

Agreement Formation in International Public Goods Provision with Heterogeneous Agents

Christine Gutekunst

Kaj Thomsson

Maastricht University

November 2012

Preliminary draft. Comments welcome.

1 Introduction

The provision of international public goods, such as financial stability, international security or environmental protection, has been one of the most salient issues in international negotiations over the past decades. International public goods have the property that all countries benefit from aggregate provisions, while the costs are only carried by those actually providing. For instance, a country's provision of military forces to a peace keeping mission also benefits countries that do not participate in the mission. Or, if some countries invest in clean technology and enforce strict regulations with respect to greenhouse gas emissions, all countries benefit from reduced pollution on a global scale while only those countries investing in emission reduction pay the costs.

The potential divergence between those who carry the costs for provision and those who benefit is prevalent in public goods problems. It leads to a strong incentive to free-ride on others' contributions, and typically induces underprovision and significant welfare losses. While on a national scale governments might be able to coerce cooperation on certain issues through specific policies, the sovereignty of nations makes this impossible on an international scale. The only way to limit inefficiencies resulting from free riding incentives in international public goods provision is therefore the formation of self-enforcing agreements.

A prominent example of such a self-enforcing agreement can be seen in the Montreal Protocol on Substances that Deplete the Ozone Layer. It was first signed in 1987 and has since experienced a steady increase in membership as well as good compliance rates. On the contrary, negotiations on climate change abatement have proven quite difficult. With the United Nations Framework Convention on Climate Change, negotiations on greenhouse gas abatement received a baseline structure in 1992. Five years later, the Kyoto Protocol was adopted in an effort to restrict greenhouse gas emissions in developed countries. However, compliance rates with the Kyoto Protocol have been low, and finding a follow up agreement which includes a common standard of greenhouse gas abatement for all countries has proven difficult.

In this paper, we analyze the impact heterogeneity in underlying country characteristics has on the possibility of finding a provision standard for international public goods. In doing so, we focus on differences in provision costs, endowment and valuation of the public good. We commence with an analysis of the direct influence of heterogeneity on negotiation outcomes. Subsequently, we take the analysis a step further in focusing on the role the political process plays in shaping the way negotiation outcomes are influenced by heterogeneity.

First, we consider the influences of heterogeneity in either the provision costs or the valuation of the public good on negotiation outcomes. We observe that a player with higher costs or a lower valuation is less likely to join an agreement. Using this observation, we can show that an increase in cost or valuation heterogeneity makes agreement formation more difficult. We furthermore derive results on the effects of level changes that impact all players. A shift towards lower costs due to the introduction of new technologies leads to a decreased incentive for a high cost player to participate, that is, it makes coalition formation more difficult. A shift towards higher valuation due to, for instance, new information about the relevance of the public good has an equivalent impact on coalition formation.

Second, we look at the impact of endowment heterogeneity on negotiation outcomes. We show that countries with large endowments are more willing to participate in negotiations. Based on this reasoning we argue that a shift towards higher endowments leads coalition formation to be easier. Subsequently, we look at the way endowment heterogeneity interacts with cost and valuation heterogeneity. Introducing endowment heterogeneity in addition to cost heterogeneity might induce a high cost country to be

more willing to participate if endowments and costs are positively correlated. Equivalently, a negative correlation between endowment and valuation can induce a low valuation country to be more likely to join.

Third, we consider the role the political process plays in shaping the way heterogeneity influences negotiation outcomes. As Stavins (2011) points out, difficulties in the formation of international agreements seem to be less induced by economical and technological feasibility questions, but rather appear to be due to political coordination problems. We therefore move beyond extending work on International Environmental Agreements by explicitly modeling the voting procedure. In our baseline model, we assume that decisions are made by majority rule, but we also consider voting rules with different distributions of political power; specifically, we compare outcomes under majority rule with weighted or unweighted voting to dictator rule results, and analyze the way the distribution of political power interacts with changes in cost and valuation heterogeneity. Among other things, we can show that the negative impact (on the possibility of agreement formation) that follows from an increase in heterogeneity is reduced if political power is more concentrated in the hands of some countries, while it is aggravated if power is concentrated in the hands of other countries.

Our work builds on, and draws together, three different strands of literature. First, the literature on public provision of public goods is relevant to our work. Public provision of public goods refers to the formation of agreements to provide public goods. A paper which considers the formation of such agreements is Kosfeld et al. (2009). Kosfeld et al. consider heterogeneity with respect to preferences. However, they do not analyze endowment or cost heterogeneity and assume a fairly simplified political process.

Second, our work builds on the literature on union formation in the context of federalism and the development of new political confederations, such as the recent work of Alesina et al. (2005). While Alesina et al. consider a setting with heterogeneity and explicitly analyze the political process, our model differs from theirs in (among other things) the assumptions about spillovers between coalition members and non-members; they focus on a scenario without such spillovers, while we consider it a crucial component of the settings we analyze. Other related papers in this strand of literature are Cremer and Palfrey (1999, 2006) and Hafer and Landa (2007).

A third basis of our model is constituted by the literature on self-enforcing inter-

national environmental agreements, which originated in the 1990s with the seminal contributions by Carraro and Siniscalco (1993) and Barrett (1994). Since, scholars in the field have looked at numerous ways of how to enhance participation and contribution incentives. However, most of this work has been dealing with homogeneous agents. Barrett (1997) is one of the first to consider heterogeneity. He looks at how benefits from cooperation can be divided between countries in a coalition and mostly uses simulations to derive results. Botteon and Carraro (2001) and Mc Ginty (2006) extend his framework to a larger number of countries, but also have to use simulations to derive results. Kolstad (2010) is one of the first to present a tractable model of agreement formation under heterogeneity. However, he limits heterogeneity to size and marginal damage and considers only two different types of countries.

We add to these strands of literature by incorporating a more general approach to heterogeneity and a more refined political process while considering a pure public good, that is, while allowing for spillovers between coalition members and non-members. We present an N -country game theoretic model of agreement formation, specifically allowing for heterogeneity over provision costs, endowment, valuation of the public good, and the distribution of political power.

2 The Model

There exist $N \geq 2$ countries which we, in slight abuse of notation, label $i = 1, \dots, N$. A country i 's preferences with respect to the provision of a public good are given by the following Cobb-Douglas function:

$$U_i = \ln \left(\frac{w_i - (1 + c_i)g_i}{n_i} \right) + a_i \ln \left(\sum_{j=1}^N g_j \right) - K \mathbf{1}_{\{g_i > 0\}} + B \mathbf{1}_{\{coalitionmember\}}$$

where w_i is endowment, c_i are costs, a_i is valuation, n_i is the number of citizens in country i , and g_i is a country's contribution. The first term of the expression denotes the utility received from private consumption, that is, from the part of the endowment that is not invested in the public good. The second part of the expression is the utility received from the aggregate amount provided, weighted by the respective country's valuation for the good. We furthermore assume that there may be additional fixed costs and benefits associated with provision and coalition membership. K is a fixed cost that arises if a country provides a positive amount to the public good, either on its own or in a coalition. This fixed cost may, for instance, capture the need

for technological investments or other start-up costs associated with provision. B is an exogenous benefit from being in a coalition. This benefit may, for instance, pick up non-modelled favorable treatment for coalition member by other insiders in other policy areas such as additional trade. It may also capture a desire by countries, or more specifically country leaders, to be held in regard by other countries, to do what is perceived to be the morally correct thing, or simply to not stand out in a negative way on the international stage. We make no attempt to distinguish between these possible intrinsic benefits of coalition membership, but simply represent their joint effect with one parameter, B .

In order to analyze how an agreement for the provision of such a public good might form, we introduce a three stage setup. In the first stage, countries decide simultaneously and independently whether they wish to enter negotiations over their contribution to a public good. If at least three countries enter, they become participants of negotiations which subsequently take place in the second stage. Participants vote on a common fraction $t \in [0, 1]$ of endowment they wish to contribute¹. In our baseline model decisions are made by majority rule. Countries vote in favor if they are indifferent between forming an agreement or not. Participants of negotiations cannot leave during the second stage and non-participants cannot enter. If an agreement forms the participants of negotiations turn into members of the agreement. In the third stage countries contribute depending on whether an agreement formed and, if so, depending on whether they are members of that agreement. Countries adhere to the contributions they agreed to during negotiations.

We assume that K is such that countries have no incentive to provide if they are not members of a coalition. To be more precise, we first define U_i^{no} as the utility achieved by player i when no one is providing: $U_i^{no} = \ln(\frac{w_i}{n_i})$. We then define U_i^a as the utility achieved by player i when it alone provides and does so in an optimal way without taking K into consideration; that is, $U_i^a = \ln\left(\frac{w_i - (1+c_i)g_i^a}{n_i}\right) + a_i \ln(g_i^a)$, where $g_i^a = \arg \max_{g_i} \left\{ \ln\left(\frac{w_i - (1+c_i)g_i}{n_i}\right) + a_i \ln(g_i) \right\}$. Using these definition, we make the following assumption about K :

¹An alternative approach could be to allow for different fractions of endowment being contributed by different countries. However, we will follow Barrett (1992) who suggests that a uniform percentage reduction of abatement can be seen as a focal point in climate negotiations. We thus limit our analysis to agreements on a common fraction t of endowment.

$$K > \max \{U_1^a - U_1^{no}, \dots, U_N^a - U_N^{no}\}$$

Now, note that the utility increase the player with the greatest incentives to provide can achieve by providing is largest when no one else is providing. Hence, if that player has no incentive to provide in this setting it also has no incentive to provide in any other setting. This assumption thus implies that any player outside of a coalition will not choose to provide a positive amount of the public good.

The three stage setup we employ matches the agreement formation process outlined in most of the literature on international environmental agreements (Wagner 2001) and international public goods provision (Kosfeld et al. 2009). However, we introduce an explicit voting procedure and thereby add more detail to the political process described in the second stage. This helps us to analyze the impact of heterogeneity in costs, endowments and valuation on agreement formation and the way this underlying heterogeneity interacts with the political structure.

3 Homogeneous countries

We start by considering the case of homogeneous countries, that is, we assume all countries to have equal cost, endowment and valuation. If a coalition forms, it provides according to all countries' optimal provision amount. This follows because all voting participants have the same optimal t , so there is no disagreement and the agreed upon t trivially equals all countries' optimal t . As endowments are assumed to be the same for all countries, also the actual optimal provision amount is the same across all countries. We now look at how many countries can form a coalition, such that this coalition is stable.

Proposition 1. *If all countries are homogeneous and $B - K$ is not too small, there exists a unique number of countries that can form a stable coalition, defined by $\bar{S} \geq S^* \geq \underline{S}$.*

There exists an upper bound \bar{S} on internal stability, that is, an upper bound on the area for which no country in a coalition would have an incentive to leave the coalition. At the same time, there exists a lower bound \underline{S} on external stability, that is, a lower bound on the number of countries in a coalition for which no non-member would have an incentive to join. These bounds coincide in the sense that the smallest integer fulfilling $S \geq \underline{S}$ is at the same time the largest integer fulfilling $S \leq \bar{S}$. Thus, $\bar{S} \geq S^* \geq \underline{S}$

defines the unique stable coalition size.

The bounds \bar{S} and \underline{S} are independent of costs and endowment, but depend on the valuation of the respective public good and on the additional benefit B received if being in a coalition. If $B - K$ is larger, the stable coalition size is larger. Both boundaries are increasing in a , that is, the higher the good is valued the more countries can form a stable coalition in equilibrium.

A special case of this proposition is the formation of a grand coalition, that is, a coalition including all countries. A grand coalition forms if the number of countries is lower than the upper bound for internal stability, that is $N \leq \bar{S}$. In that case all countries prefer participating in a coalition over not doing so and a coalition with $S = N$ forms.

Corollary 1. *If $N \leq \bar{S}$ and $B - K$ not too small, a grand coalition with $S = N$ forms.*

4 Heterogeneous countries

4.1 Cost heterogeneity

We now introduce heterogeneity over costs. Countries might differ with respect to their costs of provision. That is, while one country might be able to reach some provision amount fairly easily given the production technology it uses, the same provision might cause much higher costs for another country. These differences are important to take into consideration as they largely influence countries' willingness to engage in provision.

We assume costs to be uniformly and equidistantly distributed. This allows us to order countries from highest costs to lowest costs. As preferences are single-peaked, countries can also be ordered according to their optimal t , given by

$$t_i = \frac{a}{(1+a)(1+c_i)}.$$

Larger costs induce the fraction of endowment a country wishes to contribute to decrease. As the policy space is linear and countries' preferences are single-peaked, the median voter theorem holds. If a coalition forms, it thus provides $G_C = t_M Sw$ where t_M is the fraction of endowment the median country would like to provide, and Sw is the aggregate endowment of coalition members. In the case of an even number of

countries joining negotiations, t_M is an average of the two countries around the median position.

The country with the highest costs in a conjectured coalition is the least likely to be willing to be a member of that coalition. That follows from two different aspects: a private consumption effect and a political effect. The political effect describes the change in the median t if a country does not participate. If a country with a $t > t_M$ were to not participate the new median t , denoted by \tilde{t}_M , would be lower than the old median t , denoted by t_M . Thus, the fraction contributed by the coalition would be lower. If, however, a country with $t < t_M$ does not participate, then $\tilde{t}_M > t_M$. That is, the fraction contributed by the coalition is larger. Thus, incentives to participate resulting from the political effect are larger for countries with $t > t_M$. The private consumption effect describes the decrease in private consumption which occurs due to being in the coalition. The larger a country's costs in comparison to median costs, the larger its decrease in private consumption due to contributing. Thus, the country with the highest costs incurs the largest positive change in the private consumption effect when leaving the coalition. Overall, the incentive to participate is thus smaller for countries with $t < t_M$.

Based on this reasoning, whenever countries have an incentive to form a coalition, one possible coalition forming is one comprising the countries with the lowest costs. That follows directly from the argument that a country with larger costs has a lower incentive to be in a coalition, while a country with lower costs has a larger incentive to be in a coalition. Such a coalition includes all countries from the one with lowest costs up to some threshold above which all countries are non-members, that is, it includes countries with adjacent cost levels. We call such a coalition of countries with adjacent cost levels a contiguous coalition. In the following we focus our analysis on this type of coalition.

We now analyze the impact of a median-preserving increase of cost heterogeneity on coalition formation, that is, we focus on settings where the coalition median stays the same before and after a change of heterogeneity. A median preserving increase of heterogeneity might occur either as an increase on only one side of the coalition median or as an increase on both sides. An increase in heterogeneity holds for the whole population, that is, it is not restricted to members of the coalition. The effect of an increase in heterogeneity on one side of the median depends on the side the increase occurs

on: an increase of heterogeneity on the low-cost side induces the political effect to be stronger, and an increase on the high cost side leads to a larger decrease in private consumption. Both of these effects lead coalition formation to be more difficult. An increase in heterogeneity on both sides of the median combines both effects.

Proposition 2. *A median preserving increase in the degree of cost heterogeneity makes it harder to reach an agreement.*

This proposition says that the existence of an equilibrium in which a coalition forms depends on the degree of heterogeneity between countries. If heterogeneity is sufficiently large, coalition formation becomes impossible.

Given that result on the relevance of the variation in costs, we now turn to the analysis of variation-preserving level changes in costs. That is, we look at how, for instance, a new technology that makes overall production of the public good cheaper influences coalition formation. For the case of climate change, new advances in clean energy technologies might pose an example of such a change in cost levels. If the cost level of public good production decreases, the private consumption effect is larger, that is, a player has a larger incentive not to participate. At the same time, the political effect of not being member of a coalition is larger, which creates an additional incentive not to participate. Thus, while one might on first sight expect a cost decrease to have a positive effect on coalition formation, it turns out that it actually makes coalition formation more difficult.

Corollary 2. *A decrease in the level of production costs of the public good makes it harder to reach an agreement.*

4.2 Valuation heterogeneity

Countries might have fairly different valuations of a public good. For instance, with respect to climate change some countries are likely to be more impacted than others and, at the same time, some countries seem to have a higher intrinsic value they attach to sustainable behavior. Similar considerations hold true for the case of Ozone depletion and other related international public goods. Differences along dimensions such as geographic position, economic development or general social norms prevalent in a country's culture are likely to lead to quite different valuations for a public good.

We distinguish a private consumption effect, a political effect and a coalition size

effect. As in the case of cost heterogeneity, the private consumption effect of being in a coalition is always negative, creating an incentive to not participate in a coalition. It is solely determined by the median valuation and becomes more negative if the median valuation is larger. The coalition size effect is always positive, that is, it gives an incentive to be in a coalition. In opposition to the dynamics of cost heterogeneity, however, the sign of the political effect switches depending on whether country i 's valuation is larger or smaller than the median valuation. If $a_i < a_M$ the effect is negative, while it is positive if $a_i > a_M$. In order to ensure that the combined political and coalition size effects yield a positive incentive to join the coalition, we impose the following assumption:

$$\text{A1:} \quad \frac{S}{S-1} \geq \frac{\frac{\tilde{a}_m}{a_m + 1}}{\frac{\tilde{a}_m + 1}{a_m}}$$

We assume that the a_i 's are closely packed and $a_i \in (0, \infty)$ in order to ensure that A1 holds for the broadest range of coalition sizes possible.

Given that the a_i 's are uniformly and equidistantly distributed, countries can be lined up according to their valuations. Under A1, the country with the lowest valuation in any given coalition is the one who receives the lowest utility from participating. The lower a player's valuation the lower the positive impact of the combined political and coalition size effects, while the negative impact of the private consumption effect does not vary with the respective player's own valuation. Thus, the lower a country's own valuation the higher is the weight of the negative private consumption effect relative to the positive combined political and coalition size effects, creating stronger incentives not to participate. That is, the country with the lowest valuation is the least likely to participate in negotiations.

Based on this argument, we can see that if coalition formation is possible, one coalition that may form is one which encompasses the countries with the highest valuations down to some cut off value. That is, under A1 there exists a contiguous coalition whenever coalition formation is possible according to the underlying parameters. As in the case of cost heterogeneity, we focus our analysis on this type of coalition.

We then turn to the analysis of a median-preserving increase in valuation heterogeneity. We again focus on settings where the coalition median stays unchanged, and a heterogeneity increase holds true for the whole population. The size of the political effect

depends crucially on the degree of heterogeneity. If there is an increase in heterogeneity for those countries with valuations above the median, the political effect increases, while the valuation of the country with the lowest valuation stays unchanged. That is, the country with the lowest valuation has a stronger incentive not to participate than before the increase. If there is an increase in heterogeneity for those countries with valuations below the median, the political effect stays the same as before the increase. However, the own valuation of the lowest valuation country decreases, leading the combined political and coalition size effect to be less positive. The country with the lowest valuation in a given coalition therefore has a lower incentive to participate than before the increase. To analyze the impact of an increase in heterogeneity on both sides of the median, we combine the effects of a one-sided increase. As each of these effects yields a lower incentive to participate for the player with the lowest valuation, the combined effect on participation incentives is negative.

Proposition 3. *A median-preserving increase in the degree of valuation heterogeneity leads coalition formation to be more difficult.*

We now turn to the analysis of shifts in valuation levels. For instance, new information becoming available or a general increase in awareness might lead to such a shift in valuation. A distribution-preserving shift in the valuation level of its members leads a given coalition to provide more because the median t is larger than before. However, it might also decrease participation incentives for the player with the lowest valuation. The larger median t leads the private consumption effect to be more negative. The interior political effect is less negative. The own valuation of the lowest-valuation country is now higher though. As the interior political effect gets multiplied with the own valuation of the lowest-valuation country, the overall effect is negative. Thus, the lowest-valuation country has a stronger incentive not to participate.

Corollary 3. *A shift towards higher valuation of the public good leads coalition formation to be more difficult.*

4.3 Endowment heterogeneity

Differences in endowment of countries constitute the most visible form of heterogeneity on the international stage. We assume endowment to be equidistantly distributed across countries. Differences in endowment do not have any direct impact on the fraction of endowment a country is willing to contribute. However, they do influence the overall amount a country contributes and through that lead to different externalities

caused by countries participating or not. For a country with a large endowment the utility difference between being member of a coalition and not being member is larger, that is, being in a coalition is relatively more profitable for a rich country than for a poor country. Thus, a country with a larger aggregate endowment is more likely to be in an agreement than one with a lower endowment.

This reasoning allows for a clear statement on the effect of a median-preserving increase in endowment heterogeneity. Those countries above the median have a higher endowment after such an increase, leading them to be more willing to participate. Those countries below the median, however, have a lower endowment, leading them to be less likely to participate. The country with the lowest endowment is thus less likely to be a member of the coalition after a median-preserving increase of endowment heterogeneity, leading coalition formation to be more difficult.

Along the same lines of reasoning, we can make a statement about the way in which a shift in income influences coalition formation. A shift towards higher income means that every country possesses a higher endowment. That leads every individual country to be more willing to participate. Thus, a shift towards higher income leads coalition formation to be easier.

Proposition 4. *An increase in income for all countries leads coalition formation to be easier.*

4.3.1 Endowment and costs

We now analyze the effects which result from combining endowment and cost heterogeneity. Endowment and costs can be positively or negatively correlated, depending on the specific scenario. A positive correlation means that countries with higher average costs have a higher endowment, while a negative correlation means that countries with higher average costs have a lower endowment. A country with a high endowment has higher average costs than a country with a low endowment if marginal costs are increasing, while it has lower average costs if marginal costs are decreasing. A positive correlation between endowment and costs is thus equivalent to increasing marginal costs, and a negative correlation is equivalent to decreasing marginal costs.

If endowments are equal across countries, we have seen that a country with higher costs is less likely to participate. If endowments are heterogeneous such that a coun-

try with higher costs has a higher endowment than countries with lower costs, that country will be more willing to join. If on the other hand heterogeneity is such that a country with higher costs has a lower endowment, that country will be less willing to join. That is, a positive correlation between costs and endowment has a positive effect and a negative correlation has a negative effect on coalition formation.

Proposition 5. *An increase in the degree of endowment heterogeneity makes coalition formation less difficult if marginal costs are increasing in endowment. It makes coalition formation more difficult if marginal costs are decreasing in endowment.*

4.3.2 Endowment and valuation

We now consider the way endowment and valuation heterogeneity interact. In general, a larger endowment leads a player to be more likely to participate. This effect also holds in the case of additional valuation heterogeneity. While a lower valuation as such leads a player to be less willing to participate, a higher endowment can counter this effect. We can show that, if a low valuation player has a small endowment, that leads to a lower incentive to participate, while having a larger endowment leads a low valuation player to be more willing to be in an agreement.

Proposition 6. *An increase in the degree of endowment heterogeneity makes coalition formation less difficult if valuation is decreasing in endowment. It makes coalition formation more difficult if valuation is increasing in endowment.*

5 The political process

We have so far analyzed the way cost, valuation and endowment heterogeneity influence coalition formation, and we have looked at the way endowment heterogeneity interacts with cost or valuation heterogeneity. Now we turn to the question how heterogeneity over costs, valuation and endowment interact with the political process. We argue that endowment as a measure of economic power constitutes a valuable proxy for political power. That is, we see endowment as a measure for a country's relevance on the international stage: a large endowment country is likely to have a larger weight in international negotiations than a country with a small endowment. While potentially neglecting some other aspects that might cause differences in political weights, such as for instance historical relevance, this approach reflects much of the political heterogeneity observed on the international stage.

Heterogeneity in political power might make agreement formation more or less difficult, depending on the correlation between weights and costs or valuation. We therefore consider in detail the way heterogeneity over political power interacts with heterogeneity over these underlying parameters.

5.1 Costs and political power

If weight is larger for countries with high cost, it holds that $t_{M_{\text{weighted}}} < t_M$. That is, $t_{M_{\text{weighted}}}$ is closer to a high cost country's optimal t than t_M . This induces a high cost country which participates in a coalition to be better off in the case with weighted voting than in the case without. If a country with high costs opts against participating, on the other hand, it causes the median to change more, leaving it worse off not participating than what it is in the unweighted voting case. Thus, a high cost country is more likely to be in a coalition under weighted voting than under unweighted voting.

If weight is larger for countries with low costs, it holds that $t_{M_{\text{weighted}}} > t_M$. That makes a large cost country worse off in case it participates under weighted voting than what it would be under unweighted voting. If it does not participate, on the other hand, it would be better off than under unweighted voting. Not participating induces a lower change in the median t , that is, not participating causes the coalition provisions to fall less than under unweighted voting.

Proposition 7. *Weighted voting makes coalition formation easier if countries with high costs have greater political power, while it makes coalition formation more difficult if countries with low costs have greater political weight.*

We now look at an extreme case, where one country in a coalition is able to dictate the t which will be provided. In such a case the weight of all other countries in a coalition is zero, meaning that they have no influence at all on the provision outcome of the coalition. We first consider a setting where a high cost country, that is a country with costs above the median costs, is a dictator in the coalition. Then a country with high costs is better off if it is in a coalition than under equal voting weights because the chosen t is closer to its own optimal t . If it is outside of a coalition, the political effect stays unchanged and thus utility is not different from a setting with equal voting weights. Overall the existence of a high cost dictator therefore increases incentives of a high cost country to be in the coalition. Now consider a low cost dictator, that is

a country with costs below the median. In such a setting a country with high costs is worse off if it joins the coalition because the chosen t is further away from its own optimal t . If it does not participate, however, it is better off because the coalition provides a higher t than under equal voting weights.

Corollary 4. *If there is a high cost dictator in a given coalition, the number of members of the coalition weakly increases, while t weakly decreases compared to a setting with equal voting weights. If there is a low cost dictator in a coalition the number of coalition members weakly decreases while t weakly increases.*

We consider now how the effect of an increase in cost heterogeneity depends on the distribution of voting weights. In the case of a high cost dictator, an increase in heterogeneity of those countries with lower costs than the median has no effect. An increase of heterogeneity for the countries with costs higher than the median leads to a larger incentive to join the coalition. The overall effect of an increase in heterogeneity if there exists a high cost dictator is thus positive, leading to a larger incentive for high cost countries to be in a coalition. If there exists a low cost dictator, both an increase of heterogeneity for those countries with lower costs than the median and for those with higher costs leads to a lower incentive to participate for countries with high costs. The overall effect of an increase in heterogeneity if there exists a low cost dictator is thus negative, leading to a lower incentive for high cost countries to be in a coalition.

Corollary 5. *The negative effect of an increase in cost heterogeneity is less harsh if political weights are more concentrated on high cost countries. If political weight is more concentrated on low cost countries, the negative effect of an increase in heterogeneity is harsher than under equally distributed political weights.*

5.2 Valuation and political power

In order to consider the combined effect of valuation heterogeneity and political power heterogeneity, we compare the utility of a low valuation country in the case of equally distributed voting weights to the utility it receives under weighted voting. If valuation and voting weights are positively correlated, that is, if countries with a higher valuation have larger political power, a low valuation country is worse off if it participates under weighted voting, but better off if it does not participate under weighted voting. If valuation and voting weights are negatively correlated, that is if countries with a higher valuation have lower political power, the country with the largest incentive not to

participate is better off if it participates under weighted voting, but worse off if it does not participate under weighted voting compared to the unweighted case.

Proposition 8. *Weighted voting makes coalition formation easier if countries with low valuation have greater political power, while it makes coalition formation more difficult if countries with high valuation have greater political power.*

In the extreme case where all political weight is focused on one player, incentives to join a coalition for a low valuation country may be larger or smaller, depending on the position of the decisive player. If there exists a high valuation dictator, a low valuation country is worse off if it participates but better off if it does not do so than in a case with equal voting weights. If there exists a low valuation dictator in a coalition, a low valuation country is better off if it participates but worse off if it does not do so than in a case with equal voting weights. Thus, the country with the lowest participation incentives has a higher incentive to participate if there exists a low valuation dictator but a lower incentive if there exists a high valuation dictator.

Corollary 6. *If there is a low valuation dictator in a given coalition, the number of members of the coalition weakly increases, while t weakly decreases compared to a setting with equal voting weights. If there is a high valuation dictator in a coalition the number of coalition members weakly decreases while t weakly increases.*

We now consider how the existence of a dictator influences the impact an increase in heterogeneity has on coalition formation. If there is a high valuation dictator a heterogeneity increase for countries with valuations above the median does not change incentives to join a coalition for a low valuation country. A heterogeneity increase for countries with valuations below the median lowers incentives to join. This reasoning holds true equivalently for a median valuation country having dictatorial power.

Corollary 7. *The negative impact of an increase in valuation heterogeneity is stronger if political weights are more concentrated on medium or high valuation countries.*

6 Conclusion

We show that if countries are homogeneous, there exists a unique number of countries that can form a stable agreement in equilibrium. If we allow for heterogeneity over costs or valuation, agreement formation becomes more difficult. We see that the country with the highest costs or the lowest valuation is the least likely to be willing

to be a member of a coalition. The existence of an equilibrium in which a coalition forms thus depends on the degree of heterogeneity between countries. If heterogeneity is sufficiently large, coalition formation may become impossible.

Further, we look at the impact of endowment heterogeneity on agreement formation. We first consider pure endowment heterogeneity and then analyzed the way it interacts with cost or valuation heterogeneity. Endowment heterogeneity leads countries with a large endowment to be more willing to participate in negotiations. The effect of an interaction between endowment heterogeneity and cost or valuation heterogeneity depends on whether high cost countries are associated with a high or a low endowment and, respectively, whether high valuation countries are associated with a high or a low endowment. We show that introducing endowment heterogeneity in addition to cost or valuation heterogeneity might make a large cost country more willing to participate if it leads to an increase in the large cost or low valuation country's endowment.

Cost and valuation heterogeneity have a direct negative effect on agreement formation. Heterogeneity over costs or valuation causes countries to differ with respect to the fraction of endowment they are willing to contribute to a public good. Such differences in willingness to contribute lead agreement formation to be fairly difficult, especially if differences are substantial. Endowment heterogeneity, on the other hand, has a more indirect effect. It does not change countries direct willingness to contribute, but has an indirect effect which is based on the externalities caused by a country. A larger country causes larger positive externalities for other countries if it provides. If it does not provide, however, the aggregate amount provided decreases more than if a smaller country decided against participating. In that sense, a larger country can be seen as more decisive for the outcome of negotiations. That idea can be reflected through weighted voting.

We conduct a detailed analysis of heterogeneity over political power. We model political coordination through introducing an explicit voting procedure into the analysis of negotiations, where political heterogeneity is introduced through different voting weights. We are able to show that the political process plays a crucial role in agreement formation. If a country with a large endowment has a larger political weight on the international stage, negotiation outcomes are shifted in favor of high endowment countries. Introducing different voting weights might lead a country with larger costs or lower valuation to be more likely to be in an agreement than under unweighted

voting, given that it also has a larger endowment. Thus, if political power is positively correlated with costs or negatively correlated with valuation, coalition formation is less difficult under weighted voting than under unweighted voting.

In considering the extreme case where all weight is concentrated on one player we show that, if there is a high cost dictator in a coalition, the number of members of the coalition weakly increases. If there is a low cost dictator in a coalition the number of coalition members weakly decreases. Equivalently, if there is a low valuation dictator in a given coalition, the number of members of the coalition weakly increases while, if there is a high valuation dictator in a coalition, the number of coalition members weakly decreases.

The impact of an increase in heterogeneity depends largely on the specific political situation. Political weights that are more concentrated on high cost countries lead the impact of an increase in cost heterogeneity to be less harsh than under equally distributed political weights. Equivalently, political weights that are more concentrated on low cost countries lead the impact of an increase in cost heterogeneity to be harsher than under equally distributed political weights. In the case of valuation heterogeneity, the negative impact of an increase in heterogeneity is stronger if political weights are more concentrated on high or median valuation countries.

APPENDIX

Proposition 1: Homogeneity. The proposition follows from considering internal and external stability of a coalition. A coalition provides $G_C = St^*w$ where S is the number of members of the coalition and

$$t^* = \frac{a}{(1+a)(1+c)}$$

is the agreed upon fraction of endowment that is contributed. A country's utility from being in a coalition is then

$$U_C = \ln\left(\frac{w}{n} - \frac{a}{1+a} \frac{w}{n}\right) + a \ln\left(\frac{Saw}{(1+a)(1+c)}\right) + (B - K),$$

while being outside of a coalition yields utility

$$U_O = \ln\left(\frac{w}{n}\right) + a \ln\left(\frac{Saw}{(1+a)(1+c)}\right).$$

Comparing the utility a country in a coalition with S members receives to the utility a country being outside of an $S - 1$ coalition receives, yields

$$\ln\left(1 - \frac{a}{1+a}\right) + a \ln(Sw) + (B - K) \geq a \ln((S - 1)w). \quad (1)$$

Comparing the utility a country outside of a coalition with S members receives to that which a country participating in a coalition of $S + 1$ receives, yields

$$\ln\left(1 - \frac{a}{1+a}\right) + a \ln((S + 1)w) + (B - K) \leq a \ln(Sw). \quad (2)$$

A coalition is internally stable for $S \leq \bar{S}$ and externally stable for $S \geq \underline{S}$. As the largest integer fulfilling $S \leq \bar{S}$ is at the same time the smallest integer fulfilling $S \geq \underline{S}$, the unique stable coalition size is defined by $\bar{S} \geq S^* \geq \underline{S}$.

For the case where $B - K = 0$ this can be expressed as

$$\frac{1}{1 - \left(\frac{1}{1+a}\right)^{\frac{1}{a}}} \geq S^* \geq \frac{1}{(1+a)^{\frac{1}{a}} - 1} \quad \text{for } a \geq 4.$$

□

Proposition 2: Cost heterogeneity. Consider internal and external stability:

$$\begin{aligned} \ln\left(1 - \frac{1 + c_i}{1 + c_M} \frac{a}{1+a}\right) + a \ln\left(\frac{a}{(1+c_M)(1+a)}\right) + a \ln(Sw) + (B - K) &\geq \\ a \ln\left(\frac{a}{(1+c_{\bar{M}})(1+a)}\right) + a \ln((S - 1)w) &\end{aligned} \quad (3)$$

and

$$\begin{aligned} a \ln\left(\frac{a}{(1+c_M)(1+a)}\right) + a \ln(Sw) &\geq \\ \ln\left(1 - \frac{1 + c_i}{1 + \tilde{c}_M} \frac{a}{1+a}\right) + a \ln\left(\frac{a}{(1+\tilde{c}_M)(1+a)}\right) + a \ln((S + 1)w) + (B - K) &\end{aligned} \quad (4)$$

If heterogeneity on the low cost side increases, the distance between countries with respect to costs increases. That induces the political effect to be stronger: the difference between t_M when the country with the highest costs in a conjectured coalition participates and \tilde{t}_M when it does not participate increases. That makes it more profitable for the country to not participate.

If heterogeneity on the high cost side increases, the difference between the median player and the player with the largest costs increases. Thus, $\frac{1+c_i}{1+c_M}$ increases, which leads the utility received from private consumption to decrease more due to contributing. This makes a country more likely not to participate.

In the case of a two-sided heterogeneity increase both effects work together. \square

Proposition 3: Valuation heterogeneity. Consider internal and external stability:

$$\begin{aligned} \ln\left(1 - \frac{a_M}{1 + a_M}\right) + a_i \ln\left(\frac{a_M}{(1+c)(1+a_M)}\right) + a_i \ln(Sw) + (B - K) &\geq \\ a_i \ln\left(\frac{\tilde{a}_M}{(1+c)(1+\tilde{a}_M)}\right) + a_i \ln((S - 1)w) &\end{aligned} \quad (5)$$

and

$$\begin{aligned} a_i \ln\left(\frac{a_M}{(1+c)(1+a_M)}\right) + a_i \ln(Sw) &\geq \\ \ln\left(1 - \frac{\tilde{a}_M}{1 + \tilde{a}_M}\right) + a_i \ln\left(\frac{\tilde{a}_M}{(1+c)(1+\tilde{a}_M)}\right) + a_i \ln((S + 1)w) + (B - K) &\end{aligned} \quad (6)$$

An increase of heterogeneity on the right side of the median induces the difference between \tilde{a}_M and a_M to be larger and, therefore, has a negative political effect. As the relevant a_i stays unchanged, the overall effect is negative.

An increase of heterogeneity on the left side of the median leaves the difference between \tilde{a}_M and a_M unchanged and, therefore, has a neutral political effect. The relevant a_i is lower, however, and therefore the overall effect is negative on the parameter set that fulfills A1.

A heterogeneity increase on both sides of the median induces a negative political effect and a lower relevant a_i . The overall effect of an increase in heterogeneity is thus negative. \square

Proposition 4: Endowment heterogeneity. Consider internal and external stability:

$$\ln\left(1 - \frac{a}{1 + a}\right) + a \ln\left(\sum_{j \in S} w_j\right) + (B - K) \geq a \ln\left(\sum_{j \in S-1} w_j\right) \quad (7)$$

and

$$a \ln\left(\sum_{j \in S} w_j\right) \geq \ln\left(1 - \frac{a}{1+a}\right) + a \ln\left(\sum_{j \in S+1} w_j\right) + (B - K). \quad (8)$$

The difference between the utility resulting from being in a coalition and not being in a coalition is larger for a country with a larger endowment. That is, a country with a larger endowment has a larger incentive to be in the coalition. \square

Proposition 5: Combined cost and endowment heterogeneity. Consider internal and external stability:

$$\begin{aligned} \ln\left(1 - \frac{1+c_i}{1+c_M} \frac{a}{1+a}\right) + a \ln\left(\frac{a}{(1+c_M)(1+a)}\right) + a \ln\left(\sum_{j \in S} w_j\right) + (B - K) &\geq \\ a \ln\left(\frac{a}{(1+c_{\tilde{M}})(1+a)}\right) + a \ln\left(\sum_{j \in S-1} w_j\right) & \end{aligned} \quad (9)$$

and

$$\begin{aligned} a \ln\left(\frac{a}{(1+c_M)(1+a)}\right) + a \ln\left(\sum_{j \in S} w_j\right) &\geq \\ \ln\left(1 - \frac{1+c_i}{1+c_{\tilde{M}}} \frac{a}{1+a}\right) + a \ln\left(\frac{a}{(1+c_{\tilde{M}})(1+a)}\right) + a \ln\left(\sum_{j \in S+1} w_j\right) + (B - K) & \end{aligned} \quad (10)$$

Increasing marginal costs lead to higher average costs for a large endowment country. Decreasing marginal costs lead to lower average costs. If a country with high average costs has a high endowment it is more likely to participate than under equal distribution of endowments, while if it has a low endowment it is even less likely to participate than it would be under income homogeneity. \square

Proposition 6: Combined valuation and endowment heterogeneity. Consider internal and external stability:

$$\begin{aligned} \ln\left(1 - \frac{a_M}{1+a_M}\right) + a_i \ln\left(\frac{a_M}{(1+c)(1+a_M)}\right) + a_i \ln\left(\sum_{j \in S} w_j\right) + (B - K) &\geq \\ a_i \ln\left(\frac{a_{\tilde{M}}}{(1+c)(1+a_{\tilde{M}})}\right) + a_i \ln\left(\sum_{j \in S-1} w_j\right) & \end{aligned} \quad (11)$$

and

$$\begin{aligned}
& a_i \ln \left(\frac{a_M}{(1+c)(1+a_M)} \right) + a_i \ln \left(\sum_{j \in S} w_j \right) \geq \\
& \ln \left(1 - \frac{\tilde{a}_M}{1 + \tilde{a}_M} \right) + a_i \ln \left(\frac{\tilde{a}_M}{(1+c)(1+\tilde{a}_M)} \right) + a_i \ln \left(\sum_{j \in S+1} w_j \right) + (B - K) \quad (12)
\end{aligned}$$

In any given coalition, the country with the lowest valuation is the least likely to participate. A larger endowment leads that country to be more willing to participate, while a smaller endowment leads that country to be less likely to participate. That is, an increase in endowment heterogeneity makes coalition formation easier or harder, depending on the direction of correlation between endowment and valuation. \square

Proposition 7: Cost heterogeneity and weighted voting. We first consider the case where weight is larger for countries with large costs and then the case where weight is larger for countries with low costs.

1. Weight is larger for countries with large costs: Assume there exists some coalition S . It holds that $t_{M_{weighted}} < t_M$. For countries with a $t^* \geq t_{M_{weighted}}$ not participating would yield a lower utility than participating. Thus, these countries are more likely to be members of a coalition than to be non-members. We can therefore focus on countries with $t^* < t_{M_{weighted}}$. As $t_{M_{weighted}} < t_M$, $t_{M_{weighted}}$ is closer to these countries' optimal t than t_M . Thus, a high cost country which participates in a coalition is better off in the case with weighted voting than in the case without. At the same time, the political effect is stronger if a high weight country does not participate, that is, it causes the median to change more if it decides against participating. A country with a larger weight is thus worse off not participating than in the unweighted voting case, while it is better off participating than in the unweighted voting case. Therefore, a large cost country is more likely to be in a coalition than under unweighted voting.

2. Weight is larger for countries with small costs: Assume there exists some coalition S . It holds that $t_{M_{weighted}} > t_M$. As above, for countries with a $t^* \geq t_{M_{weighted}}$ not participating would yield a lower utility than participating. Thus, these countries are more likely to be members of a coalition than to be non-members. We can therefore again focus on countries with $t^* < t_{M_{weighted}}$. As $t_{M_{weighted}} > t_M$, $t_{M_{weighted}}$ is further away from these countries' optimal t . That makes a large cost country worse off in

case it participates under weighted voting than what it would be under unweighted voting. If it does not participate, on the other hand, it would be better off than under unweighted voting because it induces a lower increase in the median t through not participating. \square

Proposition 8: Valuation heterogeneity and weighted voting. We first consider the case where weight is larger for countries with low valuation and then the case where weight is larger for countries with high valuation.

1. Weight is larger for countries with low valuation: Assume there exists some coalition S . It holds that $t_{M_{weighted}} < t_M$. For countries with a $t^* \geq t_{M_{weighted}}$ not participating would yield a lower utility than participating. Thus, these countries are more likely to be members of a coalition than to be non-members. We can therefore focus on countries with $t^* < t_{M_{weighted}}$. As $t_{M_{weighted}} < t_M$, $t_{M_{weighted}}$ is closer to these countries' optimal t than t_M . Thus, a low valuation country which participates in a coalition is better off in the case with weighted voting than in the case without. At the same time, the political effect is stronger if a high weight country does not participate, that is, it causes the median to change more if it decides against participating. A country with a larger weight is thus worse off not participating than in the unweighted voting case, while it is better off participating than in the unweighted voting case. Therefore, a low valuation country is more likely to be in a coalition than under unweighted voting.

2. Weight is larger for countries with high valuation: Assume there exists some coalition S . It holds that $t_{M_{weighted}} > t_M$. As above, for countries with a $t^* \geq t_{M_{weighted}}$ not participating would yield a lower utility than participating. Thus, these countries are more likely to be members of a coalition than to be non-members. We can therefore again focus on countries with $t^* < t_{M_{weighted}}$. As $t_{M_{weighted}} > t_M$, $t_{M_{weighted}}$ is further away from these countries' optimal t . That makes a low valuation country worse off in case it participates under weighted voting than what it would be under unweighted voting. If it does not participate, on the other hand, it would be better off than under unweighted voting because it induces a lower increase in the median t through not participating. \square

References

- [1] Alesina, A. and Angeloni, I. and Etro, F.: *International Unions*, American Economic Review, Vol.95, No.3, pp. 602 - 615 (2005)
- [2] Barrett, S.: *The Problem of Global Environmental Protection*, Oxford Review of Economic Policy, Vol.6, No.1, pp. 68 - 79 (1990)
- [3] Barrett, S.: *Self-enforcing International Environmental Agreements*, Oxford Economic Papers, Vol. 46, pp. 878 - 894 (1994)
- [4] Barrett, S.: *Heterogeneous International Environmental Agreements*, in Carraro, C. (ed.): *International Environmental Negotiations: Strategic Policy Issues*, Edward Elgar (1997)
- [5] Botteon, M. and Carraro, C.: *Environmental Coalitions with Heterogeneous Countries: Burden Sharing and Carbon Leakage*, in Ulph, A. (ed): *Environmental Policy, International Agreements, and International Trade*, Oxford University Press, New York (2001)
- [6] Carraro, C. and Siniscalco, D.: *Strategies for the International Protection of the Environment*, Journal of Public Economics, Vol. 52, No. 3, pp. 309 - 328 (1993)
- [7] Cremer, J. and Palfrey, T.R.: *Political Confederation*, American Political Science Review, Vol.93, No.1, pp. 69 - 83 (1999)
- [8] Cremer, J. and Palfrey, T.R.: *An Equilibrium Voting Model of Federal Standards with Externalities*, Journal of Public Economics, Vol.90, pp. 2091 - 2106 (2006)
- [9] Hafer, C. and Landa, D.: *Public Goods in Federal Systems*, Quarterly Journal of Political Science, Vol.2, pp. 253 - 275 (2007)
- [10] Kolstad, C.: *Equity, Heterogeneity and International Environmental Agreements*, The B.E. Journal of Economic Analysis and Policy, Vol.10, No.2 (2010)
- [11] Kosfeld, M., A. Okada and A. Riedl: *Institution Formation in Public Goods Games*, American Economic Review, Vol. 99, No. 4, pp. 1335 - 1355 (2009)
- [12] McGinty, M.: *International Environmental Agreements Among Asymmetric Nations*, Oxford Economic Papers, Vol.59, pp. 45 - 62 (2007)

- [13] Stavins, R.N.: *The Problem of the Commons: Still Unsettled after 100 Years*, American Economic Review, Vol.101, No.1, pp. 81 - 108 (2001)
- [14] Wagner, U.J.: *The Design of Stable International Environmental Agreements*, Journal of Economic Surveys, Vol. 15, No. 3, pp. 377 - 411 (2001)