The stability and effectiveness of climate coalitions: A comparative analysis of multiple integrated assessment models

Kai Lessmann · **Ulrike Kornek** · **Valentina Bosetti** · **Rob Dellink** · **Johannes Emmerling** · **Johan Eyckmans** · **Miyuki Nagashima** · **Hans-Peter Weikard** · **Zili Yang**

Abstract We report results from a comparison of numerically calibrated game theoretic integrated assessment models that explore the stability and performance of international coalitions for climate change mitigation. We identify robust results concerning the incentives of different nations to commit themselves to a climate agreement and estimate the extent of greenhouse gas mitigation that can be achieved by stable agreements. We also assess the potential of transfers that redistribute the surplus of cooperation to foster the stability of climate coalitions. In contrast to much of the existing analytical game theoretical literature, we find substantial scope for self-enforcing climate coalitions in most models that close much

Kai Lessmann · Ulrike Kornek
Potsdam-Institute for Climate Impact Research (PIK), PO Box 60 12 03, D-14412 Potsdam, Germany E-mail: lessmann@pik-potsdam.de

Valentina Bosetti
Department of Economis, IEFE, and IGIER, Università Bocconi and Fondazione Eni Enrico Mattei (FEEM), Milan, Italy

Rob Dellink
Environmental Economics and Natural Resources Group, Department of Economics, Wageningen University, The Netherlands

Johannes Emmerling
Fondazione Eni Enrico Mattei (FEEM) and Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), Italy

Johan Eyckmans
KU Leuven – Center for Economics and Corporate Sustainability (CEDON), Warmoesberg 26, B-1000 Brussels, Belgium

Miyuki Nagashima
Research Institute of Innovative Technology for the Earth (RITE), Kyoto, Japan

Hans-Peter Weikard
Environmental Economics and Natural Resources Group, Department of Economics, Wageningen University, The Netherlands

Zili Yang
Department of Economics, State University of New York at Binghamton, Binghamton, New York, USA

of the abatement and welfare gap between complete absence of cooperation and full cooperation. This more positive message follows from the use of appropriate transfer schemes that are designed to counteract free riding incentives.

Keywords coalition stability, international environmental agreements, numerical modeling, transfers

1 Introduction

1.1 Motivation

This paper reports results from a comparison of models that explore international coalitions for climate policy. These models investigate the incentives of different nations to commit themselves to a climate agreement and examine the extent to which climate agreements are stabilized when participation is driven only by the self-interest of its members.

International climate policy suffers from the adverse incentive structure of public good provision, and it is well known that non-cooperative behavior results in under-provision of such goods. How much a climate coalition improves upon this dilemma depends on the costs and benefits of the individual nations. It is particularly dependent on their heterogeneity and whether nations can compensate each other, i.e. the existence of transfer schemes. In this paper we investigate this for the first time using an ensemble of five numerical models of climate coalition formation. Numerical models give particularly valuable insights beyond those of their analytical counterparts, when the analysis depends on regional heterogeneities in costs and benefits, quantitative estimates (i.e. their order of magnitude), or detailed representations of reaction functions.

The models in this study are diverse both in their modeling approaches and in the data sources used for calibration, representing a range of estimates for the costs and benefits of real-world regions and their dynamics. Consequently, the models do not necessarily agree in their assessment of the stability of specific coalitions.

The aim of this study is to make the differences in the underlying assumptions of costs and benefits, and their implications for key model results, transparent. We quantify these differences using new, model independent metrics, and we subsequently ask what these metrics can contribute to understanding diverging coalition stability results and the effect of transfers across the models.

Early theoretical investigations of coalition stability have often begun with the assumption of identical (symmetric) players. This has resulted in a pessimistic assessment of the scope of self-enforcing agreements to improve cooperation on international environmental issues. In their seminal paper Carraro and Siniscalco (1993) find stable coalitions to be generally small in size. While stable coalitions may be large in the model setup of Barrett (1994), this only holds if the gains from cooperation are small. Much of the ensuing research has investigated this dilemma and ways around it.[1]

The theoretical potential for transfer payments to increase participation and environmental performance of a coalition has been long known (Carraro and Siniscalco 1993). Yet when real world heterogeneity is included, transfers that do not consider strategic implications are unlikely to improve an agreement and may even damage its success (Weikard et al 2006; Nagashima et al 2009). In contrast, transfers designed specifically to make cooperation more attractive than the free-riding alternative greatly improve the success of coalitions

---

[1] For a survey of this literature see Finus (2008), Eyckmans (2012) and Benchekroun and Long (2012).

(Carraro et al 2006; Nagashima et al 2009; Weikard 2009). Similarly, McGinty (2007) shows the beneficial effect of such transfers within the modeling framework of Barrett (1994).

Regardless of the design of transfers, the asymmetries in the costs and benefits of climate policy are shown to be important for the achievement of cooperation. Barrett (2001) stresses that asymmetry of players complements the effect of transfers; together, asymmetry and transfers may well improve cooperation. Weikard (2009) shows that higher levels of participation under transfers that target the free-riding incentive are spurred by stronger asymmetry among the coalition members, including the possibility of full cooperation. Fuentes-Albero and Rubio (2010) confirm this and show that differences in marginal damages (rather than abatement costs or the level of damages) are key to this result. Recently Karp and Simon (2013) obtained comprehensive results showing how the functional specification of costs and benefits of abatement impacts on the equilibrium coalition size.

While these studies firmly establish a more optimistic prospect for cooperation and highlight the importance of heterogeneity, the degree of asymmetry often remains a conceptual assumption. Notable exceptions are Carraro et al (2006) and Nagashima et al (2009), which rely on integrated assessment models to quantify asymmetries.

To improve the understanding of real-world asymmetries, both the mechanics and the calibration of the models are of central importance. However, the uncertainties are large, both within each model (most prominently concerning model parameters) and between models (concerning model structure). Previously, the issue of how uncertainty and the prospect of future learning about climate change impacts affects international cooperation, has been addressed by including uncertainty in the structure of the coalition formation game (Kolstad and Ulph 2008, 2011; Finus and Pintassilgo 2013). In this paper, uncertainty is reflected in the diversity of assumptions of the models compared. The strength is threefold: it makes uncertainty more transparent; it helps identify robust results across modeling assumptions and parameterizations; and it enables us to learn from the differences.

Our contribution is a better understanding of the well-known cooperation failure, particularly in the heterogeneous setting provided by these numerical models. In addition, we identify transfers and assess their magnitude and direction when used as a tool to enhance cooperation.

1.2 International Climate Agreements

Central to this study is the concept of "self-enforcing agreements" or "coalition stability". A climate coalition is a subset of the world's regions that agree to cooperate on climate change mitigation policies. More specifically, we assume that within the coalition, climate change is addressed in an efficient manner, i.e. in a manner that maximizes coalitional welfare.[2] The coalition adopts a joint climate policy and interacts with the remaining regions as a single player. Each player is assumed to act selfishly with respect to the others.

Coalition formation is hence modeled as a game with two stages: a "participation stage game," where players decide to either become members of the coalition or to remain outside; and an "emission stage game" where players choose economic strategies which (directly or indirectly) determine the emissions and abatement of greenhouse gases. Strategies therefore result from a combination of the membership decision in the first stage and the economic

---

[2] Most models implement Pareto-efficiency through maximization of the utilitarian sum of individual welfare per region. MICA computes Pareto-efficient strategies by solving a competitive equilibrium on international commodity markets with full internalization of the climate change externality.

strategies of the players in the second stage of the game.[3] This setup is similar to the Partial Agreement Nash Equilibrium (PANE) concept introduced by Chander and Tulkens (1995).

The equilibrium concept in stage 1 is cartel stability (d'Aspremont and Gabszewicz 1986) in all models. A coalition is considered "stable" if it satisfies two conditions. First, it is internally stable, meaning that no member is willing to leave the coalition. Second it is externally stable, meaning that all non-members prefer to remain singletons.[4] Formally, any given coalition S is stable if for the payoff of player i facing coalition S, $\pi_i(S)$, we have:

$$\pi_i(S) \geq \pi_i(S \setminus \{i\}) \text{ for all } i \in S, \quad \pi_j(S) \geq \pi_j(S \cup \{j\}) \text{ for all } j \notin S \qquad (1)$$

Cartel stability allows us to analyze the whole range of partial cooperation. Note that the notion of "core stability" was simultaneously developed for this class of models (Chander and Tulkens 1995). Core stability derives directly from classical cooperative game theory. Its specific properties are quite different from those of cartel stability so its analysis is left for future research but the tools that facilitate this model comparison are also applicable to core stability, see Kornek et al (2014) and footnote 13. We refer the interested reader to Bréchet et al (2011) who take a comparative look at both concepts.

Thus building solely on cartel stability, we aim to explore the drivers of cooperation. In particular we also examine the effects of transfer schemes on the prospects for cooperation. To investigate transfers, we employ the concept of potential internal stability (PIS, as defined in Carraro et al 2006). A coalition is said to satisfy the PIS property if a transfer scheme exists that can redistribute payoffs within the coalition such that the coalition is internally stable. Formally, PIS requires the existence of a vector of transfers τ_i with $\sum_i \tau_i = 0$ such that:

$$\pi_i(S) + \tau_i \geq \pi_i(S \setminus \{i\}) \text{ for all } i \in S \qquad (2)$$

Note that simple addition of transfers is only appropriate in models with transferable utility. For models that do not assume transferable utility but feature a transferable commodity (e.g. consumption), transfers can be implemented at the commodity level. Here, a transfer scheme consists of a redistribution of the commodity between regions for each time period.[5]

The rest of the paper is structured as follows. An overview of the different integrated assessment models used in this analysis is given in section 2. Section 3 focuses on the role of transfers. Section 4 summarizes results and concludes.

2 Characterization of the Models

In all models in this study, economic strategies are derived with respect to climate change mitigation from a dynamic optimization framework. Each model combines the two level game described above with an integrated climate economy model in the second stage. These models are solved numerically, and thus their parameters need to be calibrated. Reflecting

---

[3] The strategy set in stage 2 depends on the specific features of the models. These range from choosing abatement directly to the indirect approach of choosing to invest in a broad variety of capital stocks (including energy and abatement technologies). See the appendix for details.

[4] This notion of cartel stability was first applied to international environmental agreements by Hoel (1992), Carraro and Siniscalco (1993), and Barrett (1994).

[5] We apply the procedure thus outlined for the models MICA, WITCH and RICE, for details see Kornek et al (2014).

uncertainty in knowledge over climate change economics, calibration varies with the different empirical estimates of costs and benefits of climate change abatement used for the specific models. More detailed information on the models can be found in appendix A; model equations are documented in Online Resource 1.

In the Stability of Coalitions model (STACO) each region's monetary payoff equals regional benefits (avoided damages) less costs of abatement (Nagashima et al 2009, 2011). The time dependence of both benefits and costs is calculated through an approximation of a Ramsey type growth model (Ramsey 1928). STACO does not model the consumption/savings decision endogenously. Instead it uses exogenous baseline projections for economic development and carbon emissions.

The other four models determine each region's payoff through the use of dynamic, long-term, perfect foresight, Ramsey-type optimal growth models which determine savings behavior and abatement endogenously.

The first multi-region economic-climate model following this approach was RICE (Nordhaus and Yang 1996) and we use an updated version in this study (Yang 2008). RICE examines the relationship between economic growth, greenhouse gas emissions and climate change by explicitly modeling the stock of emissions in the atmosphere and the resulting global temperature.

Closely related to this is the ClimNeg World Simulation (CWS) model, a modified version of the RICE model, updated with new data on its cost and damage parameters (Eyckmans and Tulkens 2003; Eyckmans and Finus 2006; Bréchet et al 2011). As one of its prominent distinctions, the payoff in CWS is in monetary units rather than an abstract utility metric, which facilitates an implementation of transfer payments at the payoff level.

The Model of International Climate Agreements (MICA) follows the same economic framework as RICE but with different assumptions about costs and benefits. Formally, its main distinction is to include the international goods markets (Lessmann et al 2009; Lessmann and Edenhofer 2011; Kornek et al 2013).

The aforementioned models rely on stylized abatement cost functions to model emissions reductions. In contrast, WITCH incorporates an explicit representation of mitigation options, particularly in the energy system (Bosetti et al 2006). With this level of detail comes a trade-off: the higher computational complexity necessitates the use of selected coalitions.

2.1 Non-cooperative and fully cooperative solutions

The modeling assumptions, model structures, and data sources of the five models reflect quite different views of the world economy and its development (see Table 1). A key difference between the models is the way they value the present against the future. For monetary values such as abatement costs or climate change damages, this is determined by each models' endogenous interest rate. Simple Ramsey models suggest that this interest rate depends on the pure rate of time preference and, if the intertemporal elasticity of substitution is strictly positive, on consumption growth. This follows from the Keynes-Ramsey rule $\dot{c}/c = 1/\eta \, (r - \rho)$ with per capita consumption c and $\dot{c}$ its derivative with respect to time, and two preference parameters: the elasticity of marginal utility η and pure rate of time preference ρ. At the interest rate r households are indifferent between one unit now or $(1 + r)$ units later. Table 1 (top section) shows how models differ in their preference parameters.

Together with assumed projections of technological progress, these preference parameters determine the growth rates of economic output (Table 1, middle section) which range

from 1.2 percent to 2.1 percent per year over the first century. The pure rate of time preference is highest for MICA, WITCH, and RICE at 3 percent, which has a direct consequence on the interest rates in these models (around 5 percent).[6] For these models, all costs and benefits occurring in the future will be discounted at this higher rate. For STACO, the pure rate of time preference is lower (at 1.5 percent) but the (exogenous) assumption of relatively strong growth in the coming decades leads to a high initial discount rate, especially for emerging economies, which declines over time to values of around 3 percent, giving an average of 4.2 percent. Finally, in CWS the interest rate, at 1.5 percent, is the same as the pure rate of time preference, which is the lowest of all the models.

A non-cooperative equilibrium is one where no coalition forms. In the non-cooperative equilibrium, greenhouse gas emissions are of the same order of magnitude in all models, with moderately lower values in MICA and RICE. As regional damages are internalized in the non-cooperative equilibrium, emissions are lower compared to a hypothetical case without damages. We measure these emission reductions relative to a no damage business-as-usual scenario (BAU) and find that they are of comparable magnitude in 4 out of 5 models (about 10 percent of emissions), and 5 percent in RICE.

In the cooperative solution for the grand coalition, emissions are substantially lower than those in the non-cooperative equilibrium (Table 1, bottom section). This is again with the exception of RICE, which is probably due to the extent of climate change damages. In the other models, emission reductions bring down climate change damages by several percentage points in 2100. In STACO and WITCH in particular, high damages occurring in the non-cooperative equilibrium are reduced by about four percentage points. In contrast, the formation of the grand coalition in RICE leaves climate change damages almost unchanged. When we derive a metric for damages in the next sections, we will see that damage estimates are indeed relatively low in RICE.

2.2 Cost/benefit information

In this section, we introduce two metrics to characterize the severity of climate change damages and abatement costs in the models, both globally and on the regional scale. Perhaps the most intuitive metric would be to compare marginal cost functions and marginal damage functions. Unfortunately this information is not easy to extract from, or make comparable between, the models. Instead we propose alternative metrics based on model output rather than assumptions regarding functional forms and parameter values.

Measuring marginal climate change damages

Figure 1 compares aggregate discounted damages in the non-cooperative equilibrium across the models. The figure highlights that the damage calibration is low in RICE, also reflected in the relatively low carbon price in the cooperative solution; the average price in 2100 is 208 $/tC compared to a range from 369 to 966 $/tC in the other models (Table 1).[7]

In order to compare the marginal damages from climate change between regions for each model, we take a slightly different approach to that provided by the total damages shown in Figure 1. Instead, we say that a region has high marginal damages if the carbon

---

[6] To be precise, the pure rates of time preference are constant in RICE and MICA, but diminish in WITCH from an initial 3 percent to 2 percent over the course of a century.

[7] Technically, the STACO model only considers benefits from abatement, and payoffs do not depend on the level of damages.

Table 1 Modeling assumptions and key numbers of non-cooperative equilibrium and fully cooperative solution

Modeling assumptions	MICA	STACO	CWS	WITCH	RICE
Initial year	2005	2011	2000	2005	2000
Time Horizon (years)	190	95[a]	330	145	245
Number of regions	11	12	6	13	6
Pure rate of time preference (percent)	3.0	1.5 [b]	1.5	3.0	3.0
Elast. of marginal utility	1.0	1.0 [b]	0.0	1.0	1.0
Non-cooperative equilibrium	**MICA**	**STACO**	**CWS**	**WITCH**	**RICE**
Mean GDP growth rate[c]	2.06	1.97	1.54	1.56	1.24
Mean interest rate[c,d]	5.26	4.17	1.50	5.35	4.98
GHG emissions (GtC) 2015-2100	1516	1827	1754	1963	1404
Non-cooperative GHG reductions (percent)[e]	9.8	12.1	10.2	13.0	5.0
Mean GHG intensity (GtC/tn$)	0.12	0.14	0.13	0.15	0.13
Climate change damage in 2100 (percent)[f]	5.8	7.8	3.2	9.3	1.6
Carbon price 2100: reg. mean ($/tC)	12	89	49	38	8
Cooperative solution	**MICA**	**STACO**	**CWS**	**WITCH**	**RICE**
GHG emissions (GtC) 2015-2100	953	984	1094	1122	1242
Climate change damage in 2100 (percent)[f]	3.8	4.0	1.9	4.9	1.5
Carbon price 2100: reg. mean ($/tC)	369	966	529	858	208
Carbon price growth rate to 2100 (percent)	1.90	1.69	0.90	1.02	1.02

a In STACO, climate change damages are considered for a 300 year period.

b STACO derives the interest rate for discounting payoffs using the Keynes-Ramsey rule to ensure consistency with a logarithmic utility function and a pure rate of 1.5 percent

c Using a time horizon of 100 years

d The endogenous rate at which monetary values are discounted in the model, averaged over regions and time

e Emission reduction in the non-cooperative equilibrium relative to a business-as-usual scenario without climate change damages

f Damages are reported as a share of 2100 economic product
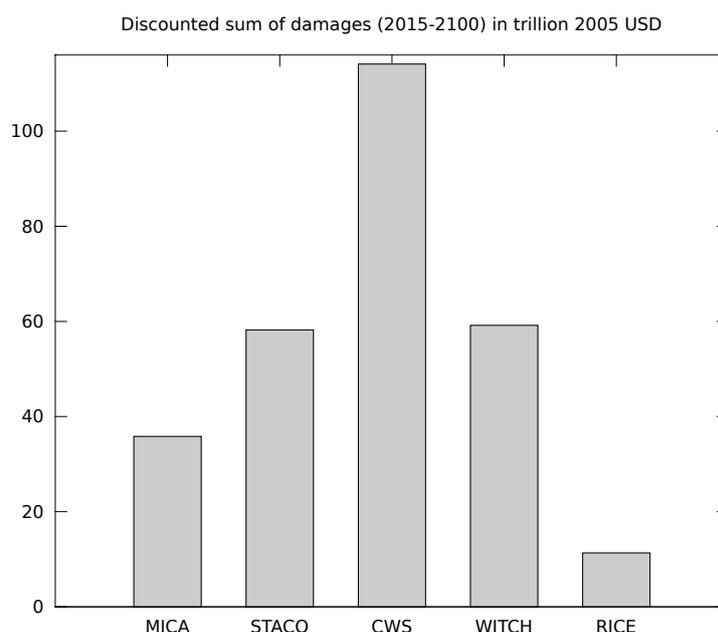
Fig. 1 Aggregate damages 2015–2100 in the non-cooperative equilibrium in trillion US$, discounted at the model specific discount rates

price of the grand coalition is significantly reduced when the region in question leaves. The impact on the coalitional carbon price is a good indicator of how much the region is affected by climate change; this region's damages are reflected in the carbon price if and only if the region is part of the coalition.

We take the discounted sum of this carbon price difference for each region in the grand coalition compared to the sub-coalition when the one region in question is leaving, and average the differences. Taking the average only matters when coalitions do not establish a uniform carbon price, which is the case for WITCH and RICE. We normalize this metric to the maximal difference over all regions (Figure 2).

Each model yields a different distribution of marginal damages across regions. In particular there is no agreement between the models about the region incurring highest marginal damages. This reflects the fact that the assumptions made by the models about the regional distribution of damages differ greatly. These differences concerning the marginal damage assumptions are a primary driver of the results of the comparison.

Measuring regional marginal abatement costs

For our metric of regional abatement costs, we look at the cumulative emissions reduction in each region at a uniform carbon price. Higher emissions reductions at such a global carbon price signal a flatter marginal abatement cost curve up to this point, and hence lower abatement costs for a prescribed abatement target. Since high figures indicate low marginal costs, we talk about this metric as being the region's "abatement potential".

Technically, all models implement the common carbon price scenario by imposing the same global emissions tax trajectory under conditions of disabled climate change damages. The cumulative abatement is the absolute emissions reduction, summed over each model's time horizon.

We find the global abatement potential to be largest in case of MICA, followed by WITCH, STACO, CWS, and RICE in declining order. Only about two thirds of the abatement prescribed by MICA are achieved in RICE. Despite this, we see that the potential and hence abatement costs are of the same order of magnitude in all five models.

Figure 2 shows the abatement potential indicator for each model. The abatement potential indicator is the tax scenario normalized to the maximum abatement level over all regions. The indicator shows that China and India always rank high on abatement potential while for Japan the mitigation costs are perceived to be amongst the highest. We will use the information from this table extensively when discussing the main objectives of this paper: incentives of specific regions to join a climate agreement and the characterization of transfers. In general, one can say that the models seem to be in good agreement over their assumptions on the costs of abatement.

2.3 Incentives of regions

The incentive to remain in a coalition (or in short: incentive to stay) is defined as the payoff received as a member of a given coalition minus the payoff of being outside the coalition (i.e. as a free-riding non-member). For the following discussion, we want to structure the driving forces that determine the incentive to stay for a given region in the following way:

1. First, the benefit of joining the coalition, which is in turn influenced by:
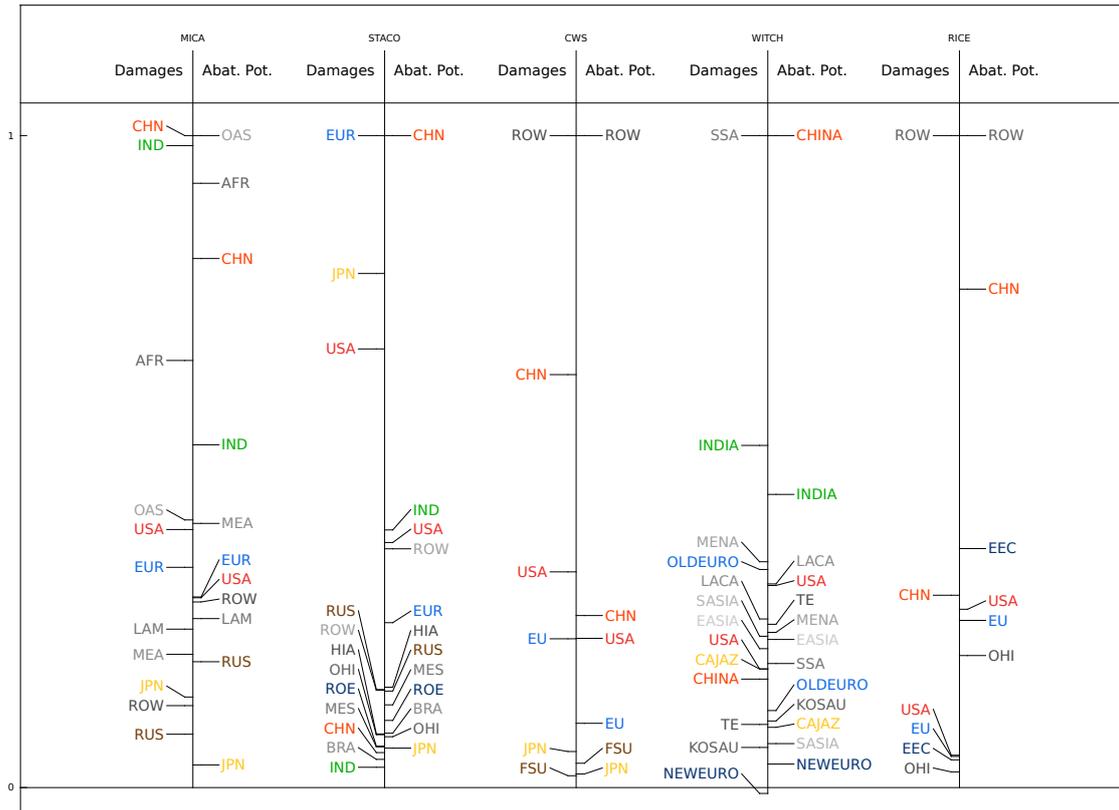
Fig. 2 Climate change damages and abatement potential indicators scaled to $[-1, 1]$. The climate change marginal damage indicator for a particular region was calculated by taking the average of the difference in discounted carbon prices of the grand coalition and the grand coalition minus the particular region at hand. This indicator was normalized relative to the maximum average difference over all regions. Abatement potential was calculated by implementing a common carbon tax trajectory for all regions in every model. The resulting abatement trajectory (measured in tons) was integrated over each model's time horizon and scaled according to the maximum abatement level. Model regions are specified in appendix.

(a) the extent of climate change damages in this region. When a region joins the coalition, any damages it incurs are henceforth internalized by all coalition members. Thus, the higher the marginal damages in the joining region, the greater the abatement of the coalition as a whole. Any region benefits from such additional abatement, and this is particularly pronounced for a region with high marginal damage.

(b) the response of non-members. For example, the free-riding non-members are likely to raise their emissions in reaction to the reduced emissions of the coalition. Such "leaked" emissions offset the abatement of the coalition and therefore reduce the benefit of joining such a coalition.

2. Second, the additional costs incurred by the region upon joining the coalition. We distinguish the abatement costs of a coalition member and other opportunity costs as follows:

(a) abatement costs are a result of the distribution of emission reductions which are determined by efficiency in abatement (i.e. the lower the marginal costs, the more a region needs to abate), and the overall ambition of the coalition, which depends on the collective marginal climate change damages of all coalition members.

(b) other opportunity costs emerge when regions are coupled through more channels than just the externality. For example, when carbon pricing affects the world demand for fossil resources, price changes in such resources will represent a cost of participating in the coalition for net exporters of such fuels.
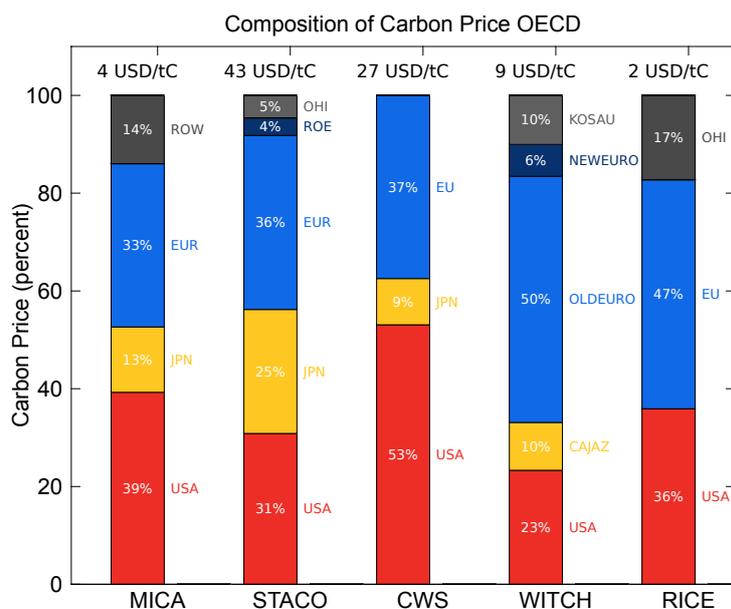
Fig. 3 Carbon price in the OECD coalition (as average net present value). Percentages indicate how climate change damages in the member regions contribute to the overall carbon price.

The extent of 1a and 2a in the models is covered by the indicators derived above and summarized in Figure 2 . We will see that often these indicators suffice to understand model behavior. Before we take a look at the incentives, we discuss how damages, abatement and leakage vary by considering the OECD coalition. For our purposes, all regions generating at least half their economic output from OECD countries are considered members of this coalition.
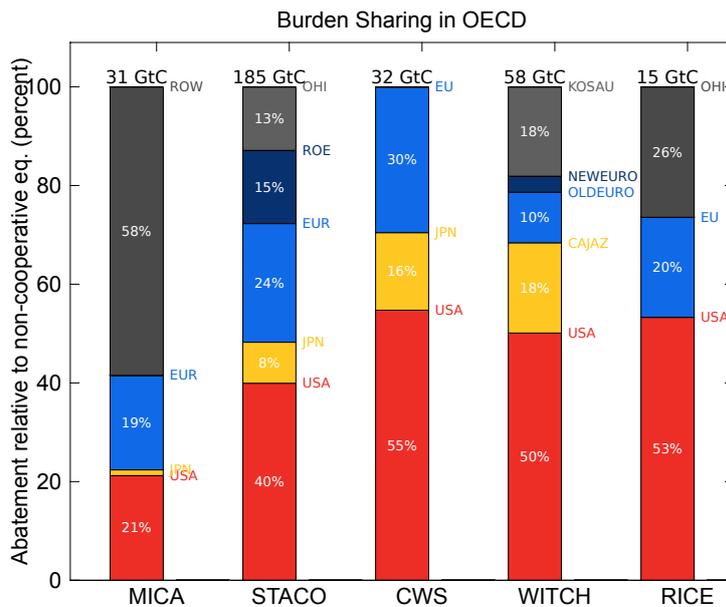
Distribution of Damages

The extent to which a region benefits from abatement is measured through the carbon price (see discussion of Figure 2). Figure 3 reports the percentage decrease in abatement price when a nation leaves the OECD coalition.[8]

The absolute level of the carbon price, given above each bar, shows the ambition in the coalition's emissions reductions across the models. In the case of the STACO model, the OECD coalition combines regions with high damages. This makes the carbon price of this coalition very high, and free-riding on this coalition's abatement very attractive. In contrast, the ambition of the OECD coalition and hence the incentive to free-ride is much lower in MICA.

There is reasonable agreement across the models that it is the USA and Europe[9] which contribute the most to damages: they both score highly in every model. The USA and Europe would therefore gain much from the abatement undertaken by the coalition. Their share of coalitional abatement must, of course, be taken into account to determine the incentive of these regions.

---

[8] Since marginal damages are not entirely flat in any of the models but STACO, this procedure is just an approximation of the decomposition of the cooperative carbon price but since the abatement of the OECD coalition is unambitious and leakage is small, the error is negligible.

[9] Specifically, we will refer to model regions EUR, EU, and OLDEURO as "Europe".

Fig. 4 Allocation of emission reductions in the OECD coalition as percentages of overall emissions reduction, the time horizon is one century.

Distribution of Abatement

There are many ways to distribute overall abatement amongst the members of climate coalitions, guided by many criteria, e.g. pragmatic, normative or incentive compatible. The default distribution in coalition models is normative: the maximization of coalition welfare.[10] We follow this approach.

Figure 4 shows that the OECD coalition's total abatement over the first century is quite different across the models, partly because the composition of the coalition is different between models. However, since differences turn out to be large even when regions are identical, we conclude that much of the variation in abatement allocation is due to different cost and benefit assumptions.

The distribution of abatement for a given country varies substantially across the models. For example, the USA share falls anywhere within the range of 20-60 percent, that of Japan between 1 and 18 percent, and for Europe within the range 10-30 percent. All models agree that the largest share of abatement ought to be achieved in the USA, often followed by Europe.[11]

Leakage Emissions

Carbon emissions are said to "leak" out of the climate coalition if non-members increase their emissions in response to the coalition's abatement. The amount of leakage depends on the sensitivity of the reaction functions of nations not in the coalition. These reaction functions depend largely on model features that determine the ways in which non-members are affected by the coalition.

---

[10] Some models use a weighted sum in the social welfare function, see model factsheets in the appendix.

[11] In MICA, the largest share falls onto the rest-of-the-world (ROW) region, which includes several non-OECD countries.
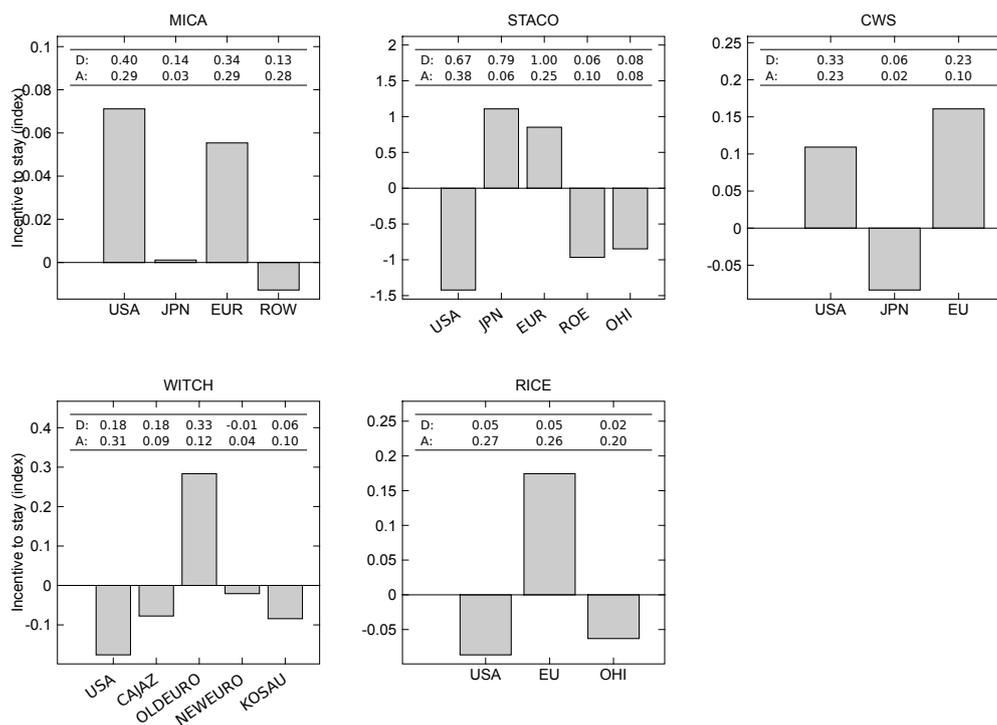
Fig. 5 Incentive to stay in the OECD coalition, calculated as the difference between inside and outside payoff and scaled to the gap in global aggregated payoff between non-cooperative equilibrium (=0) and cooperative solution for grand coalition (=100). The inset tables list the regional indicators of climate change damages and abatement potential, denoted "D" and "A", respectively.

There is zero leakage in STACO. This is a consequence of assuming constant marginal damages, which implies that abatement is chosen independently by each region. In all other models, the regions react to the abatement decisions of the others. MICA, CWS, and RICE show only very moderate leakage: leakage rates per region are less than one percent of the coalition's abatement (not shown).

Regions in WITCH show the strongest free-riding behavior in terms of leakage, with to-tal leakage rates of 16 percent of the OECD's coalition abatement (not shown). In WITCH, the coalition affects non-members through an additional channel: energy markets (see Bosetti and De Cian 2013, for details). The coalition's abatement effort drives down oil prices and free-riders increase their consumption of the carbon-intensive oil grades in particular.

Overall incentives

The interplay of all the drivers discussed above jointly shapes the incentive to join or leave a given coalition. We consider the OECD coalition and the grand coalition in turn.

Figure 5 shows the incentive to stay inside the OECD coalition for its members. If the incentive to stay was positive for all members, the coalition would be internally stable with-out transfers. Conversely, the figure shows which regions are responsible when the OECD coalition fails to be internally stable in any of the models. For easy reference, the indicators of abatement potential and climate change damages from Figure 2 are repeated in Figure 5.

While the models all agree that the OECD coalition is not internally stable, each model suggests that different regions want to leave the coalition. For example, the USA would sup-port the OECD coalition in MICA and CWS, but would not support it in STACO, WITCH, or RICE.

In general, low abatement potential and high marginal damages have a positive effect on the incentive to stay. Low abatement potential indicates a steep marginal abatement cost function and therefore a low mitigation burden. High marginal damages indicate that a region will benefit much from increased coalitional abatement.

According to this logic, the incentives for Europe appear relatively simple: in all models Europe is a coalition member characterized by relatively high marginal damages and low abatement potential. Such players have much to gain from cooperation, but have low costs as their share of the mitigation burden is small. The models therefore agree that Europe would want to remain in the OECD coalition.

This is not the case for the USA, which has a strong motivation to defect from the OECD coalition. This is because the USA's estimated abatement potential is high, which means that the USA would carry a large share of the emission reductions in the coalition. In fact three of the models find that the USA would not support the OECD coalition (STACO, WITCH, and RICE). In MICA and CWS however, the costs are more than compensated for by large benefits. In these models the USA incur the highest marginal damages of all coalition members and thus have an incentive to remain in the coalition. It also helps that in these two models the ambition level of the OECD coalition is low, resulting in a low burden for its members as most high marginal damage regions are outside the OECD coalition.

Japan is modeled as a single country region in three of the five models. The models unanimously see little abatement potential in Japan alone, mainly due to the relatively small size of its economy. Thus Japan would carry only a small burden, which makes it better off inside the coalition in MICA and STACO. In CWS, the estimated marginal damages are also very low for Japan. Japan can therefore defect (and save on abatement costs) without substantially lowering the ambition level of the coalition, which turns out to be preferred by Japan. The larger Canada/Japan/New Zealand aggregate region of WITCH incurs substantial abatement costs, tipping the balance towards defection as marginal damages are only average compared to the other OECD players.

Moving to the grand coalition, positive net incentives to stay become rare (not shown). This is a consequence of more ambitious emission reductions in this coalition, which places a larger burden on all members. The few exceptions are either of the high damage/low burden type discussed above (e.g. Japan and Europe in STACO, and Sub Saharan Africa (SSA) in WITCH) or very large players (ROW in CWS and RICE). The latter aggregate much of the world's damages and abatement potential due to their sheer size.

In WITCH, net revenues from trade in oil are part of the region's income and an additional driver of incentives. Coalitions that strongly abate emissions consume substantially less oil leading to a drop in prices. In turn, this increases the outsiders' consumption. Therefore oil-rich regions, while cutting their own oil consumption when joining the coalition, receive large revenues. Interestingly, the model shows that extraction does not change very much and the price differences are only minor. However, the pattern of consumption changes between different grades of oil, leading to increased exports of low carbon intensive grades. The top three regions showing the strongest increase in oil revenues prefer to stay in the grand coalition.[12] This effect is negative for the regions Canada/Japan/New Zealand and South Korea/South Africa/Australia.

---

[12] Model regions Middle East and North Africa (MENA), Non-EU Eastern Europe (TE), and Sub Saharan Africa (SSA).

Table 2 Stable coalitions for internal/external stability and potential internal stability.

Model	Concept[a]	Number stable		Max. size	Max. abat.[b]	Max. welf.[c]
MICA	IES	1	(0.05%)	3	0.06	0.09
	IS	54	(2.64%)	4	0.17	0.24
	PIS[e]	481	(23.50%)	6	0.31	0.47
STACO	IES	1	(0.02%)	2	0.03	0.03
	IS	23	(0.56%)	2	0.07	0.07
	PIS	2130	(52.01%)	9	0.59	0.68
CWS	IES	1	(1.59%)	2	0.67	0.77
	IS	5	(7.94%)	2	0.67	0.77
	PIS	55	(87.30%)	6	1.00	1.00
WITCH[d]	IES	1		2	0.03	0.05
	IS	1		2	0.03	0.05
	PIS[e,f]	5		4	0.17	0.38
RICE	IES	0	(0.00%)			
	IS	3	(4.76%)	2	0.03	0.06
	PIS[e,f]	7	(11.11%)	2	0.12	0.11

a Stability concepts are abbreviated IES (internal/external stability), IS (internal stability), and PIS (potential internal stability)

b The maximum global abatement achieved by a coalition is measured by the *closing the gap* indicator from 0=non-cooperative equilibrium to 1=cooperative solution for the grand coalition.

c Maximum global welfare is measured by the *closing the gap* indicator.

d In WITCH, only seven selected coalitions were analyzed.

e In order to determine the PIS-property, the maximization procedure described in Kornek et al (2014) was employed.

f For the maximization procedure, the discount-rate was held fixed at the level of the PANE-solution.

3 Transfers

3.1 Stable coalitions

In this section, we compare two stability concepts without transfers (internal/external stability, IES, and internal stability, IS) with potential internal stability (PIS).[13] Table 2 shows the number of stable coalitions for each stability concept, along with maximum coalition size, maximum abatement achieved and maximum welfare achieved. For the latter two measures, we use the "closing the gap" indicator to characterize the performance of coalitions. This indicator relates global emission reductions and welfare to the gap between the non-cooperative equilibrium – set to zero – and to the cooperative solution for the grand coalition, set to unity (cf. Eyckmans and Finus 2007).

We find coalitions that are internally/externally stable to be small and achieve little, which is in line with existing literature. CWS is an interesting exception: the best internally/externally stable coalition achieves 77 percent of the global welfare gains of the grand coalition. This is due to the very large region ROW which enables a two player coalition to abate a large share of global emissions. The best performing coalition that does not include ROW achieves a closing the gap indicator for welfare of only 21 percent.

---

[13] These are the most commonly used concepts for this set of models in previous studies. The analysis could be extended to include blocking power (or core stability) which, for the sake of brevity, we leave for future work.

Table 3 Permit allocation schemes. The permit allocation for a coalition S is determined as follows: each member i of S receives $q_{it} = \omega_{it} \cdot Q_t$ where $Q_t = \sum_{j \in S} e_{jt}$ are the available permits within S. Population, emissions and economic product are abbreviated pop, e, and y. Non-cooperative equilibrium is abbreviated NC, business-as-usual ND. The schemes are taken from Altamirano-Cabrera and Finus (2006)

Scheme	Distribution key
Egalitarian	$\omega_{it} = \text{pop}_{it} / \sum_{j \in S} \text{pop}_{jt}$
Grandfathering	$\omega_{it} = e_{i,t_0} / \sum_{j \in S} e_{j,t_0}$
Quota Nash	$\omega_{it} = e_{it}^{NC} / \sum_{j \in S} e_{jt}^{NC}$
Quota BAU	$\omega_{it} = e_{it}^{ND} / \sum_{j \in S} e_{jt}^{ND}$
Historic responsibility	$\omega_{it} = \left(e_{i,t_0}^{ND}\right)^{-1} / \sum_{j \in S} \left(e_{j,t_0}^{ND}\right)^{-1}$
Ability to pay	$\omega_{it} = (y_{it}/\text{pop}_{it})^{-1} / \sum_{j \in S} (y_{it}/\text{pop}_{it})^{-1}$
Ability to pollute	$\omega_{it} = (e_{it}/\text{pop}_{it})^{-1} / \sum_{j \in S} (e_{it}/\text{pop}_{it})^{-1}$
Energy efficiency	$\omega_{it} = (e_{it}/y_{it})^{-1} / \sum_{j \in S} (e_{it}/y_{it})^{-1}$

When we focus on internal stability alone, more coalitions are stable, and their performance improves. We highlight two observations: first, participation remains almost unchanged (with the exception of an increase from 3 to 4 players in one model, MICA); second, the performance improvement when ignoring external stability is substantial for some models, and negligible in others.

Turning to coalitions with PIS, the transfers implicit in this concept have a strong effect: the number of stable coalitions increases by 1-2 orders of magnitude. The corresponding improvement in the closing the gap indicator is also large. The CWS model even finds the grand coalition to have the PIS property. In MICA, STACO, and WITCH, PIS transfers improve the closing the gap indicator from single digit values to values roughly half that of the grand coalition (47, 68, and 38 percent, respectively). In STACO, the coalition with PIS generating the highest global welfare is not only internally stable after receiving the implied transfers but is also externally stable.[14]

Thus, the model comparison shows that transfers exist that make it possible to stabilize coalitions that substantially close the welfare gap. This is a considerably more optimistic message than the traditional conclusion derived from analytical models so far. With our multi-model approach, we conclude that this claim is robust with respect to modeling approaches and parameterizations.

3.2 Transfers and stable coalitions

The PIS transfers in the preceding section are determined by stability considerations. In contrast, transfers based on conventional burden sharing rules (Table 3) are designed to be either equitable or pragmatic. How does this departure from incentive compatibility affect the ability of conventional burden sharing schemes to induce stable coalitions?

To evaluate how burden sharing affects stability of coalitions, we convert permit allocations to monetary transfers using the carbon price of the coalition. The monetary transfers are added either to the consumption streams or payoff (in case of CWS and STACO).[15]

---

[14] Technically, this is because the STACO model is characterized by superadditivity, which means that the total worth of a group of players involved in a merger does not decrease, see Eyckmans et al (2013) for details.

[15] In two models, there is no single carbon price within the coalition (WITCH and RICE) because maximization of social welfare for the coalition balances marginal value of emissions in terms of utility but not

Table 4 How transfers affect coalition stability.

Transfer: Model	grandfathering coal[a]	Δcoal[b]	egalitarian coal[a]	Δcoal[b]	historic responsibility coal[a]	Δcoal[b]	ability to pay coal[a]	Δcoal[b]
MICA	6	-48	4	-50	5	-49	4	-50
STACO	3	-20	9	-14	0	-23	11	-12
CWS	16	11	0	-5	0	-5	3	-2
WITCH[c]	1	0	0	-1	0	-1	1	0
RICE	1	-2	0	-3	0	-3	0	-3

a Number of internally stable coalitions
b Number of internally stable coalitions relative to no-transfers
c Only selected coalition were analyzed in WITCH

In a first look at the implications of the conventional transfer schemes, we analyze how a selection of four schemes from Table 3 affects internal stability. Table 4 shows the number of internally stable coalitions under these transfers and how this number changes in relation to the scenario without transfers.

The main conclusion is that transfer schemes that were designed without coalition stability in mind have an almost unanimously adverse effect on stability. This is evident from the decrease in the number of internally stable coalitions (cf. the almost exclusively negative numbers in the column Δcoal). An exception is "grandfathering" in CWS.

In Table 5 we compare two additional statistics of the conventional transfer schemes to the PIS-transfers to investigate the poor performance of conventional transfer schemes. The first column of each model shows the share of members of coalitions with PIS where the direction of transfers coincides with PIS transfers i.e. regions that need a positive transfer are receivers. By definition, PIS transfers reach the perfect score of 100 percent and other transfers score lower.

We find that most conventional transfer schemes stay well below 100 percent for this indicator and often around 50 percent or lower. Most models find that specific transfer schemes do better; however, the models disagree on which transfer scheme performs best.

The second column displays the average flow of money between the regions across the ensemble. Models agree that the stability-enhancing PIS-transfers need relatively small flows of money. Of the other schemes, transfers for "quota BAU" and "grandfathering" are roughly in the same order of magnitude. These two schemes often also score high on the direction indicator. In the other transfers schemes, more money is transferred than is necessary for internal stability.

Thus, we have identified two problems of the conventional transfer schemes with respect to their negative effect on stability, namely the direction of the induced transfers and their magnitude.

To investigate how PIS transfers depend on the properties of coalition members, we relate properties of coalition members to the frequency with which they receive a positive transfer. The results are found in Table 6.

We find significant agreement between the models regarding the relationship between damages and transfers. The more damages a region incurs from climate change, the more likely it is that this region has a surplus to share with other members, i.e. PIS transfers will be negative. We find no significant correlation between transfers and either abatement potential or abatement per unit damage, although the results do indicate that in most cases the direction of the relationship is the same.

4 Summary and Conclusions

We have compared five different models to explore the stability and performance of international coalitions for climate change mitigation in a setting where regional heterogeneity reflects real-world asymmetries of regions. To facilitate comparison of the impact of modeling assumptions on the costs and benefits of mitigation we developed two indicators measuring, first, the regional abatement potential and, second, regional exposure to climate change damages. While the models' estimates for abatement potentials are in agreement for

---

monetary units. This is different in MICA (where international trade balances marginal utility of consumption) and CWS (which uses a linear utility function). In WITCH and RICE, we instead use the social cost of carbon for the conversion (computed as the marginal utility of carbon inside the coalition divided by the marginal utility of average per-capita consumption inside the coalition).

Table 5 Permit transfer schemes compared to PIS-transfer schemes

Transfer	MICA		STACO		CWS		WITCH		RICE	
	Direction[a]	Magnitude[b]	Direction	Magnitude	Direction	Magnitude	Direction	Magnitude	Direction	Magnitude
PIS-transfer[c]	100	0.018	100	0.532	100	8.334	100	0.239	100	0.002
egalitarian	48	0.219	56	1.537	40	18.382	23	0.383	21	0.044
historic responsibility	71	0.513	57	3.968	85	58.451	54	1.511	64	0.144
grandfathering	50	0.157	41	0.866	87	12.135	54	0.279	50	0.008
quota bau	64	0.024	54	0.195	48	3.562	46	0.096	79	0.001
ability to pay	67	0.386	75	2.770	66	25.592	8	0.542	50	0.103

a Column "Direction" gives the share for members of PIS coalitions where the specified transfer scheme had the same direccion as the PIS-transfer (in percent)

b Column "Magnitude" specifies the magnitude of discounted transfers in trillion USD, averaged over all PIS coalitions

c Transfers so that each member attains exactly her outside option payoff, that is the surplus is not distributed

Table 6 Characterization of PIS transfers with properties of players. We show the correlation coefficients of the percentage of coalitions in which a player receives positive transfers to this player's abatement potential and damages indicators and their ratio. The significance of the one-sided correlation is indicated with a "*" for the p=0.05 level and a "**" for the p=0.01 level.

	Percentage of positive transfer received		
	Abatement Potential	Damages	Damages/ Abatement
MICA	-0.502	$-0.802**$	-0.593
STACO	0.250	$-0.960**$	$-0.827**$
CWS	$-0.857*$	$-0.914*$	-0.114
WITCH	0.499	0.273	0.078
RICE	-0.014	-0.186	-0.357

key world regions, we find substantial differences in the climate change damage estimates that the models produce for certain regions. To a large extent, the differences reflect the variations in the literature sources upon which the model parameterizations are based, and therefore they reflect the uncertainty over costs and benefits of climate change mitigation in the literature (cf. Metz et al 2007).

It is therefore not surprising that the models differ in their assessment of whether certain coalitions are stable, and whether certain world regions or nations have an incentive to be members of a given coalition. A notable exception is the assessment of the EU, for which all the models unanimously attest an incentive to support a coalition of OECD countries. However, when we turn from the identity of the players to their cost-benefit characteristics in terms of the two indicators suggested in this study, the models are remarkably consistent in their predictions. We find that the indicators of a region's abatement potential and its exposure to climate change damages substantially reflect its incentives and allow us to understand its preference for or against membership in a coalition. When regional abatement potential is low (implying a steep marginal abatement cost function) or regional marginal climate change damages are high, there is a greater chance for a positive incentive to stay in a coalition.

In the absence of transfers, all models agree that stable coalitions tend to be small and achieve little, due to a lack of internal stability of larger, more ambitious coalitions. This is in accordance with the theoretical literature and therefore not surprising.

Transfers designed to minimize free-riding incentives as far as possible achieve much more: the models find that coalitions with PIS are substantially larger than internally stable coalitions and achieve about half or more of what full cooperation would achieve both in welfare and in terms of greenhouse gas abatement.

In contrast, conventional transfer schemes do not improve cooperation; they often even undermine existing stable coalitions. The reason is, of course, that conventional transfers do not reflect incentives; among other things they are frequently too large in magnitude and transfer wealth in the wrong direction, i.e. regions that need transfers to be convinced to stay in a coalition are effectively paying regions that have no incentive to defect from the coalition. Conversely, we conclude that when transfers are designed to take incentives into account, the financial flows need to be small compared to the cases, for example, of an allocation based on historic emissions.

Finally, we examine how the properties of coalition members affect the PIS transfers necessary to stabilize the coalition. We find that players with high damages tend to benefit

enough from cooperation to allow them to share some of the gains, and thus compensate those players with high abatement potential that provide the necessary mitigation.

For future research, there are several possible extensions of this analysis. First, our stability analysis focused on one particular non-cooperative concept (internal/external stability) but could be extended to alternative game theoretic stability concepts.

Second, we found that on many issues, the different models were remarkably unanimous. However, where results were substantially different it could be argued that this is caused by differences in assumptions. In this paper we preferred to represent a broad range of possible future economic dynamics. More insights into the modeling details, or the parameters causing models to diverge in their assessment could, however, be gained by a close harmonizing of baseline assumptions of the models.

Third, our analysis has stressed that assumptions on climate change damages are very influential but at the same time highly uncertain. This concerns both the data and the modeling. Additional empirical studies on the economic impact of climate change are badly needed, and the modeling of climate change damages could be improved beyond the standard assumption found in the participating models.

A Appendix: Model Factsheets

Model: MICA (Model of International Climate Agreements), PIK, Germany

Model description: Kornek et al (2013)

Model concept Multi-region optimal growth model with climate externality and international trade	**Solution method** Competitive equilibrium, full internalization of climate change damage within the coalition; implemented as non-linear optimization problem solved with a modified Negishi algorithm
Welfare concept Discounted utilitarianism in each region, joint welfare maximization with constant Negishi weights for the coalition	**Parametric specification** Pure rate of time preference $\rho = 3\%$, elasticity of marginal utility $\eta = 1$
Markets and Trade Consumption good	**Model anticipation** Perfect foresight

Number of region: 11

AFR	Sub-Saharan Africa without South Africa
CHN	China
EUR	EU-27
IND	India
JPN	Japan
LAM	All American countries except Canada and the United States
MEA	North Africa, Middle Eastern and Arab Gulf countries, resource exporting countries within the former Soviet Union, and Pakistan
OAS	South East Asia, North Korea, South Korea, Mongolia, Nepal, Afghanistan
ROW	Australia, Canada, New Zealand, South Africa and non-EU27 European states except Russia
RUS	Russia
USA	United States of America

Base year 2005	**Time horizon and step** 2005–2195, 10 years
Climate Greenhouse Gases: CO_2	**Climate change** Temperature response model Carbon dioxide concentration (ppm) Temperature change (C°)
Mitigation options Abatement cost function for CO_2 based on mitigation cost information from the REMIND model (Luderer et al 2013)	**Climate impacts** Region-specific quadratic damage function in temperature increase Damage as [%] of GDP based on Fankhauser (1995) following Finus et al (2006)
Land use —	**Resources considered** —

Model: STACO-3 (Stability of Coalitions), Wageningen University, The Netherlands

---

Model description: Nagashima et al (2011), Dellink et al (2015)

Model concept	**Solution method**
Combined game-theoretic and integrated assessment model with regional benefits (avoided damages) and abatement costs of greenhouse gas emissions	Partial agreement Nash equilibrium between signatories and singletons

---

Welfare concept: Discounted net present value of regional payoff in each region, joint payoff maximization within coalitions; no full welfare evaluation but the Keynes-Ramsey rule used for discounting payoffs is consistent with a logarithmic utility function.

Parametric specification:Pure rate of time preference $\rho = 1.5\%$; implicitly $\eta = 1$

Markets and Trade	**Model anticipation**
Carbon Trade is modelled as transfers between coalition members	Perfect foresight

---

Number of region: 12

BRA	Brazil
CHN	China
EUR	EU and EFTA (EU-27, Iceland, Liechtenstein, Norway, Switzerland)
HIA	High-income Asia (Indonesia, Malaysia, Philippines, Singapore, South Korea, Taiwan, and Thailand)
IND	India
JPN	Japan
MES	Middle Eastern countries
OHI	Other high income countries (including for example, Australia, Canada and New Zealand)
ROE	Rest of Europe
ROW	Rest of the world
RUS	Russia
USA	United States of America

Base year	**Time horizon and step**
2011	2011 - 2106, 5 years

Climate	**Climate change**
Greenhouse Gases: CO_2, CH_4, N_2O, PFCs, HFCs, SF_6, based on EPPA-5 model to calibrate the regional GHGs BAU emission paths	CO_2-e concentration (ppm) Radiative Forcing (W/m^2) Temperature change (°C)

Mitigation options: Regional abatement cost functions for GHGs are based on regional cost parameters from the EPPA model (Morris et al 2008)

Climate impacts: Benefits are calculated as the net present value of the stream of future avoided damages from a unit of abatement in the current period, taking regional GDP growth and inertia in the climate system into account (Nagashima et al 2011). This function is calibrated to a simple climate module, based on the DICE model (Nordhaus 2008), but calibrated on the EPPA-5 model (Paltsev et al 2005; Paltsev 2010). The linear global benefit function is based on estimates of climate damage by Tol (2009). Benefits are allocated across regions with a share for each region (Finus et al 2006).

Land use	**Resources considered**
—	—

Notes: Exogenous technological change (annual efficiency improvement in abatement costs) is considered to be 0.5 - 2 % per year (regionally differentiated)

Model: CWS (ClimNeg World Simulation model, Version 2.0), KU Leuven, Belgium

Model description: Bréchet et al (2011)

Model concept Multi-region optimal growth model with climate externality, maximization of coalition welfare, internalization of climate externality for coalition members	**Solution method** Nash equilibrium of carbon emission game solved by tatonnement algorithm between coalition and non-members
Welfare concept Discounted utilitarianism in each region, coalition's welfare maximization with equal welfare weights for all members	**Parametric specification** $\rho = 1.5\%$ (constant), $\eta = 0$ (linear in consumption)
Markets and Trade no trade in goods, trade in carbon emission permits (optional)	**Model anticipation** Perfect foresight

Number of region: 6

USA:	United States of America
JAP:	Japan
EU:	South East Asia
China:	China
FSU:	Former Soviet Union
ROW:	Rest of the World

Base year 2000	**Time horizon and step** 2000–2310, 10 years
Climate Greenhouse Gases: CO_2	**Climate change** 3-box model of carbon cycle (atmosphere, lower and upper ocean) CO_2-e concentration (ppm) Radiative Forcing (W/m^2) Atmospheric and ocean temperature change (C°)
Mitigation options Exogenous emission efficiency improvement over time plus region-specific abatement cost functions (power functions, exponent 2.887)	**Climate impacts** Region-specific damage function, power function of atmospheric temperature change (exponent 3.0), damage as [%] of GDP
Land use –	**Resources considered** –

Model: WITCH (World Induced Technical Change Hybrid model, version 2012), FEEM, Italy

Model description: Bosetti et al (2014)

Model concept Hybrid Optimal growth model, including a bottom-up energy sector and a simple climate model, embedded in a game theoretic setup	**Solution method** Regional growth models solved by non-linear optimization and game theoretic setup solved by tatonnement algorithm between coalitions and non-members (Nash equilibrium)
Welfare concept Discounted Utilitarianism, coalition's welfare maximization with equal welfare weights for all members	**Parametric specification** $\rho = 3\%$ decreasing, $\eta = 1$ (log of consumption)
Markets and Trade Oil	**Model anticipation** Perfect foresight

Number of region: 13
CAJAZ: Canada, Japan, New Zeland
CHINA: China, including Taiwan
EASIA: South East Asia
INDIA: India
KOSAU: South Korea, South Africa, Australia
LACA: Latin America, Mexico and Caribbean
MENA: Middle East and North Africa
NEWEURO: EU new countries, CHE, NOR oldeuro: EU old countries (EU-15)
SASIA: South Asia
SSA: Sub Saharan Africa
TE: Non-EU Eastern Europe, including Russia
USA: United States of America

Base year 2005	**Time horizon and step** 2005–2150, 5 years
Climate Greenhouse Gases: CO_2, CH_4, N_2O, HFCs, CFCs, SFs Aerosols considered: yes	**Climate change** 3-box model of carbon cycle CO_2-e concentration (ppm) Radiative Forcing (W/m^2) Temperature change (C°)
Mitigation options Abatement cost functions for non-CO2 GHGs Land use Decarbonization options in the Energy system (Renewables, Nuclear, Biomass, CCS)	**Climate impacts** Region-specific damage function with linear and power function term with exponent 2.2 in the temperature increase Damage as [%] of GDP
Land use Emissions from land use change are considered	**Resources considered** Coal, Oil, Gas, Uranium, Biomass

---

Model: RICE Regional Integrated model of Climate and the Economy, SUNY Binghampton, USA

Model description Yang (2008)

Model concept	**Solution method**
Multi-region Ramsey-type growth model with joint production of GHG emission that causes climate externality	the non-cooperative Nash equilibrium; cooperative solutions under various assumptions of incentive compatibilities; coalition solutions (called "hybrid" Nash equilibria in Yang (2008))
Welfare concept	**Parametric specification**
Discounted sum of regional utility functions.	Pure rate of time preference $\rho = 3\%$, elasticity of marginal utility $\eta = 1$
Markets and Trade	**Model anticipation**
-	Perfect foresight

Number of region: 6
CHN   China
EEC   Eastern European countries and the former Soviet Union
EU   European Union
OHI   Other high-income countries
ROW   Rest of the world
USA   United States of America

Base year	**Time horizon and step**
2000	2000-2245 (5 years)
Climate	**Climate change**
CO_2 emissions and other exogenously set GHG emissions	the Schneider box model
Mitigation options	**Climate impacts**
Mitigation cost functions based on Nordhaus and Yang (1996) and updated with Yang (2008) which contains updates provided by Nordhaus	Climate damage functions based on Nordhaus and Yang (1996) and updated with Yang (2008) which contains updates provided by Nordhaus
Land use	**Resources considered**
exogenously set	Availability of fossil fuel resources at global level has been checked implicitly

References

Altamirano-Cabrera J, Finus M (2006) Permit trading and stability of international climate agreements. Journal of Applied Economics 9(1):19–47

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

Barrett S (2001) International cooperation for sale. European Economic Review 45(10):1835–1850

Benchekroun H, Long N (2012) Collaborative environmental management: A review of the literature. International Game Theory Review 14(4):1240,002

Bosetti V, De Cian E (2013) A good opening: the key to make the most of unilateral climate action. Environmental and Resource Economics 55:44–56

Bosetti V, Carraro C, Galeotti M, Massetti E, Tavoni M (2006) WITCH: A World Induced Technical Change Hybrid model. The Energy Journal 27(Special Issue 2):13–38

Bosetti V, Carraro C, Massetti E, Tavoni M (eds) (2014) Climate Change Mitigation, Technological Innovation And Adaptation - A New Perspective on Climate Policy. Edward Elgar Publishing

Bréchet T, Gerard F, Tulkens H (2011) Efficiency vs. stability in climate coalitions: a conceptual and computational appraisal. Energy Journal 32(1):49

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

Carraro C, Eyckmans J, Finus M (2006) Optimal transfers and participation decisions in international environmental agreements. Review of International Organizations 1(4):379–96

Chander P, Tulkens H (1995) A core-theoretic solution for the design of cooperative agreements on transfrontier pollution. International Tax and Public Finance 2:279–93

d'Aspremont C, Gabszewicz JJ (1986) New Developments in the analysis of market structures, Macmillan, New York, chap On the stability of collusion, pp 243–64

Dellink R, de Bruin K, Nagashima M, van Ierland EC, Weikard HP (2015) STACO technical document 3: Model description and documentation of STACO-3, WASS Working Paper 2015-11, Wageningen University

Eyckmans J (2012) Review of applications of game theory to global climate agreements. Review of Business and Economic Literature 57(2):122–142

Eyckmans J, Finus M (2006) Coalition formation in a global warming game: how the design of protocols affects the success of environmental treaty-making. Natural Resource Modeling 19(3):323–358

Eyckmans J, Finus M (2007) Measures to enhance the success of global climate treaties. International Environmental Agreements: Politics, Law and Economics 7(1):73–97

Eyckmans J, Tulkens H (2003) Simulating coalitionally stable burden sharing agreements for the climate change problem. Resource and Energy Economics 25:299–327

Eyckmans J, Finus M, Mallozzy L (2013) A new class of welfare maximizing stable sharing rules for partition function form games, working Paper

Fankhauser S (1995) Valuing climate change: the economics of the greenhouse. Routledge

Finus M (2008) Game theoretic research on the design of international environmental agreements: Insights, critical remarks, and future challenges. International Review of Environmental and Resource Economics 2:29–67

Finus M, Pintassilgo P (2013) The role of uncertainty and learning for the success of international climate agreements. Journal of Public Economics 103:29–43

Finus M, van Ierland E, Dellink R (2006) Stability of climate coalitions in a cartel formation game. Economics of Governance 7:271–91

Fuentes-Albero C, Rubio SJ (2010) Can international environmental cooperation be bought? European Journal of Operational Research 202(1):255–264

Hoel M (1992) International environment conventions: The case of uniform reductions of emissions. Environmental and Resource Economics 2(2):141–159

Karp L, Simon L (2013) Participation games and international environmental agreements: A non-parametric model. Journal of Environmental Economics and Management 65(2):326–344

Kolstad C, Ulph A (2008) Learning and international environmental agreements. Climatic Change 89:125–141

Kolstad C, Ulph A (2011) Uncertainty, Learning and Heterogeneity in International Environmental Agreements. Environmental and Resource Economics 50:389–403

Kornek U, Steckel J, Edenhofer O, Lessmann K (2013) The climate rent curse: New challenges for burden sharing, Presented at 20th Annual Conference of the European Association of Environmental and Resource Economists, 26–29 June 2013, Toulouse, France, available at http://www.webmeets.com/EAERE/2013/prog/viewpaper.asp?pid=757

Kornek U, Lessmann K, Tulkens H (2014) Transferable and non transferable utility implementations of coalitional stability in integrated assessment models, CORE Discussion Paper nb 35

Lessmann K, Edenhofer O (2011) Research cooperation and international standards in a model of coalition stability. Resource and Energy Economics 33(1):36–54

Lessmann K, Marschinski R, Edenhofer O (2009) The effects of tariffs on coalition formation in a dynamic global warming game. Economic Modelling 26(3):641–649

Luderer G, Pietzcker RC, Bertram C, Kriegler E, Meinshausen M, Edenhofer O (2013) Economic mitigation challenges: how further delay closes the door for achieving climate targets. Environmental Research Letters 8(3):034,033

McGinty M (2007) International environmental agreements among asymmetric nations. Oxford Economic Papers 59(1):45–62

Metz B, Davidson O, Bosch P, Dave R, Meyer L (eds) (2007) Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Morris J, Paltsev S, Reilly J (2008) Marginal abatement costs and marginal welfare costs for greenhouse gas emissions reductions: results from the EPPA model, MIT Joint Program on the Science and Policy of Global Change. Report 164, Cambridge, MIT

Nagashima M, Dellink R, Van Ierland E, Weikard HP (2009) Stability of international climate coalitions—a comparison of transfer schemes. Ecological Economics 68(5):1476–1487

Nagashima M, Weikard HP, de Bruin K, Dellink R (2011) International climate agreements under induced technological change. Metroeconomica 62(4):612–634

Nordhaus W (2008) A question of balance: Economic modelling of global warming. Yale University Press, New Haven, CT

Nordhaus WD, Yang Z (1996) A regional dynamic general-equilibrium model of alternative climate-change strategies. The American Economic Review 86(4):741–765

Paltsev S (2010) Baseline projections for the EPPA-5 model, personal communication

Paltsev S, Reilly J, Jacoby HD, Eckaus RS, McFarland J, Sarofim M, Asadoorian M, Babiker M (2005) The MIT emissions prediction and policy analysis (EPPA) model: version 4, MIT Joint Program on the Science and Policy of Global Change. Report 125, Cambridge, MIT

Ramsey F (1928) A mathematical theory of saving. Economic Journal 38(152):543–559

Tol R (2009) The economic effects of climate change. Journal of Economic Perspectives 23(2):29–51

Weikard HP (2009) Cartel stability under an optimal sharing rule. The Manchester School 77(5):575–593

Weikard HP, Finus M, Altamirano-Cabrera J (2006) The impact of surplus sharing on the stability of international climate agreements. Oxford Economic Papers 58(2):209–232

Yang Z (2008) Strategic Bargaining and Cooperation in Greenhouse Gas Mitigations: An Integrated Assessment Modeling Approach. MIT Press