T = \eta(N_1 e_1 + N_2 e_2)$ represents the total trillions of tonnes of carbon-dioxide-equivalent emissions over the next 80 years, for some scaling parameter η that depends on how the units of N_i and e_i are defined.[7]

Returning to the structure of the strategic model, I assume a Stackelberg setting in which country 1 is the leader: country 1 chooses e_1, and then country 2 observes that choice and selects e_2. To solve using backward induction, I start with the optimization problem for

[6] I have also considered an alternative version of the model which assumed a binary setting with some risk of a "climate disaster", as in Bohn and Stuart (2015). Results for optimal policies – which are available upon request – were qualitatively similar to those from the current model. One could also imagine that climate damages might be uncertain, leading to a distribution of possible values of ψ_i, but this would complicate the analysis without introducing new insight, as I would still need to calibrate that distribution to match the baseline emissions target of 4 trillion tonnes of carbon-dioxide-equivalent emissions as described below.

[7] This simple linear specification corresponds well to the predictions from models with more detailed climate modules: Figure SPM.5 in IPCC (2014) suggests that 4 trillion tonnes of carbon-dioxide-equivalents would lead to the atmospheric concentration stabilizing in the 720-1000 parts-per-million range, and simulations in Table 3 of Gillingham et al. (2018) find that concentrations in the middle of that range correspond to warming of about 3.6 to 3.87 degrees. IPCC (2014) further notes that "Multiple lines of evidence indicate a strong, consistent, almost linear relationship between cumulative CO2 emissions and projected global temperature change to the year 2100."

country 2:

$$\max_{e_2} V_2 = \frac{N_2}{1-\alpha}\left[\left(1-\psi_2\left(1+0.75\eta(N_1e_1+N_2e_2)\right)^2\right)T_2\left(\delta+(1-\delta)\left(\frac{e_2}{E}\right)^{\gamma}\right)\right]^{1-\alpha}.$$

It is easy to verify that the second-order condition for a maximum is satisfied, and the first-order condition is given by:

$$1.5\psi_2 t\eta N_2\left(\delta+(1-\delta)\left(\frac{e_2}{E}\right)^{\gamma}\right) = \left(1-\psi_2 t^2\right)(1-\delta)\frac{\gamma}{E}\left(\frac{e_2}{E}\right)^{\gamma-1}$$

and while there is no closed-form solution, this condition implicitly defines $e_2(e_1)$.

Working backwards, country 1 chooses e_1 taking into account country 2's reaction:

$$\max_{e_1} V_1 = \frac{N_1}{1-\alpha}\left[\left(1-\psi_1\left(1+0.75\eta(N_1e_1+N_2e_2(e_1))\right)^2\right)T_1\left(\delta+(1-\delta)\left(\frac{e_1}{E}\right)^{\gamma}\right)\right]^{1-\alpha}.$$

This leads to the following first-order condition:

$$1.5\psi_1 t\left(N_1+N_2\frac{de_2}{de_1}\right)\left(\delta+(1-\delta)\left(\frac{e_1}{E}\right)^{\gamma}\right) = \left(1-\psi_1 t^2\right)(1-\delta)\frac{\gamma}{E}\left(\frac{e_1}{E}\right)^{\gamma-1}.$$

Given the complicated solutions for $e_2(e_1)$ and e_1, it is clear that this is not an analytically simple problem. However, the basic intuition is clear: each country will choose their emissions per person to maximize their own utility, taking into account the other's choice (prior or subsequent) as well as the environmental constraint embodied in the damage function $D_i(t)$. As is usual in a model with a negative externality, there will be a tendency for both countries to choose excessive emissions, leading to inefficiently high climate damages which reduce gross output and thus consumption per capita.

Having characterized the solution to the baseline model, I can then ask several welfare- and policy-related questions:

1. How does the Stackelberg outcome compare to the global emissions optimum, or the values of $\{e_1, e_2\}$ that maximize $W \equiv V_1 + V_2$?[8]

[8] As discussed in the introduction, I use a simple utilitarian global welfare function, rather than weighting countries' utility by the inverse of their marginal utility of income. Ackerman et al. (2013) and Budolfson et al. (2021) evaluate optimal allocations in models that also do not use such "Negishi weights", though under different assumptions about available policy tools: Ackerman et al. (2013) find that large capital flows to poor countries are optimal, while Budolfson et al. (2021) rule out such redistribution and find that carbon prices should vary significantly by region. My analysis is different: instead of the policy that maximizes global welfare, I focus primarily on the policy that is selfishly optimal from the rich country's perspective – and then show that that policy actually generates a Pareto improvement.

2. How does the Stackelberg outcome compare to the outcome if the rich country wants to maximize global welfare rather than just their own utility, accounting for the poor country's response? That is, if the rich country is benevolent, what would they choose for their own emissions e_1?

3. What is the optimal selfish policy of the rich country if, as well as choosing e_1, they can also choose a subsidy rate p to pay to the poor country for reductions in their emissions below the Stackelberg level?

4. What is the optimal selfish policy of the rich country if, as well as choosing e_1, they can also choose a technology transfer rate q to provide to the poor country for reductions in their emissions below the Stackelberg level?

To answer the third question, I introduce a new policy instrument: p will be a conditional monetary subsidy, for which I define the problem as follows:

$$\max_{e_2} V_2 = \frac{N_2}{1-\alpha}\left[\left(1-\psi_2\left(1+0.75\eta(N_1e_1+N_2e_2)\right)^2\right)T_2\left(\delta+(1-\delta)\left(\frac{e_2}{E}\right)^\gamma\right)+p(e_2^S-e_2)\right]^{1-\alpha}$$

$$\max_{e_1,p} V_1 = \frac{N_1}{1-\alpha}\left[\left(1-\psi_1\left(1+0.75\eta(N_1e_1+N_2e_2)\right)^2\right)T_1\left(\delta+(1-\delta)\left(\frac{e_1}{E}\right)^\gamma\right)-\frac{pN_2}{N_1}(e_2^S-e_2)\right]^{1-\alpha}$$

where e_2^S is the per-person emissions of country 2 in the Stackelberg equilibrium, which I assume is used as the baseline for the subsidy payments. I assume that country 1 pays country 2 a total amount equal to $p(N_2e_2^S - N_2e_2)$, or the subsidy rate p times the reduction in total emissions, so that per-person payments have to be divided by each country's population size. Importantly, I assume that country 1 can commit to a subsidy structure as a function of the realized emissions of country 2; my model has a one-period structure, but such an assumption could be justified from the fact that, with a sufficiently high discount factor in a repeated game, the Folk Theorem implies that such an outcome could be supported as an equilibrium. I also must assume that the emissions of country 2 are observable and verifiable by country 1; this may necessitate the installation of a system of monitors, but the same assumption of observability is standard in all papers that consider international environmental agreements in which countries commit to their respective amounts of emissions or abatement. Finally, I assume that the subsidy is linear starting from e_2^S for simplicity and practicality, though alternative subsidy structures could be imagined.

For the final question, I instead assume that the rich country will pay to increase the poor country's technology T_2 based on the latter's reduction in emissions. To be precise, if q is the technology transfer rate, suppose that the technology level of country 2 is increased to $\hat{T}_2 = T_2 + q(N_2 e_2^S - N_2 e_2)$. Meanwhile, to calculate the cost to country 1 of such a transfer, I assume that, to raise T_2 by t percent of the way from the baseline T_2 to the rich country's value T_1, the rich country must spend $k(t) = \frac{\ln(1-t)+t}{-\mu}$. Therefore, the problem for the two countries is:

$$\max_{e_2} V_2 = \frac{N_2}{1-\alpha}\left[\left(1-\psi_2\left(1+0.75\eta(N_1e_1+N_2e_2)\right)^2\right)\hat{T}_2\left(\delta+(1-\delta)\left(\frac{e_2}{E}\right)^\gamma\right)\right]^{1-\alpha}$$

$$\max_{e_1,q} V_1 = \frac{N_1}{1-\alpha}\left[\left(1-\psi_1\left(1+0.75\eta(N_1e_1+N_2e_2)\right)^2\right)T_1\left(\delta+(1-\delta)\left(\frac{e_1}{E}\right)^\gamma\right)+\frac{\ln(1-t)+t}{\mu N_1}\right]^{1-\alpha}$$

where $\hat{T}_2 = T_2 + q(N_2 e_2^S - N_2 e_2)$ and $t = \frac{q(N_2 e_2^S - N_2 e_2)}{T_1 - T_2}$.

To answer the question of what monetary subsidy rate p or technology transfer rate q would be selfishly optimal for country 1, an analytical approach will be insufficient, given the complexity of the problem, and my analysis will be primarily numerical, based on a calibrated version of the model that I will present in the following section; as noted by Sælen (2016), "many studies use numerical simulations because of difficulties of solving analytical models of IEAs with heterogeneous actors". However, at least in the case of the monetary subsidy p, I am able to derive several interesting comparative statics in my model, which are presented in appendix A: I can show that $\frac{de_2}{dp} < 0$, and that $\frac{\partial V_1}{\partial p} > 0$ at $p = 0$ (and the optimized Stackelberg value of e_1), so that country 1's utility maximizing subsidy rate will always be positive. I also show that $\frac{de_2^2}{dpdT_2} > 0$ at $p = 0$, so that a smaller value of T_2 (a poorer, lower-technology poor country) tends to make the negative effect of p on e_2 stronger: when the rich country has significantly better technology than the poor country – and thus is significantly richer – the subsidy means a lot to the poor country, and the rich country can get a larger value for money from a subsidy in terms of lower poor-country emissions. The comparative statics for the technology transfer q are much more complicated and do not yield simple theoretical results.

4 Calibration of Model

To evaluate the welfare implications of the various scenarios discussed above, I now calibrate the model to data from the real world. For any values of the full set of parameters $\{\alpha, \psi_1, \psi_2, T_1, T_2, \delta, \gamma, N_1, N_2, \eta, E\}$, I can solve the model numerically for the Stackelberg equilibrium, and I want that equilibrium to match the data as closely as possible. I continue to use the two-country model – though I will consider a larger number of countries in the extensions in section 6 – and so I need to divide the world into one rich country and one poor country. Due to the availability of good-quality statistics from the OECD, I will model the OECD countries collectively as the rich country, whereas the poor country will represent all non-OECD countries.

Table 1 presents the moments that I will use in the calibration: the values that I am trying to match are in column 1, while the values generated by the simulated model are in column 2. Meanwhile, the values of the parameters used in the calibrated model can be found in Table 2: those in panel A are chosen directly to match data or other evidence, as described below, while those in panel B are chosen collectively to minimize the sum of squared percentage deviations between the moments in the data and the analogous moments simulated in the model (the two columns of Table 1).

Table 1: Moments from Data and Calibrated Model

Moment	Model	Data	Source
country 1 emissions per capita (tonnes/year)	10.560	10.560	See table notes
country 2 emissions per capita (tonnes/year)	5.619	5.619	See table notes
country 1 GDP per capita (US $000s)	39.532	39.532	World Bank Indicators 2019
country 2 GDP per capita (US $000s)	5.366	5.366	World Bank Indicators 2019

Notes: Total emissions per year are taken from Climate Watch (2023): in 2019, CO2 equivalents were 10.56 tonnes per capita for the OECD and 6.48 for the world as a whole. Populations are then taken from the UN World Population Prospects for 2022 (United Nations, 2022) for the world as a whole, and from the OECD's Population Projections (Organisation for Economic Co-operation and Development, 2022) for the OECD: the starting populations in 2022, of 1.384 billion for the OECD and 6.558 billion for the rest of the world, are used to calculate emissions per capita outside the OECD.

In panel A of Table 2, the populations are the projected values for 2050 from Table I.1 of United Nations (2022) (for the entire world) and from Organisation for Economic Co-operation and Development (2022) (for the OECD); the same sources are used in Table 1 for

Table 2: Calibrated Model Parameters

Description	Name	Value
Panel A: Selected or Matched Directly to Data		
coefficient of relative risk-aversion	α	1.45
level/slope parameter of production function	δ	0.8848
curvature parameter of production function	γ	0.2442
2050 population of country 1 (billions)	N_1	1.487
2050 population of country 2 (billions)	N_2	8.200
emissions scaling parameter	η	0.0647
emissions output weight	E	10.560
Panel B: Calibrated Parameters		
climate change damage parameter of country 1	ψ_1	0.0071
climate change damage parameter of country 2	ψ_2	0.0013
production technology in country 1	T_1	44.631
production technology in country 2	T_2	5.575

Notes: This table contains the parameter values of the model, where panel A presents parameters for which values are selected from external estimates, and panel B contains parameters whose values are calibrated to minimize the sum of squared percentage deviations between the moments in Table 1 and the analogous moments simulated in the model.

the values for 2022, which are used to calculate emissions per capita. Emissions per capita and GDP per capita in Table 1 are current estimates, and are derived from sources described in the notes to Table 1; I effectively assume that the ratio of emissions and GDP in rich versus poor countries are not projected to change across generations. The population values are from 2050 to represent something close to an average over the next 80 years (accounting for the fact that the rate of population growth is slowing and is project to turn negative before 2100).

The rest of the values from panel A of Table 2 are chosen as follows. The value of α – the coefficient of relative risk-aversion, or more appropriately the rate of inequality aversion in this non-stochastic environment – is set to 1.45 as in Nordhaus (2018), though a sensitivity analysis in appendix B shows that my main results are robust to a smaller value of α. I assume that E is equal to the target emissions for country 1, so that $\frac{e_1}{E} = 1$ at baseline, and then the values of δ and γ are chosen to match as closely as possible the abatement function from Nordhaus (2018), where the percentage production reduction from a $\mu(t)$ percent emission abatement is $\Lambda(t) = \theta_1(t)\mu(t)^{\theta_2}$, where $\theta_1(t)$ starts at 0.0741 and declines

with time t and $\theta_2 = 2.6$. This equation from Nordhaus (2018) is typical of recent DICE models, but I cannot simply use this function, because I may need to allow for an increase in production when a country's emissions increase above the initial level. Therefore, using the starting value of 0.0741 for the θ_1 from Nordhaus (2018), I choose the values of δ and γ such that the integral of $\delta + (1-\delta)(1-\mu)^\gamma$ is as close as possible to that of $1 - \theta_1 \mu^{\theta_2}$ over $\mu \in [0, 1]$, which gives me $\delta = 0.8848$ and $\gamma = 0.2442$.

Finally, I need to choose an appropriate value of η, the scaling parameter that turns emissions from the model into worldwide trillions of tonnes of carbon-dioxide-equivalents over 80 years. I define N_i as billions of people and e_i as tonnes per year per person, and this defines η as the scaling parameter that turns tonnes per person per year into world-wide trillions of tonnes over 80 years. A natural starting point would be a value of 0.08, since the calculation involves tonnes times billions of people times 80 years divided by 1 trillion; however, projected emissions tend to decrease over time, and I want to correspond to the "baseline-fast" scenario in Xu and Ramanathan (2017) in which total emissions between now and 2100 are 4 trillion tonnes of carbon-dioxide-equivalent emissions. Therefore, I calculate the η for which $e_T \equiv \eta\,(N_1 e_1 + N_2 e_2)$ is equal to 4 at the target values of e_1 and e_2, which is $\eta = 0.0647$.

The 4 parameters that are chosen to match the moments in Table 1 are the climate change damage parameters ψ_i and the technology parameters T_i. As expected, the OECD's technology parameter is significantly higher than that of the rest of the world; and the damage parameter is also higher in the OECD, and higher than is typical in the literature.[9] By choosing the damage parameters in this way, I implicitly assume that governments are correctly accounting for the climate damages caused by their pollution when choosing their Stackelberg equilibrium values of e_i. An alternative approach would be to take estimated "correct" values for the damage parameters from the literature and assume that the governments may behave according to incorrect values of ψ_i; however, this approach would conflate benefits from international coordination with benefits from efficient internalization of externalities within each country. Nonetheless, appendix B presents the results from a sensitivity analysis in which I allow for incorrect beliefs about ψ_i: in particular, I permit that the true

[9]For example, Nordhaus (2018) uses a value of 0.00227.

value of ψ_2 is equal to 0.0071 as for the OECD, and solve for policies in that setting, with ambiguous results that depend on country 2's ability to modify their domestic policy to act upon the correct larger value of ψ_2.

5 Optimal Policy Results

Using the values of the parameters from the previous section, I can now simulate the model in various scenarios, and solve for optimal outcomes and policies, which are presented in Table 3. Column (1) contains the results from the baseline Stackelberg equilibrium, which is the one that was used to calibrate the model in section 4. Along with the familiar values of emissions and income per capita, utility is reported for each country, as well as the combined global welfare $W \equiv V_1 + V_2$ and the temperature increase in degrees Celsius above the level in the year 1900.

Table 3: Optimal Outcomes & Policies

	(1)	(2)	(3)	(4)	(5)
Variable	Stackelberg	Optimum	Benevolent	Subsidy	Technology
country 1 emissions per capita	10.560	1.577	1.568	10.960	7.831
country 2 emissions per capita	5.619	5.016	6.359	1.983	5.018
country 1 income per capita	39.532	39.769	38.663	41.804	40.104
country 2 income per capita	5.366	5.398	5.402	5.448	5.378
country 1 utility	-0.632	-0.630	-0.638	-0.616	-0.628
country 2 utility	-8.556	-8.533	-8.530	-8.498	-8.547
combined welfare	-9.187	-9.163	-9.168	-9.114	-9.174
temperature increase	4.000	3.111	3.645	2.581	3.564
subsidy as % of GDP				0.018	0.004
welfare gain as % of GDP		0.006	0.005	0.018	0.003

Notes: I use the same units for population, emissions per capita, and GDP per capita as in Table 1.

The subsequent column (2) presents the global emissions optimum: the values of $\{e_1, e_2\}$ that maximize combined welfare W. In comparison to the Stackelberg outcome in column (1), we can see that emissions are lower in both countries, but particularly in country 1. This result is due to the relatively large value of α, which implies that the marginal utility of income is much lower in country 1, and that the latter should therefore reduce their emissions more significantly. If α were smaller, the reduction in emissions would be smaller in country 1 and larger in country 2, as shown in appendix B, because technology is better in the OECD

(country 1), making it feasible for them to produce a larger amount of consumption per unit of emissions.[10] Utility is higher for both countries at the emissions optimum than at the Stackelberg equilibrium, as the pollution externality is resolved and the temperature increase is significantly reduced, but the utility gain is particularly large for country 2. The overall welfare gain is presented in the final row of the table, and is equivalent to a 0.6% increase in consumption in both countries.

This global emissions optimum might seem like the natural target for an IEA, but since the OECD gains little from implementing this global optimum, they are unlikely to want to pursue it if alternatives exist; it is of interest particularly as a point of comparison for the subsidies that I consider later. Column (3) presents another unrealistic scenario, in which the rich country implements a benevolent policy in which they want to maximize global welfare rather than just their own utility, accounting for the poor country's response. In this scenario, I find that emissions per capita in the OECD should decrease by slightly more than at the global optimum, to make more room for the poor country to slightly increase their emissions in response, even while the temperature increase drops somewhat to about 3.6 degrees. Obviously, the rich country is worse off in this scenario than at baseline, and the poor country somewhat better off, but the overall welfare effect is lower than at the optimum, because the poor country will react to any reduction in rich-country emissions by increasing e_2, preventing a large reduction in the temperature increase.

Next, I suppose that both countries are selfish as in the baseline Stackelberg model, but that the OECD can implement a monetary subsidy rate p on country 2's reductions in total emissions from the Stackelberg baseline. Formally, I assume that for each unit of reductions in emissions, the OECD will transfer p, costing them $\frac{p}{N_1}$ per person, and individuals in the poor country will thus receive a transfer of $\frac{p}{N_2}$ per person and per unit of reduction in emissions. Of course, I abstract here from questions of distribution and the various possibilities of how the transfer would be used in country 2: I assume that the transfer would simply be distributed among the population of country 2 and that everyone in the poor country would benefit equally. In terms of the subsidy that maximizes the utility of

[10]The use of Negishi weights would also alter this outcome, but as previously mentioned, I do not weight each country's utility by the inverse of their marginal utility.

country 1, the results in column (4) are dramatic: the OECD would choose to offer a subsidy rate of $p = 0.037$, which would cause emissions per capita in the poor country to drop to only 35% of their baseline level, while emissions per capita increase by 3.8% in the OECD in reaction to the reduction elsewhere. The temperature increase is only 2.58 degrees, much lower than both the Stackelberg equilibrium and the global emissions optimum.

In this scenario, both the rich and poor countries are better off than in the global emissions optimum,[11] because a new policy instrument has been introduced – and it is a powerful instrument. The basic intuition is that the poor country's emissions per capita are too high, given their technology, and so it is beneficial for the OECD to pay them to produce less and live partly off the transfers from the OECD. Even the OECD is better off in this scenario because, even though they make large transfers to the poor country, doing so both provides more carbon space for the OECD to increases its emissions and significantly reduces the climate damages from the temperature increase; both forces lead to an increase in consumption per capita even after the transfer. The poor country is poor enough that they respond quickly to the subsidy, making it a cheap way for the rich country to purchase carbon space. And because the subsidy combines climate action with redistribution from rich to poor, it represents a Pareto improvement not only relative to the Stackelberg equilibrium, but also relative to the global emissions optimum![12]

The optimal transfer is quantitatively very large, as the total amount transferred is equivalent to $1.11 trillion per year, or 1.8% of total OECD GDP, while the subsidy rate translates into a value of $37.34 per tonne. The overall transfer is much larger than in any paper in the literature, where global values of optimal transfers are never higher than the values of $380 billion per year or 0.5% of GDP found by Sælen (2016).[13] The amount is also much larger than any transfers which have taken place in the world so far: Organisation for Economic Co-operation and Development (2023) reports that total climate finance from developed to developing countries, from both public and private sources, reached a new high

[11]The "global emissions optimum" is the optimal allocation when countries can choose their emissions but no other policy variables are permitted. As a result, it is not necessarily the optimum once an emissions-reduction subsidy is introduced.

[12]Yang (1999) also points out that transfers from rich to poor would improve equity.

[13]Strand (2020) considers "transformational" climate finance, which seems like an appropriate word to describe the subsidy policy here; however, Strand (2020) focusses on mitigation projects with increasing returns to scale and doesn't perform any quantitative analysis.

of $89.6 billion in 2021.

These results indicate that an emissions-reduction subsidy given by country 1 to country 2 is capable of generating significant welfare gains for both countries; and it is important to note that this is not a global optimum requiring negotiation between countries, but rather an incentive-compatible mechanism introduced by the rich country for their own benefit (which also happens to benefit the poor country). Of course, one could question certain aspects and assumptions underlying this approach: would poor countries really be willing to reduce their emissions by about 65%? Accounting for the reduction in climate damages relative to the Stackelberg case, this corresponds to a 1% reduction in output (or a 2.3% reduction ignoring the reduction in climate damages), which is not enormous, but it would introduce a significant reliance on a transfer from the rich country which would amount to 2.6% of poor-country GDP. It is possible that poor countries would find that such an approach would leave them too vulnerable to global events (such as the Covid-19 pandemic) which would affect the rich countries' ability or willingness to continue the transfers. At the same time, one could question whether rich countries would be able to work together to make such a transfer; this point will be addressed below when I extend the model to more than 2 countries. The political acceptability of such transfers may also be questionable, as highlighted by Buntaine and Prather (2018) and Gaikwad et al. (2023).[14] However, regardless of the precise quantitative results, the qualitative point is remarkable: large transfers from rich countries to poor countries to subsidize emissions reductions can be incentive-compatible and beneficial to the rich countries.

If the justification for a subsidy is that poor countries have worse technology and thus find it cheaper (in terms of lost output) to reduce emissions than rich countries, this suggests that a technology transfer from rich to poor countries could be alternative way for the OECD to get poor countries to reduce emissions. Therefore, for the final scenario in column (5), I assume that country 1 can act to raise the poor country's technology T_2 from its baseline value, but at some cost to the rich country, as they must pay for transportation and adaptation of technological components. Specifically, I assume that country 1 chooses

[14]There is also the issue of whether the rich country can really commit to the subsidy structure, but the same issue of commitment arises in any model of negotiation over climate pledges.

a conditional technology transfer rate q, and transfers technology according to the poor country's reduction in emissions, raising the latter's technology level to $\hat{T}_2 = T_2 + q(N_2 e_2^S - N_2 e_2)$. The cost of this transfer to country 1 is $k(t) = \frac{\ln(1-t)+t}{-\mu}$ for an increase in T_2 by t percent of the way from the baseline T_2 to the rich country's value T_1. To calibrate μ, I use the fact that the latest Nationally Determined Contributions represent a 7% drop in emissions from 2019 by 2030, as reported by SDG Knowledge Hub (2022), and that Organisation for Economic Co-operation and Development (2023) states that developing countries will need \$2.4 trillion annually (during 2026-30) to tackle global warming. I therefore choose μ such that the cost of a technology transfer that offsets (in terms of production) a 7% reduction in emissions is \$2.4 trillion per year.

The results in column (5) indicate that the technology transfer is dominated by the monetary subsidy, at least at the rough estimate of the cost described above. Both countries prefer the monetary subsidy, although the technology transfer does make at least the rich country better off than at the emissions optimum. The total cost of the subsidy to the OECD is \$223 billion, improving the poor country's technology by 0.05% and raising global welfare by an amount that is equivalent to a 0.3% increase in consumption, while the temperature increase in this scenario is about 3.6 degrees.[15] The feasibility of this scenario is unclear, and the value for the cost parameter is very approximate, so it is conceivable that a much less expensive technology transfer could be more successful at raising welfare, but at the current estimates, its value appears to be much less than that of the monetary subsidy.

The results from this section indicate that the rich country has an incentive to offer a large per-unit monetary subsidy to the poor country for reductions in emissions, and that such a unilateral policy on the part of the rich country could significantly raise utility for both the poor and rich countries while lowering the resulting temperature increase. However, these results rely on a simplified setting in which there are only two countries in the world. As a result, the next section of the paper will extend the analysis to a broader set of circumstances, including multiple poor and rich countries.

[15]It is also possible to consider an unconditional technology transfer, but since T_2 cancels out of the first-order condition for e_2 in section 3, the specification of my model rules out T_2 having an effect on the emissions of the poor country – or more accurately, the "rebound effect" which concerned Stephan and Müller-Fürstenberger (2015) is complete – and so the rich country has no incentive to provide what is purely a donation of improved technology to the poor country.

6 Model Extensions: Policy Results

First of all, before expanding the number of countries in the model, I can also consider a setting in which country 2 is the Stackelberg leader. I keep the same calibration of the model, which is based on an equilibrium in which the rich country is the leader, but I then assume that country 2 can take over the leadership role and offer subsidies. In results that are available upon request, the various policies don't work anywhere near as well: the poor country isn't wealthy enough to give large monetary subsidies to the rich country, and it is harder for such subsidies to reduce emissions in the rich country since the subsidy has less impact in a country with a lower marginal utility from consumption. As a result, the optimal monetary subsidy from the poor country to the OECD is much, much smaller, at about 0.01% of poor country GDP, or only $4.5 billion (less than 1% of the OECD-to-poor-country subsidy); and the subsidy's effect on outcomes is also quite modest, at only 27% of the welfare gain from moving to the global optimum. The conditional technology transfer, meanwhile, is impossible if the transferring country has worse technology than the recipient. These results further highlight the mechanism driving the usefulness of a rich-to-poor subsidy: the rich country's subsidy can have a large effect on emissions in a country that is relatively poor and has high marginal utility from consumption.

A more interesting extension considers a world with more than two countries. To start, I consider a setting with one rich country and many poor countries, where the poor countries play a Nash equilibrium subgame over their emissions at the second step of the model. I recalibrate the model to this case, with 10 poor countries of equal size (which is equal to 10% of the size of the single poor country in the baseline model) for simplicity, to avoid having to solve 10 different maximization problems, and the results of the calibration can be found in Table 4. While the technology parameters are similar to those from Table 2, the climate change damage parameter is much higher for country 2, which is more in line with the usual assumption in the literature. When there is more than one poor country, less of the climate externality is internalized inside each country, and it becomes easier to understand why poor-country emissions are as high as they are without requiring a low climate damage parameter.

Table 4: Moments from Data and Model & Calibrated Model Parameters with 10 Poor Countries

Panel A: Moments from Data and Model		
Moment	Model	Data
country 1 emissions per capita (tonnes/year)	10.560	10.560
poor country emissions per capita (tonnes/year)	5.619	5.619
country 1 GDP per capita (US $000s)	39.532	39.532
poor country GDP per capita (US $000s)	5.366	5.366

Panel B: Calibrated Parameters		
Description	Name	Value
climate change damage parameter of country 1	ψ_1	0.0081
climate change damage parameter of country 2	ψ_2	0.0112
production technology in country 1	T_1	45.455
production technology in country 2	T_2	6.651

Notes: This table contains the parameter values of the model with 10 poor countries, where panel A presents the moments from the data and the calibrated model, and panel B contains parameters whose values are calibrated to minimize the sum of squared percentage deviations between those two sets of moments. The parameters for which values are selected from external estimates in panel A of Table 2 are unchanged except that the population of each poor country is 820 million, or 10% of the total from Table 2.

The optimal policies from the perspective of the rich country are fairly similar to the baseline model, as shown in Table 5: in the case of the monetary subsidy, I find that the OECD should provide an emission-reduction subsidy to the poor countries amounting to $564 billion or 0.9% of GDP, which is smaller than the baseline result but remains quantitatively significant. This subsidy doesn't quite raise welfare to the level of the global emissions optimum in column (2), though it comes close, and the welfare gain from the subsidy is much larger than before, at 12.2% of consumption. Meanwhile, the conditional technology transfer once again provides a welfare gain relative to the Stackelberg equilibrium, but underperforms both the monetary subsidy and the global emissions optimum. This multiple-poor-country scenario features a larger baseline international climate externality, because rather than having one large poor country that internalizes much of the externality, there are many poor countries that act strategically towards each other. This implies that the existing levels of emissions are consistent with a larger climate damage parameter, and that the welfare gain from solving this externality is larger than before, either using the subsidies or in the global emissions optimum.

Table 5: Optimal Outcomes & Policies: 10 Poor Countries

Variable	(1) Stackelberg	(2) Optimum	(3) Benevolent	(4) Subsidy	(5) Technology
country 1 emissions per capita	10.560	0.390	0.256	11.619	6.651
poor country emissions per capita	5.619	1.191	6.612	0.594	5.091
country 1 income per capita	39.532	41.777	37.734	43.601	40.232
poor country income per capita	5.366	6.174	5.585	6.030	5.628
country 1 utility	-0.632	-0.616	-0.645	-0.604	-0.627
poor country utility	-0.856	-0.803	-0.840	-0.812	-0.837
combined welfare	-9.187	-8.648	-9.048	-8.722	-9.001
temperature increase	4.000	1.502	3.651	2.076	3.507
subsidy as % of GDP				0.009	0.004
welfare gain as % of GDP		0.144	0.034	0.122	0.047

Notes: I use the same units for population, emissions per capita, and GDP per capita as in Table 1.

If, along with multiple poor countries, there are also multiple rich countries, the model becomes significantly more complicated and requires an assumption about the nature of the strategic interaction between rich countries. I have calibrated the model with 10 poor countries and 2 rich countries, where the latter specifically represent the United States (country 0) and the rest of the OECD (country 1). There are a variety of ways that this game could be structured, but I have assumed that the US is the Stackelberg leader, and that the other rich country and the 10 poor countries all move simultaneously in a second-stage Nash equilibrium subgame.[16] I also assume that the emissions output weight E is now equal to the target emissions for country 0. The calibration results can be found in Table 6, where GDP per capita, emissions per capita, and starting and 2050 populations are taken from the same sources as before (but dividing the OECD into the US and the rest). The results for the technology parameters are unsurprising: T_2 is almost unchanged, but T_0 is higher and T_1 is lower than the prior values for T_1. ψ_2 is similar to the value from Table 4, but the estimated climate damages are higher in the US than in the rest of the OECD; this is due

[16]The alternatives of a three-stage game, or a two-stage game with both rich countries in the first stage, run into the obstacle of smoothing countries' utility functions. In stages prior to the final one, countries must account for how their choices will affect the choices of countries that act in later stages, and even infinitesimal numerical imprecision in the later stages can generate considerable instability in earlier stages; this problem is minimized by placing all countries but one in the final stage. An extension to all of the countries in the world would also make the numerical problem much more time-consuming, since I would need to solve separate maximization problems for each country, and would run into issues of data availability for many countries.

to the assumption that the US is the Stackelberg leader, as such a small leader could have emitted more unless they were responding to a high expected cost from climate damages.

Table 6: Moments from Data and Model & Calibrated Model Parameters with 2 Rich Countries & 10 Poor Countries

Panel A: Moments from Data and Model		
Moment	Model	Data
country 0 emissions per capita (tonnes/year)	17.580	17.580
country 1 emissions per capita (tonnes/year)	8.300	8.300
poor country emissions per capita (tonnes/year)	5.619	5.619
country 0 GDP per capita (US $000s)	65.120	65.120
country 1 GDP per capita (US $000s)	31.295	31.295
poor country GDP per capita (US $000s)	5.366	5.366

Panel B: Calibrated Parameters		
Description	Name	Value
climate change damage parameter of country 0	ψ_0	0.0169
climate change damage parameter of country 1	ψ_1	0.0061
climate change damage parameter of country 2	ψ_2	0.0103
production technology in country 0	T_0	89.158
production technology in country 1	T_1	35.369
production technology in country 2	T_2	6.606

Notes: This table contains the parameter values of the model with 2 rich countries and 10 poor countries, where panel A presents the moments from the data and the calibrated model, and panel B contains parameters whose values are calibrated to minimize the sum of squared percentage deviations between those two sets of moments. The parameters for which values are selected from external estimates in panel A of Table 2 are unchanged except that the population of each poor country is 820 million, the population of country 0 (the US) is 389 million, the population of country 1 (the rest of the OECD) is 1.098 billion, and the emissions scaling parameter is now 0.0645.

The results for the optimal policies from the point of view of the US can be found in Table 7, where I assume that the US can offer emissions-reduction incentives separately to country 1 and the poor countries: p_1 and q_1 for country 1, p_2 and q_2 for the poor countries. The results are generally similar to those from Table 5, as the optimal monetary subsidy from the US to the poor countries is once again large, at $531 billion or about 1.6% of US GDP, leading to a large reduction in emissions in the poor countries; there is also a smaller subsidy of $176 billion to the rest of the OECD, and the US significantly increases their emissions per capita. The temperature increase is 2.0 degrees rather than 4.0, and welfare increases significantly, with all countries benefitting, though the welfare gain is again slightly less than that from the emissions optimum in column (2). Meanwhile, the technology transfer again

produces a smaller welfare gain than the monetary subsidy; the technology transfer is only given to the poor countries, because a similar conditional transfer to the rest of the OECD is ineffective at actually reducing the latter's emissions.

Table 7: Optimal Outcomes & Policies: 2 Rich Countries & 10 Poor Countries

	(1)	(2)	(3)	(4)	(5)
Variable	Stackelberg	Optimum	Benevolent	Subsidy	Technology
country 0 emissions per capita	17.580	0.259	0.172	23.160	12.698
country 1 emissions per capita	8.300	0.437	8.799	6.208	9.356
poor country emissions per capita	5.619	1.164	6.006	0.555	4.968
country 0 income per capita	65.120	79.467	61.649	82.081	67.279
country 1 income per capita	31.295	32.501	31.585	33.785	31.864
poor country income per capita	5.366	6.095	5.450	5.987	5.504
country 0 utility	-0.132	-0.121	-0.135	-0.119	-0.130
country 1 utility	-0.518	-0.509	-0.516	-0.501	-0.514
poor country utility	-0.856	-0.808	-0.850	-0.814	-0.846
combined welfare	-9.206	-8.709	-9.147	-8.764	-9.103
temperature increase	4.000	1.490	3.852	1.986	3.706
country 1 subsidy as % of GDP				0.005	
poor country subsidy as % of GDP				0.016	0.009
welfare gain as % of GDP		0.131	0.014	0.116	0.025

Notes: I use the same units for population, emissions per capita, and GDP per capita as in Table 1.

One possible concern which could arise when considering two (or more) rich countries is whether there would be any conflict over which country should take the lead in providing the subsidy. In the case of both the monetary subsidy and the technology transfer, there could be a conflict, but not in the way that one might initially suspect: being the Stackelberg leader is appealing to both rich countries, and each of the two rich countries would prefer to be the leader and offer the subsidy to the other countries. So, in this model at least, there is no reason to fear a conflict in which no rich country wants to offer emissions-reduction incentives to poor countries.

The results of this section confirm the general conclusions from section 5: emissions-reduction incentives offered by a rich country to poor countries (and possibly other rich countries), particularly in the form of a monetary subsidy, appears to be a promising direction for climate policy, with the possibility of generating significant welfare gains both from a reduction in the global temperature increase and from global redistribution from rich to poor. Such an approach could avoid the long, complicated and messy negotiations that have

characterized international action against climate change for several decades. Importantly, such subsidies would actually be in the selfish interest of rich countries, providing them with utility gains that may approach or even surpass what they could hope to obtain from an ideal (and probably infeasible) climate treaty. In fact, it is possible for the global welfare gains from a unilateral subsidy to be greater than the gain from implementing the optimal hypothetical negotiated climate treaty, at least in the absence of side payments. These conclusions are largely qualitatively robust to allowing for multiple poor and rich countries in the model, with quantitative impacts and welfare gains that may be even larger than with two countries.

7 Conclusion

This paper considers optimal international emission-reduction incentives in a two-country model of climate change, in which both a rich country and a poor country choose their emissions subject to an environmental constraint. I assume a Stackelberg setting in which the rich country moves first, and I calibrate the model to data from the OECD (the rich country) and non-OECD countries (the poor country). I show that the optimal monetary subsidy from the rich country to the poor country is large, at over $1 trillion per year or about 2% of OECD GDP, and generates a Pareto improvement and a global welfare gain that is even larger than the gain from an optimal climate change agreement without subsidies. A conditional technology transfer could also generate welfare gains, although at the rough estimate of the cost of such a transfer, it underperforms the monetary subsidy. Extensions to more than two countries generate similar results, with large emissions-reduction subsidies from rich countries to poor countries, with even larger estimated welfare gains from the selfishly-optimal rich-country policy.

Essentially, the paper indicates that it is in rich countries' best interests to pay poor countries to emit less, rather than waiting for a hypothetical optimal climate treaty. Of course, there are obstacles to the implementation of such a policy, including political obstacles in the form of public opinion which may be skeptical of paying or receiving such large transfers. But papers like Gaikwad et al. (2023) give us reason to hope that transfers can be designed in such a way as to overcome such obstacles, and it remains much simpler than

negotiating a truly binding climate treaty among many countries. Another caveat is related to the assumption that transfers to poor countries will actually benefit their populations; as noted by Goussebaïle et al. (2023), "donor countries allocate aid to recipient countries' governments and not directly to the poor, which generates inefficiencies". Nonetheless, if a subsidy would lead to emissions reduction, it would already achieve a great deal, and despite all of the caveats, there is reason to hope that a unilateral subsidy approach could work towards the joint objectives of solving climate change and reducing global poverty.

References

Ackerman, F., E. A. Stanton, and R. Bueno (2013): "CRED: A New Model of Climate and Development," *Ecological Economics*, 85, 166–176.

Arce M., D. G. (2001): "Leadership and the Aggregation of International Collective Action," *Oxford Economic Papers*, 53, 114–137.

Aronsson, T., K. Backlund, and L. Sahlén (2010): "Technology Transfers and the Clean Development Mechanism in a North-South General Equilibrium Model," *Resource and Energy Economics*, 32, 292–309.

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

——— (2001): "International Cooperation for Sale," *European Economic Review*, 45, 1835–1850.

——— (2003): *Environment and Statecraft: The Strategy of Environmental Treaty-Making*, Oxford: Oxford University Press.

——— (2013): "Climate Treaties and Approaching Catastrophies," *Journal of Environmental Economics and Management*, 66, 235–250.

Bayramoglu, B. and J.-F. Jacques (2015): "International Environmental Agreements: The Case of Costly Monetary Transfers," *Environmental and Resource Economics*, 62, 745–767.

Bernauer, T. and R. Gampfer (2015): "How Robust is Public Support for *Unilateral* Climate Policy?" *Environmental Science & Policy*, 54, 316–330.

Boadway, R. and M. Hayashi (1999): "Country Size and the Voluntary Provision of International Public Goods," *European Journal of Political Economy*, 15, 619–638.

Bohn, H. and C. Stuart (2015): "Calculation of a Population Externality," *American Economic Journal: Economic Policy*, 7, 61–87.

Buchholz, W. and K. A. Konrad (1995): "Strategic Transfers and Private Provision of Public Goods," *Journal of Public Economics*, 57, 489–505.

Buchholz, W., K. A. Konrad, and K. E. Lommerud (1997): "Stackelberg Leadership and Transfers in Private Provision of Public Goods," *Review of Economic Design*, 3, 29–43.

Budolfson, M. B., D. Anthoff, F. Dennig, F. Errickson, K. Kuruc, D. Spears, and N. K. Dubash (2021): "Utilitarian Benchmarks for Emissions and Pledges Promote Equity, Climate and Development," *Nature Climate Change*, 11, 827–833.

Buntaine, M. T. and L. Prather (2018): "Preferences for Domestic Action Over International Transfers in Global Climate Policy," *Journal of Experimental Political Science*, 5, 73–87.

Carlsson, F., M. Kataria, E. Lampi, Å. Löfgren, and T. Sterner (2022): "Leading by Example? EU Citizens' Preferences for Climate Leadership," Working Paper in Economics No. 828, University of Gothenburg.

Carraro, C., J. Eyckmans, and M. Finus (2006): "Optimal Transfers and Participation Decisions in International Environmental Agreements," *Review of International Organizations*, 1, 379–396.

Chatterji, S. and S. Ghosal (2009): "Technology, Unilateral Commitments and Cumulative Emissions Reduction," *CESifo Economic Studies*, 55, 286–305.

Climate Watch (2023): "Historical GHG Emissions," `https://www.climatewatchdata.org/ghg-emissions?breakBy=regions&calculation=PER_CAPITA&end_year=2019&regions=WORLD%2COECD&start_year=1990`, World Resources Institute, viewed on March 6, 2023.

de Zeeuw, A. (2015): "International Environmental Agreements," *Annual Review of Resource Economics*, 7, 151–168.

DeCanio, S. J. and A. Fremstad (2013): "Game Theory and Climate Diplomacy," *Ecological Economics*, 85, 177–187.

Eichner, T. and R. Pethig (2015): "Self-Enforcing International Environmental Agreements and Trade: Taxes versus Caps," *Oxford Economic Papers*, 67, 897–917.

Fuentes-Albero, C. and S. J. Rubio (2010): "Can International Environmental Cooperation be Bought?" *European Journal of Operational Research*, 202, 255–264.

Gaikwad, N., F. Genovese, and D. Tingley (2023): "Climate Action from Abroad: Assessing Mass Support for Cross-Border Climate Compensation," Unpublished Paper.

Gerlagh, R., V. Lupi, and M. Galeotti (2023): "Fertility and Climate Change," *Scandinavian Journal of Economics*, 125, 208–252.

Gillingham, K., W. Nordhaus, D. Anthoff, G. Blanford, V. Bosetti, P. Christensen, H. McJeon, and J. Reilly (2018): "Modeling Uncertainty in Integrated Assessment of Climate Change: A Multimodel Comparison," *Journal of the Association of Environmental and Resource Economists*, 5, 791–826.

Goussebaïle, A., A. Bommier, A. Goerger, and J.-P. Nicolaï (2023): "Altruistic Foreign Aid and Climate Change Mitigation," *Environmental and Resource Economics*, 84, 219–239.

Heitzig, J., K. Lessmann, and Y. Zou (2011): "Self-Enforcing Strategies to Deter Free-Riding in the Climate Change Mitigation Game and Other Repeated Public Good Games," *PNAS*, 108, 15739–15744.

Hovi, J., H. Ward, and F. Grundig (2015): "Hope or Despair? Formal Models of Climate Cooperation," *Environmental and Resource Economics*, 62, 665–688.

IPCC (2014): "Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change," Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.). `https://www.ipcc.ch/site/assets/uploads/2018/02/SYR_AR5_FINAL_full.pdf`.

Kerr, S., S. Lippert, and E. Lou (2019): "Financial Transfers and Climate Cooperation," Working Paper 19-04, Motu Economic and Public Policy Research.

Landis, F. and T. Bernauer (2012): "Transfer Payments in Global Climate Policy," *Nature Climate Change*, 2, 628–633.

Li, Q. and T. Fujita (2021): "Transfers in International Environmental Agreements under Heterogeneity," *Applied Economics*, 53, 3900–3911.

Madani, K. (2013): "Modeling International Climate Change Negotiations More Responsibly: Can Highly Simplified Game Theory Models Provide Reliable Policy Insights?" *Ecological Economics*, 90, 68–76.

Maggi, G. and R. W. Staiger (2022): "International Climate Agreements and the Scream of Greta," Working Paper no. 30681, NBER, Cambridge, MA.

McGinty, M. (2007): "International Environmental Agreements among Asymmetric Nations," *Oxford Economic Papers*, 59, 45–62.

Michaelowa, A. (2015): "Opportunities for and Alternatives to Global Climate Regimes Post-Kyoto," *Annual Review of Environment and Resources*, 40, 395–417.

Moellendorf, D. (2014): *The Moral Challenge of Dangerous Climate Change: Values, Poverty, and Policy*, Cambridge: Cambridge University Press.

Molina, C., E. Akçay, U. Dieckmann, S. A. Levin, and E. A. Rovenskaya (2020): "Combating Climate Change with Matching-Commitment Agreements," *Scientific Reports*, 10, 10251.

Nordhaus, W. (2010): "Excel file for RICE model as of April 26, 2010," https://doi.org/10.3886/ICPSR28461.v1, Inter-university Consortium for Political and Social Research [distributor].

——— (2017): "Integrated Assessment Models of Climate Change," *NBER Reporter*, 3, 16–20.

——— (2018): "Projections and Uncertainties about Climate Change in an Era of Minimal Climate Policies," *American Economic Journal: Economic Policy*, 10, 333–360.

Organisation for Economic Co-operation and Development (2022): "Population Projections," https://stats.oecd.org/Index.aspx?DataSetCode=POPPROJ#, OECD.Stat, viewed on March 6, 2023.

——— (2023): "Climate Finance Provided and Mobilised by Developed Countries in 2013-2021: Aggregate Trends and Opportunities for Scaling Up Adaptation and Mobilised Private Finance, Climate Finance and the USD 100 Billion Goal," https://doi.org/10.1787/e20d2bc7-en, OECD Publishing, Paris.

Rive, N. and D. T. Rübbelke (2010): "International Environmental Policy and Poverty Alleviation," *Review of World Economics*, 146, 515–543.

Sælen, H. (2016): "Side-Payments: An Effective Instrument for Building Climate Clubs?" *International Environmental Agreements: Politics, Law and Economics*, 16, 909–932.

Sarr, M. and T. Swanson (2017): "Will Technological Change Save the World? The Rebound Effect in International Transfers of Technology," *Environmental and Resource Economics*, 66, 577–604.

SDG Knowledge Hub (2022): "WRI Analysis of Updated NDCs Reveals Need for Transformational Change," https://sdg.iisd.org/news/wri-analysis-of-updated-ndcs-reveals-need-for-transformational-change/, International Institute for Sustainable Development, viewed on March 7, 2023.

Sheeran, K. A. (2006): "Side Payments or Exemptions: The Implications for Equitable and Efficient Climate Control," *Eastern Economic Journal*, 32, 515–532.

Stanton, E. A. (2011): "Negishi Welfare Weights in Integrated Assessment Models: The Mathematics of Global Inequality," *Climatic Change*, 107, 417–432.

Stephan, G. and G. Müller-Fürstenberger (2015): "Global Warming, Technological Change and Trade in Carbon Energy: Challenge or Threat?" *Environmental and Resource Economics*, 62, 791–809.

Strand, J. (2020): "Transformational Climate Finance," Policy Research Working Paper 9251, World Bank Group, Development Economics, Development Research Group.

Underdal, A., J. Hovi, S. Kallbekken, and T. Skodvin (2012): "Can Conditional Commitments Break the Climate Change Negotiations Deadlock?" *International Political Science Review*, 33, 475–493.

United Nations (2022): "World Population Prospects 2022: Summary of Results," https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf, Department of Economic and Social Affairs, viewed on March 6, 2023.

United Nations Security Council (2021): "Climate Change 'Biggest Threat Modern Humans Have Ever Faced', World-Renowned Naturalist Tells Security Council, Calls for Greater Global Cooperation," https://press.un.org/en/2021/sc14445.doc.htm, United Nations, viewed on August 4, 2023.

Xu, Y. and V. Ramanathan (2017): "Well Below 2 °C: Mitigation Strategies for Avoiding Dangerous to Catastrophic Climate Changes," *PNAS*, 114, 10315–10323.

Yang, Z. (1999): "Should the North Make Unilateral Technology Transfers to the South? North-South Cooperation and Conflicts in Responses to Global Climate Change," *Resource and Energy Economics*, 21, 67–87.

A Algebraic Results from Two-Country Model

In this appendix, I prove that $\frac{de_2}{dp} < 0$, and that $\frac{\partial V_1}{\partial p} > 0$ and $\frac{de_2^2}{dpdT_2} > 0$ at $p = 0$, in my baseline two-country model. To begin with, consider country 2's problem in the presence of a conditional monetary subsidy p:

$$\max_{e_2} V_2 = \frac{N_2}{1-\alpha}\left[\left(1-\psi_2\left(1+0.75\eta(N_1e_1+N_2e_2)\right)^2\right)T_2\left(\delta+(1-\delta)\left(\frac{e_2}{E}\right)^\gamma\right)+p(e_2^S-e_2)\right]^{1-\alpha}.$$

This leads to the following first-order condition:

$$1.5\eta N_2\psi_2 tT_2\left(\delta+(1-\delta)\left(\frac{e_2}{E}\right)^\gamma\right)+p=\left(1-\psi_2t^2\right)T_2(1-\delta)\frac{\gamma}{E}\left(\frac{e_2}{E}\right)^{\gamma-1}.$$

Differentiating with respect to p (but leaving e_1 fixed), I find:

$$\begin{aligned}1.5\eta N_2\psi_2T_2&\left[0.75\eta N_2\left(\delta+(1-\delta)\left(\frac{e_2}{E}\right)^\gamma\right)+t(1-\delta)\frac{\gamma}{E}\left(\frac{e_2}{E}\right)^{\gamma-1}\right]\frac{de_2}{dp}+1\\&=T_2(1-\delta)\frac{\gamma}{E}\left[-1.5\eta N_2\psi_2t\left(\frac{e_2}{E}\right)^{\gamma-1}+\left(1-\psi_2t^2\right)\frac{(\gamma-1)}{E}\left(\frac{e_2}{E}\right)^{\gamma-2}\right]\frac{de_2}{dp}\end{aligned}$$

which can be rearranged to give:

$$\begin{aligned}T_2\Bigg[1.5\eta N_2\psi_2&\left(0.75\eta N_2\left(\delta+(1-\delta)\left(\frac{e_2}{E}\right)^\gamma\right)+2t(1-\delta)\frac{\gamma}{E}\left(\frac{e_2}{E}\right)^{\gamma-1}\right)\\&+\left(1-\psi_2t^2\right)(1-\delta)\frac{\gamma(1-\gamma)}{E^2}\left(\frac{e_2}{E}\right)^{\gamma-2}\Bigg]\frac{de_2}{dp}=-1\end{aligned}$$

and it is obvious that everything in the square brackets on the left-hand side is positive, and therefore $\frac{de_2}{dp} < 0$.

To show that $\frac{de_2^2}{dpdT_2} > 0$ is then straightforward if $p = 0$: while e_2 does generally depend on T_2, the latter cancels out of the first-order condition when $p = 0$. As a result, I know that $\frac{de_2}{dT_2} = 0$ if $p = 0$, and T_2 only appears once on the left-hand side of the final equation above. Therefore, differentiating with respect to T_2 and cancelling out the term in square brackets above, I have:

$$\frac{de_2}{dp} + T_2 \frac{de_2^2}{dpdT_2} = 0$$

and therefore $\frac{de_2^2}{dpdT_2} > 0$ if $p = 0$. If $p > 0$, there will be a further term of ambiguous sign in the latter expression, reflecting the impact of the positive effect of T_2 on e_2 in the presence of the subsidy ($\frac{de_2}{dT_2} > 0$ if $p > 0$).

Finally, I want to show that $\frac{\partial V_1}{\partial p} > 0$ at $p = 0$ and the optimized Stackelberg value of e_1. I begin with the maximization problem for country 1:

$$\max_{e_1,p} V_1 = \frac{N_1}{1-\alpha} \left[\left(1 - \psi_1 \left(1 + 0.75\eta(N_1 e_1 + N_2 e_2)\right)^2\right) T_1 \left(\delta + (1-\delta)\left(\frac{e_1}{E}\right)^\gamma\right) - \frac{pN_2}{N_1}(e_2^S - e_2) \right]^{1-\alpha}$$

which leads to the following first-order condition for p:

$$\frac{\partial V_1}{\partial p} = N_1 \left[\cdot\right]^{-\alpha} \left[-1.5\eta N_2 \psi_1 t \frac{de_2}{dp} T_1 \left(\delta + (1-\delta)\left(\frac{e_1}{E}\right)^\gamma\right) - \frac{N_2}{N_1}(e_2^S - e_2) + \frac{pN_2}{N_1}\frac{de_2}{dp} \right]$$

where $[\cdot]$ represents the terms in square brackets in the maximization problem. Obviously, if $p = 0$, the final term vanishes, and the second term is also equal to zero given $p = 0$ and the optimized Stackelberg value of e_1, since in that case $e_2 = e_2^S$. Therefore, if $p = 0$, $\frac{\partial V_1}{\partial p} > 0$.

B Sensitivity Analyses

In this appendix, I start from the baseline 2-country model and show that results are similar with a smaller value of α; I also present results for a scenario in which country 2 has incorrect beliefs about their value of ψ_2.

B.1 Smaller α

The results in section 5 rely on a value of 1.45 for the inequality aversion parameter (or coefficient of relative risk-aversion) α, as in Nordhaus (2018). However, a variety of values have been used for this parameter in the past, and so here I consider a smaller value of $\alpha = 1$, in which case the per-person utility function corresponds to $U(c_i) = \log(c_i)$. This does not change the calibration at all, because each country's utility is a monotonic positive transformation of c_i regardless of the value of α, so the rest of the parameters are unchanged from the baseline model.

Table 8 presents the policy results in this case. The Stackelberg equilibrium is, obviously, equivalent to that from the baseline model, except that the utility values are different due to the altered functional form. In column (2), the emissions optimum features higher emissions in country 1 and lower emissions in country 2 than in the baseline model, because the redistributive motive for lower emissions in country 1 is weaker when inequality aversion is

weaker; logarithmic utility is the point at which optimal emissions per capita are in fact equal across countries. The benevolent-rich-country scenario in column (3) also features higher emissions in country 1 and lower emissions in country 2 than in the baseline model. Finally, the optimal subsidy policies in columns (4) and (5) are essentially identical to baseline, because the poor country's response to the subsidy is not affected by the value of α, and the rich country's goal of maximizing consumption is also independent of α. However, the welfare gain from both subsidy policies are actually larger when α is smaller, because the reduction in country 2's emissions corresponds more closely to the optimal change than it did when α was larger.

Table 8: Optimal Outcomes & Policies with $\alpha = 1$

	(1)	(2)	(3)	(4)	(5)
Variable	Stackelberg	Optimum	Benevolent	Subsidy	Technology
country 1 emissions per capita	10.560	3.743	3.502	10.960	7.831
country 2 emissions per capita	5.619	3.743	6.193	1.983	5.018
country 1 income per capita	39.532	41.114	39.129	41.804	40.104
country 2 income per capita	5.366	5.376	5.395	5.448	5.378
country 1 utility	5.468	5.526	5.453	5.551	5.489
country 2 utility	13.777	13.792	13.821	13.901	13.796
combined welfare	19.244	19.318	19.273	19.452	19.285
temperature increase	4.000	2.761	3.719	2.581	3.564
subsidy as % of GDP				0.018	0.004
welfare gain as % of GDP		0.008	0.003	0.022	0.004

Notes: I use the same units for population, emissions per capita, and GDP per capita as in Table 1.

B.2 Incorrect Beliefs About ψ_2

In the analysis of the baseline model, I assumed that the values of the climate damage parameters ψ_1 and ψ_2 were common knowledge, and that each country chose their privately-optimal emissions accordingly. However, the resulting values of $\psi_1 = 0.0071$ and $\psi_2 = 0.0013$ imply that the rich country is relatively more affected by climate change, which does not correspond to the literature's standard assumption that developing countries are more affected by climate change; see the 2010 RICE model parameters at Nordhaus (2010), for example.

One possible explanation for this divergence is that my model assumes that each country sets domestic policies optimally, internalizing the climate externality at the domestic level, so that only the international externality remains. However, perhaps the developing country finds it more difficult to set optimal domestic climate policy, and thus in the Stackelberg equilibrium acts as if their damage parameter is lower than it really is. Therefore, in the current appendix, I consider the possibility that country 2's damage parameter ψ_2 is larger than the baseline value, and for simplicity I assume that it is in fact equal to $\psi_1 = 0.0071$. Once again, this requires no modification to the model's calibration, but it requires us to distinguish between equilibria of the model in which country 2 uses a correct or incorrect

value of ψ_2.

Specifically, I can solve for the Stackelberg equilibrium, the benevolent-rich-country outcome, and the subsidy policies under the assumption that country 2 acts according to the incorrect value of ψ_2 – but calculating their true final utility using the correct ψ_2. And I can also solve for each of these outcomes under the assumption that country 2 instead corrects their internal externality and acts according to the correct value of ψ_2, as well as solving for the emissions optimum using the correct ψ_2. The results from each of these scenarios can be found in Table 9, which presents results where country 2 remains biased (acting according to $\psi_2 = 0.0013$), and Table 10, where country 2 corrects their bias; welfare gains are measured relative to the biased Stackelberg equilibrium in each case.

Table 9: Optimal Outcomes & Policies with Incorrect Beliefs about ψ_2

	(1)	(2)	(3)	(4)	(5)
Variable	Stackelberg	Optimum	Benevolent	Subsidy	Technology
country 1 emissions p.c.	10.560	0.527	1.568	10.960	7.831
country 2 emissions p.c.	5.619	1.750	6.359	1.983	5.018
country 1 income p.c.	39.532	41.060	38.663	41.804	40.104
country 2 income p.c.	4.857	5.232	4.979	5.241	4.975
country 1 utility	-0.632	-0.621	-0.638	-0.616	-0.628
country 2 utility	-8.948	-8.654	-8.849	-8.647	-8.852
combined welfare	-9.580	-9.275	-9.487	-9.263	-9.479
temperature increase	4.000	1.735	3.645	2.581	3.564
subsidy as % of GDP				0.018	0.004
welf. gain as % of GDP		0.075	0.022	0.078	0.024

Notes: I use the same units for population, emissions per capita, and GDP per capita as in Table 1. These results represent the biased scenario in which country 2 acts according to the incorrect value of $\psi_2 = 0.0013$; welfare gains are measured relative to the biased Stackelberg equilibrium in column (1).

Moving from column (1) of Table 9 to column (1) of Table 10 generates a significant welfare gain, as country 2 realizes (or acts upon) their high true value of ψ_2 and drastically cuts their emissions; country 1 increases their emissions in response, but the final result is still a large drop in the temperature increase and a welfare gain that is equivalent to nearly a 5% increase in consumption. In column (2) of both tables, the emissions optimum features significant reductions in emissions per capita for both countries relative to the biased Stackelberg equilibrium, but particularly for country 1 due to the relatively high value of $\alpha = 1.45$. The most interesting results can be found in columns (4) and (5) of each table: in the presence of a biased value of ψ_2, both subsidies are even more useful than before, generating welfare gains of 7.8% of consumption for the monetary subsidy and 2.4% for the technology transfer, because they lower country 2 emissions from a level that is much higher than it should be. However, if the inefficiently low value of ψ_2 can somehow be corrected, as in Table 10, country 2's emissions are already reduced to a low level, and the value of either subsidy disappears from the perspective of country 1, as it becomes harder to significantly reduce country 2's emissions further.

As a result, the overall conclusion here becomes ambiguous: if country 2's climate damages are actually higher than what is consistent with their actions, the emissions-reduction

Table 10: Optimal Outcomes & Policies with Corrected Beliefs about ψ_2

Variable	(1) Stackelberg	(2) Optimum	(3) Benevolent	(4) Subsidy	(5) Technology
country 1 emissions p.c.	14.120	0.527	0.527	14.125	14.124
country 2 emissions p.c.	1.167	1.750	1.854	1.167	1.167
country 1 income p.c.	43.021	41.060	41.017	43.025	43.025
country 2 income p.c.	5.074	5.232	5.232	5.074	5.074
country 1 utility	-0.608	-0.621	-0.621	-0.608	-0.608
country 2 utility	-8.774	-8.654	-8.653	-8.774	-8.774
combined welfare	-9.382	-9.275	-9.275	-9.382	-9.382
temperature increase	2.484	1.735	1.776	2.485	2.485
subsidy as % of GDP				0.000	0.000
welf. gain as % of GDP	0.047	0.075	0.075	0.047	0.048

Notes: I use the same units for population, emissions per capita, and GDP per capita as in Table 1. These results represent the scenario in which country 2 acts according to the correct value of $\psi_2 = 0.0071$; welfare gains are measured relative to the biased Stackelberg equilibrium in column (1) of Table 9.

subsidies from country 1 are something like a substitute for efficient climate policy in country 2. If country 2 cannot adjust domestic policy to correspond to a higher level of climate damages, the subsidy from country 1 should remain large, but its value is much lower (and perhaps zero) if country 2 can solve their domestic policy problem on their own. It is also worth noting that this question becomes less relevant in the multi-country extensions in section 6, where the calibration leads to a value of ψ_2 that is larger than ψ_1 (though not ψ_0 in the version with 2 rich countries).