

How empirical uncertainties influence the stability of climate coalitions

Jasper N. Meya^{1,2,3}

Ulrike Kornek^{4,5}

Kai Lessmann^{5,4}

Abstract: International climate agreements are negotiated in the face of uncertainties concerning the costs and benefits of abatement and in the presence of incentives for free-riding. Numerical climate coalition models provide estimates of the challenges affecting cooperation, but often resort to assuming certainty with respect to the values of model parameters. We study the impact of uncertainty on the stability of coalitions in the Model of International Climate Agreements (MICA) using the technique of Monte-Carlo analysis. We extend the existing literature by (i) calibrating parametric uncertainty about damages and abatement costs to estimates from meta-studies and by (ii) explicitly considering uncertainty in the curvature of the damage function. We find that stability is more sensitive to uncertainty in damages than in abatement costs, and most sensitive to uncertainty about the regional distribution of damages. Our calculations suggest that heterogeneity can increase stability of coalitions; however this depends on the availability of transfers.

Keywords: International environmental agreements - Climate coalition formation - Uncertainty - Monte-Carlo analysis - Numerical modelling

JEL Classification: C72, D80, H87, Q54

Acknowledgments: We thank Achim Hagen, Andrew Halliday as well as conference participants at ICP 2015, EAERE 2016 and especially two anonymous reviewers for helpful comments on earlier drafts of this paper.

Conflict of Interest: The authors declare that they have no conflict of interest.

This article has been published online 20 October 2017 in *International Environmental Agreements: Politics, Law and Economics*, DOI 10.1007/s10784-017-9378-5.

1. Introduction

To address climate change, countries need to cooperate, as emission reductions represent a global public good. Global climate governance relies on international environmental agreements like the Kyoto Protocol to set up cooperation on climate change but participation and ambition is impeded by incentives for free-riding (Barrett 2003). Theoretical and numerical analysis have been applied to the design of agreements in order to overcome free-riding (e.g. Barrett 2005; Dellink 2011). In these studies, the incentive to sign self-enforcing climate agreements is calculated based on valuing abatement costs against avoided damages

¹ Jasper N. Meya, jasper.meya@hu-berlin.de

² Department of Economics, University of Oldenburg, Oldenburg, Germany

³ Resource Economics Group, Humboldt-Universität zu Berlin, Berlin, Germany

⁴ Mercator Research Institute on Global Commons and Climate Change, Berlin, Germany

⁵ Potsdam Institute for Climate Impact Research, Potsdam, Germany

when a country joins the agreement (Barrett 1994; Lessmann et al. 2015). However, abatement costs and especially climate damages are subject to large uncertainties (e.g. Tavoni and Tol 2010; Tol 2012) and conclusions are therefore limited. The present analysis studies the impact of uncertainty on the stability of agreements in a numerical climate coalition model - the Model of International Climate Agreements (MI-CA) - using the technique of Monte-Carlo analysis.

Monte-Carlo analysis allows us to draw conclusions about the quantitative influence of different uncertainties on the incentive to join an agreement. However, we abstract from decision making under uncertainty, which is difficult to implement in numerical models. While a recent academic debate has highlighted the importance of uncertainty for climate change decision making (Weitzman 2009; Nordhaus 2012), there are only few studies that look at the effect of uncertainty on the incentive to sign an agreement, and most of these are theoretical analyses (Kolstad 2007; Kolstad and Ulph 2008, 2011; Na and Shin 1998). Due to different assumptions about uncertainty, some studies find uncertainty to be conducive to cooperation while others find that it has a negative impact. Finus and Pintassilgo (2013) recently showed that the way in which uncertainty influences cooperation depends on the availability of transfers in the climate agreement. Depending on the realization of random variables uncertainty increases or decreases asymmetry in the distribution of damages across countries; more asymmetry hinders coalition formation without transfers, but can foster it with transfers as countries can be compensated (see also Weikard 2009).

A recent comparison among numerical models that seeks to extend these theoretical insights by incorporating quantitative results finds that model assumptions on the distribution of damages varies widely (Lessmann et al. 2015). However, representing such uncertainty explicitly in numerical models is very challenging, and only a few applications go beyond simple sensitivity analysis. These studies have been mainly applied to the Stability of Coalitions (STACO) model using Monte-Carlo analysis (Dellink et al. 2013; Dellink 2011; Dellink and Finus 2012; Dellink et al. 2008), with the exception of an application of stochastic programming in the CWS model (Bréchet et al. 2012). Uncertainty analyses in STACO show that the stability of coalitions is mainly determined by the distribution of damages rather than by the distribution of abatement costs (Dellink et al. 2008; Dellink 2011). Olieman and Hendrix (2006) introduce the concept of the “stability likelihood of an international climate agreement”. Given a probability distribution of uncertain model parameters, the indicator specifies the probability that a coalition (in which members have chosen to participate under certainty) can be labeled as stable. This metric, therefore, is driven by and reflects the uncertainty of model input, but decisions of players are taken under certainty for each possible parameter value. Dellink and Finus (2012) apply the concept and confirm the results from the theoretical literature, finding that heterogeneity across regions increases or decreases coalition stability depending on whether or not transfers are available.

The present study extends previous ones in several ways in order to draw a more comprehensive picture of the effect of uncertainty on coalition stability. Firstly, we improve the numerical representation by calibrating the uncertainty about abatement costs and climate damages to evidence from meta-studies. A robust comparison of the magnitude of effects is thereby made possible. Secondly, we consider uncertainty not only about the level of damages but also regarding the curvature of the damages function, which reflects their sensitivity to temperature increase. This allows us to assess the implications of uncertainty in the curvature of the damage function, whose importance for coalition formation has recently been highlighted (Barrett 2013). Additionally, we go beyond previous studies by incorporating the (arguably more realistic case of) concave utility functions, i.e. decreasing marginal utility of consumption.

We use the numerical Model of International Climate Agreements (MICA) to study the relative effect of different uncertain parameters on the stability of coalitions without and with transfers. We limit our analysis to studying the incentive to stay in the coalition, without and with transfers, as the incentive to remain a non-member to the coalition has been described not to be the bottleneck of stability (Dellink 2011). Uncertainty in three key parameters is specified and propagated through the model using Monte-Carlo analysis, similar to the approach previously taken by Dellink et al. (2008). Our analysis therefore abstracts from decision making under uncertainty but allows derivation of probabilities of different model outputs and comparison of the relative importance of the specified uncertainties. Specifically, we consider (i) perfectly correlated global uncertainty vs. independent uncertainties across regions; (ii) uncertainty in abatement costs vs. uncertainty in climate change damages; and (iii) uncertainty about the curvature of climate change damages. This enables us to assign stability likelihoods to specific climate coalitions and assess the extent to which transfers among signatories may contribute to a stable agreement.

As our most important result, we find that the stability likelihood of coalitions is mainly driven by uncertainty about damages through various determinants. Firstly, stability is more affected by uncertainty about damages compared to uncertain abatement costs, and mostly affected by uncertainty about the interregional distribution of damages. Secondly, transfers become an important instrument when there is uncertainty about the distribution of damages. Our scenarios show that stability is sensitive to uncertainty when no transfers between regions are allowed but that the availability of transfers increases the robustness of stability under uncertainty. Each realization of climate change damages produces distinct winners and losers within the coalition. With transfers as an additional measure, worse-off regions can be compensated to remain in the coalition by winners that experience larger gains, i.e. adverse effects of the realizations of uncertainty are shared within coalitions. Furthermore, while large coalitions with ambitious abatement targets are not stable under certainty, we find that they have a positive probability to be stable under an uncertain distribution of damages when transfers are available. Lastly, we present scenarios in which uncertainty about the curvature of damages has a positive effect on stability. A higher sensitivity of damages towards a temperature increase tends to increase stability. Steeper damages resemble a critical threshold, which is known to improve the incentives to cooperate.

The remainder of this article is structured as follows: The next section defines the concepts of coalition stability employed in the analysis. Section 3 briefly describes the MICA model of coalition formation and specifies the probability density functions for the uncertain parameters. Section 4 presents the results of the Monte-Carlo analysis. Finally, Section 5 discusses and concludes.

2. Formation of climate coalitions

2.1 Coalition stability

We follow the standard approach in the literature and describe climate coalition formation as a cartel formation game (Hoel 1992; Carraro and Siniscalco 1993; Barrett 1994). Consider a set of N heterogeneous regions, for which the incentive to sign is modelled as a two-stage game. At the first stage regions decide whether they want to become a signatory or remain a singleton and at the second stage regions choose their emission abatements. The game is solved by backward induction, assuming that signatories maximise their payoffs jointly and non-signatories maximise their payoff individually, given the abatement decisions of the coalition and all other non-signatories. Abatement decisions are taken simultaneously.

The stability and performance of coalitions then depend on a vector of model parameters $\hat{\mathbf{d}} \in R^l$. To study uncertainty, this vector is replaced by a vector \mathbf{d} that is stochastic in one dimension and has the probability density function $pdf(\mathbf{d})$. To describe the impact of uncertainty on coalition formation, we apply the concept of stability likelihood developed by Olieman and Hendrix (2006) and Dellink et al. (2008). The stability likelihood of a coalition is the “the probability that the coalition stability claim is correct” (Olieman and Hendrix 2006). Before applying the concept to our analysis, we need to introduce some more terminology.

In MICA all decisions are taken under certainty. We hence draw a realization d of \mathbf{d} as a Monte-Carlo run to calculate the model output of MICA, yielding a distribution for each stability and performance indicator. We analyse the formation of coalitions by applying the concept of cartel stability introduced by d’Aspremont and Gabszewicz (1986) and applied to international environmental agreements by Barrett (1994). Formally, a coalition is stable if no member has an incentive to leave the coalition (internal stability) and no non-member has an incentive to join (external stability). For our Monte-Carlo analysis, we only look at internal stability, without and with transfers, in order to reduce the computational burden and, more importantly, as the incentive to remain a non-member to the coalition has been described not be the bottleneck of stability (Dellink 2011). We limit our analysis to studying the incentive to stay in the coalition, without and with transfers.

The incentive to stay in a coalition S , a subset of all regions, for a signatory $i \in S$ is defined as the difference between the payoff π_i upon joining to form S and the payoff from free-riding by leaving S , for a given value of the parameter d :

$$\Delta_i(S, d) = \pi_i(S, d) - \pi_i(S \setminus \{i\}, d). \quad (1)$$

A coalition S is called *internally stable (IS)* if for every signatory $i \in S$ the coalition payoff is higher than the free-rider payoff:

$$is(S, d) := \begin{cases} 1, & \text{for } \pi_i(S, d) \geq \pi_i(S \setminus \{i\}, d) \forall i \in S \\ 0, & \text{else} \end{cases}. \quad (2)$$

Furthermore, a coalition S is *potential internally stable (PIS)* if a transfer scheme τ exists that guarantees every coalition member at least its free-rider payoff, i.e. compensation equal to the value of benefits from leaving the coalition (Carraro et al. 2006). If consumption rather than utility is transferable, as is the case in MICA (see below), the concept has to be adapted (Kornek et al. 2014). Accordingly, a coalition in MICA is PIS if a transfer scheme τ exists that redistributes consumption within the coalition, providing a payoff for every signatory that is at least equal to its free-rider payoff at all times $t \in [t_0, t_m]$ (Kornek et al. 2014):

$$pis(S, d) = \begin{cases} 1, & \text{if } \exists \tau: \pi_i(c_{i,t}(S, d) + \tau_{i,t}(S, d)) \geq \pi_i(S \setminus \{i\}, d) \forall i \in S, \sum_{i \in S} \tau_{i,t}(S, d) = 0 \forall t \in [t_0, t_m] \\ 0, & \text{else} \end{cases} \quad (3)$$

Here $c_{i,t}$ is total consumption of each player and $\tau_{i,t}(S, d)$ is the transfer paid or received.

Finally, we can apply the concept of stability likelihood developed by Olieman and Hendrix (2006) and Dellink et al. (2008) to IS and PIS. The *internal stability likelihood* and the *potential internal stability likelihood* of a coalition S are then defined as the means of IS and PIS across the Monte-Carlo runs and can be interpreted as the probability that coalition S is IS and PIS respectively for the assigned *pdfs* of the uncertain parameters (Olieman and Hendrix 2006).

2.2. Coalition performance

Next, we are interested in the ability of a coalition to enhance global welfare and reduce environmental damages. The performance of a coalition S is defined relative to the performance of the non-cooperative outcome (all countries are singletons) and the fully cooperative outcome (grand coalition) as a closing-the-gap index, originally developed by Eyckmans and Finus (2006), and applied to study coalition performance under uncertainty by Dellink et al. (2013) and Dellink and Finus (2012). In the grand coalition (GC) the climate change externality is fully internalized and hence the outcome is most efficient, while in the case of all singletons (AS) regions only internalize their own damages and the outcome is non-cooperative.

Performance regarding global welfare WP is defined for a coalition S as:

$$WP(S, d) = \frac{\sum_{i=1}^N \pi_i(S, d) - \sum_{i=1}^N \pi_i(AS, d)}{\sum_{i=1}^N \pi_i(GC, d) - \sum_{i=1}^N \pi_i(AS, d)} \quad (4)$$

Performance regarding environmental quality EP is defined as the ability to reduce global temperature increase (ΔT) by 2195:

$$EP(S, d) = \frac{\Delta T(S, d) - \Delta T(AS, d)}{\Delta T(GC, d) - \Delta T(AS, d)} \quad (5)$$

The effect of uncertainty on the level and uncertainty of WP and EP are estimated using standard statistical measures on the sample of MC runs.

3. The numerical model and uncertainty distributions

This section first introduces the numerical coalition formation model MICA and describes how the uncertain parameters enter the model (Section 3.1). Subsequently, probability distribution functions for uncertain parameters are specified and the associated scenarios are presented (Section 3.2).

3.1. Model of International Climate Agreements

We conduct our study in the Model of International Climate Agreements (MICA), developed by Lessmann et al. (2009) and extended by Kornek et al. (2017). MICA is a numerically calibrated model of coalition formation encompassing eleven heterogeneous world regions (see Table 6 in the Appendix for a description of their makeup): Africa (AFR), China (CHN), the European Union with 27 member states (EUR), India (IND), Japan (JPN), Latin America (LAM), Arabic countries (MEA), other Asian countries

(OAS), Russia (RUS), USA (USA) and the “rest of the world” (ROW).⁶ These eleven world regions yield 2037 different possible coalitions.

In our analysis we focus on twelve coalitions, which correspond to cooperation among key players in the international climate negotiations (see Table 1).⁷ The OECD coalition comprises four regions, which have been main actors in past climate negotiations and are often held to be both the most responsible and capable. Our decision whether to consider a particular aggregate world region as part of the coalition of OECD countries was based on the relative influence of OECD member countries within this region as represented by their GDP. We also consider all possible coalition structures (ten in total) in which two or three of the OECD regions cooperate, in order to elucidate the effect of coalition size on internal stability. Additionally, we study a scenario in which the most relevant current and future emitters – China, India, Europe, USA as well as Africa and Russia – cooperate. Within our model the countries of this coalition of ‘major emitters’ are responsible for 67.59% of the global emissions in 2015 and for 62.15% of predicted emissions in 2105 (in the absence of a climate agreement). The large size of this coalition provides strong incentives for free-riding. Finally, we also calculate the fully cooperative outcome of the grand coalition (GC), where all regions cooperate in climate change mitigation, as a reference point for coalition effectiveness.

Table 1: Composition of the studied coalitions

Coalition	Signatories
EJ	EUR, JPN
ER	EUR, ROW
EU	EUR, USA
JR	JPN, ROW
JU	JPN, USA
RU	ROW, USA
EJR	EUR, JPN, ROW
EJU	EUR, JPN, USA
ERU	EUR, ROW, USA
JRU	JPN, ROW, USA
OECD	EUR, JPN, ROW, USA
Major Emitters	AFR, CHN, EUR, IND, RUS, USA
GC	AFR, CHN, EUR, IND, JPN, LAM, MEA, OAS, RUS, USA, ROW

MICA combines an integrated assessment model (IAM) with a coalition formation game.⁸ Thereby MICA uses a Ramsey-type optimal growth model with a representative agent for each region. The payoff π_i of region i is regional welfare defined as the aggregation of (discounted) instantaneous utility of consumption

⁶ For detailed definition of the regions in MICA and a full model description see Kornek et al. (2017).

⁷ The computational burden of Monte-Carlo analysis for coalition stability is substantial. Altogether, this manuscript is based on four Monte-Carlo ensembles: for parameters in mitigation costs, marginal damages (perfectly correlated and independent), and the slope of marginal damages. We execute the model 500 times per parameter and for 21 coalition equilibria. Thus we arrive at a total of $4 * 500 * 21 = 42,000$ Monte-Carlo runs and, at approximately 1 minute CPU time per shot, 42,000 minutes or 700 hours, or about 30 days of CPU time. To explore, for example, the stability of the grand coalition would add its ten subcoalitions to the list, raising the computation time by 50 percent.

⁸ For a discussion of the basic model structure see Lessmann et al. (2009).

of the representative agent of the region, with utility rising in per capita consumption and marginal utility decreasing in per capita consumption:

$$\pi_i = \sum_{t=t_0}^{t_m} \frac{1}{(1 + \rho)^{t-t_0}} p_{it} \log\left(\frac{c_{it}}{p_{it}}\right), \quad (6)$$

where c_{it} is the consumption and p_{it} is the population in region i at time t and ρ is the pure rate of time preference. The time horizon t goes from $t_0 = 2005$ to $t_m = 2195$ in steps of ten years.

Economic output can either be consumed or invested and is reduced by damages caused by climate change and expenditures for abatement. To assess the stability and performance of self-enforcing coalitions, two parts of the model are of particular importance: climate damages and abatement costs.

For each region the damages Ω_{it} of temperature increase is represented by

$$\Omega_{it} = \theta_{1,i} \Delta T_t^{\theta_2}, \quad (7)$$

where ΔT_t is the temperature increase at time t compared to the pre-industrial level and $\theta_{1,i}$ and θ_2 are parameters. This functional form and the quadratic exponent $\theta_2 = 2$ go back to the RICE model⁹ of Nordhaus and Yang (1996). The coefficients of the damage functions $\theta_{1,i}$ are taken from Dellink et al. (2004) and were spatially rescaled to the eleven world regions.

Abatement costs A_{it} are a function of the rate of emission control μ_{it} :

$$A_{it} = b_{1,it} \mu_{it}^{b_{2,i}}. \quad (8)$$

Again, the relationship employed originates from the RICE model (Nordhaus and Yang 1996). The regional parameters of the abatement cost function $b_{1,it}$ and $b_{2,i}$ are calibrated to the Regionalized Model of Investments and Development (Leimbach et al. 2010), a large-scale, state-of-the-art IAM.¹⁰

In principle all model parameters are uncertain. But incorporating the full range of uncertainty would place a tremendous computational burden on the Monte-Carlo analysis. To focus on key variables, we restrict our analysis to uncertainty in the abatement cost and damage function parameters $b_{1,it}$, $\theta_{1,i}$, θ_2 . Abatement costs and climate damages are identified as the determinants of coalition stability in the majority of the literature (for a review see Finus 2008), and Dellink et al (2008) also limit their analysis to costs and damages. In MICA, these three parameters selected for analysis are main determinants of coalition formation: The balance between costs of abatement and benefits, in terms of reduced damages, of internalizing emission externalities among members of the coalition determines changes in consumption and hence welfare. These changes are, in turn, assumed to determine decisions by members on whether or not to remain in the coalition.

⁹ Regional Integrated Model of Climate and the Economy

¹⁰ For a description of the calibration procedure see Kornek et al. (2017).

3.2. Uncertainty distributions

3.2.1 Level of abatement costs

We are interested in how uncertainty in the distribution of abatement costs shapes coalition formation. To this end we incorporate independent uncertainties in the coefficient of the abatement cost function $b_{1,i}$, while the functional form is taken as certain (as in e.g. Dellink et al. 2008; Nordhaus 2008).

To specify the probability distribution for the abatement cost coefficient $b_{1,it}$, we make the common assumption of a normal distribution (Dellink et al. 2008) but assign zero probability to negative values, as negative abatement costs throughout the whole study period 2005–2195 seem unrealistic.

As it is not evident *ex ante* which probability distribution is most suitable to describe uncertainty in abatement costs over a long time horizon, assumption of a normal distribution appears to us to be the most conservative approach. We set parameter means equal to the deterministic version of MICA, since values are calibrated to a multi-regional IAM with technologically disaggregated abatement costs. It is assumed that the uncertainty is identical across regions, represented by the same coefficient of variation. Furthermore, it seems unrealistic that the uncertainty in the abatement costs for every time step is independent from the abatement costs in the previous period. Thus, rather than drawing a $b_{1,it}$ for each time step from probability density functions (*pdf*), we assume that there is uncertainty about the path of abatement costs as a whole.

We therefore introduce a random variable $k \sim N(1,0.61)$. The variable has a multiplicative effect on costs, hence we choose a unity mean, i.e. no effect on average. We choose a standard deviation that reflects the variation reported in the meta-study by Tavoni and Tol (2010) of the results from the EMF 22 International Scenarios (Clarke et al. 2009) as follows: In the model comparison ten IAMs were running scenarios with different emissions stabilization targets for 2100, with and without the possibility of temporarily overshooting these targets and delayed or full participation in an international climate agreement (Clarke et al. 2009). We take the average coefficient of variation (CV) 0.24 of the global abatement costs of the two scenarios from Tavoni and Tol (2010) most closely related to MICA¹¹. Based on this we calibrate the standard deviation of the global abatement costs, measured as a fraction of output from 2005 to 2105 and discounted at a rate of 5%. Finally, the corresponding distribution of k is calculated, yielding a standard deviation of 0.61686. In each Monte-Carlo run we draw a realization of k independently for every region and calculate the corresponding values of $b_{1,it}$, so that regional heterogeneity varies between runs.

3.2.2 Level of damages

Damages caused by climate change are particularly uncertain regarding both their global level and their regional distribution. Not only might climate change damages become more or less severe in all regions, but also the regional distribution of damages might differ from today's best guess estimates. As both uncertainties are likely to affect coalition formation differently, we study uncertainty in global climate change damages and their regional distribution separately.

¹¹ These scenarios are 650 CO₂e/ full participation/ no overshoot and 550 CO₂e/ full participation/ overshoot with a CV of 0.27 and 0.21, respectively.

To this end we simulate uncertainty in the level of damages in two ways. Firstly, we model uncertainty in *global* damages as perfectly correlated uncertainty in the damage function coefficient $\theta_{1,i}$. Therefore, in each Monte-Carlo run a quantile is randomly drawn and the corresponding values for $\theta_{1,i}$ are taken from their regional distributions. This approach leaves the heterogeneity of damages across regions unchanged. Secondly, we model uncertainty in *regional* damages as independent uncertainty by drawing the values of the damage function coefficient $\theta_{1,i}$ independently from their regional distributions. Consequently, uncertainty in regional damages changes the heterogeneity across regions. The two scenarios are directly comparable, as regional distributions are calibrated to produce the same mean global damages across Monte-Carlo runs.

We take the distribution of estimates from 14 major IAMs as reviewed by Tol (2012) for a temperature increase of 2.5 °C as proxy for the probability of different climate change damages. This implicitly assumes that each of the 14 estimates is certain. We then model the damage coefficients $\theta_{1,i}$ to be distributed according to a *pdf* that reproduces the mean and standard deviation of the global damages from Tol (2012). The mean values for $\theta_{1,i}$ are taken from the deterministic version of MICA and are rescaled by a common factor to meet the mean global damages for a temperature increase of 2.5 C° in Tol (2012). For uncertainty in both global and regional damages, all regions are exposed to the same degree of uncertainty, represented by a uniform coefficient of variation across regions. Again, the standard deviations of $\theta_{1,i}$ are calibrated to produce the standard deviation of global damage estimates as reviewed in Tol (2012). This procedure gives higher standard deviations for uncertainty in regional damages compared to uncertainty in global damages (see Table 2). This is because regional uncertainties partly cancel each other out in the calculation of global damages.

Table 2: Distribution of $\theta_{1,i}$ for uncertainty in global and regional damages

Region	Uncertainty in global damages			Uncertainty in regional damages	
	Mean μ	Standard deviation σ	Coefficient of variation	Standard deviation σ	Coefficient of variation
ROW	0.00073	0.00072	0.97872	0.00177	2.41984
AFR	0.00111	0.00109	0.97872	0.00269	2.41984
LAM	0.00094	0.00092	0.97872	0.00229	2.41984
IND	0.00322	0.00315	0.97872	0.00779	2.41984
CHN	0.00266	0.00260	0.97872	0.00643	2.41984
MEA	0.00117	0.00114	0.97872	0.00283	2.41984
OAS	0.00118	0.00115	0.97872	0.00285	2.41984
JPN	0.00136	0.00133	0.97872	0.00329	2.41984
RUS	0.00228	0.00223	0.97872	0.00551	2.41984
USA	0.00088	0.00086	0.97872	0.00213	2.41984
EUR	0.00080	0.00079	0.97872	0.00194	2.41984

We employ different *pdfs* to model uncertainty in global and regional damages. The choice was between normal or lognormal distributions, both of which have been applied in previous studies of climate change damages (Tol 1995, Crost and Traeger 2011). For uncertainty in global damages, we assume a normal

distribution and assign zero probability to negative damages, as it is very unlikely that a temperature increase of 2.5 degrees would be beneficial across all eleven world regions. A similar approach was previously taken by Hwang et al. (2013). By contrast, we assume a lognormal distribution for uncertainty in regional damages. In principle, the idea of negative damage costs might seem less problematic when modelling uncertainty in regional damages. However, negative coefficients in the damage function would imply regional benefits from climate change even at very high temperatures. Tol’s (2014) survey of economic impacts of warming does not show a single study that projects benefits for the range of warming in our scenarios, and we therefore find it prudent to preclude benefits in our study. Moreover, in this case, the assumption of a truncated normal distribution would result in very distorted global damages, as a lot of probability mass would be cut off due to the large standard deviations—which are substantially larger for regional uncertainty than for uncertainty in global damages. We therefore opt for a lognormal distribution of the damage function coefficient $\theta_{1,i}$ for uncertainty in regional damages, as this allows us to reproduce the mean and standard deviation of the damages reviewed by Tol (2012) exactly and satisfies our assumption of zero likelihood of negative damages. This assumption follows earlier studies on coalition formation that also implicitly assume zero likelihood of negative damages (Dellink et al. 2008). Both *pdfs* for uncertainty in regional and global damages are in line with the result across IAMs that uncertainty about damages is right-skewed (Tol 2012).

3.3.3 Curvature of the damage function

Lastly, we look at how uncertainty about the curvature of the damage function affects coalition formation. This is particularly relevant, as this ‘damage function’ is both subject to large uncertainties (Weitzman 2012) and strongly influences model results (IPCC 2001), in particular for large temperature increases (Weitzman 2010). Most commonly, regionally disaggregated IAMs assume a polynomial of second degree (Bosetti et al. 2006; 2013; Bréchet et al. 2011; 2012; Nordhaus 2010). However, the functional form is controversial. In particular, Weitzman (2009) has highlighted the importance of low-probability high-impact events, which might be underestimated by quadratic damage functions.

We therefore employ a lognormal distribution for the exponent of the damage function θ_2 , as specified by Dietz (2013), based on the following assumptions: (1) zero warming has zero economic cost, (2) the best guess is $\theta_2 = 2$ and (3) the 90th percentile is $\theta_2 = 3$. For computational reasons we truncate the lognormal distribution at $\theta_2 = 4$, in effect incorporating the assumption that (4) values above 4 have a probability of zero¹². Note that this procedure neglects recently highlighted issues for cost-benefit analyses associated with ‘fat-tail’ damage probability distributions (Weitzman 2007, 2009), as Weitzman’s ‘Dismal Theorem’ only holds for unlimited parametric uncertainty (Nordhaus 2011).

While other similar studies incorporate uncertainty indirectly in models by increasing global damages (Dellink et al. 2013), we take a different approach and isolate the effect of the functional form from the level of global damages. Therefore, for every random draw of θ_2 we recalibrate the best guess of $\theta_{1,i}$ so that the global damages for a temperature increase of 2.5 °C are equal to the mean of the damages reviewed by Tol (2012).

It is common practise in regionally disaggregated IAMs to assume a constant curvature for the damage function across regions (e.g. Bosetti et al. 2013; Bréchet et al. 2011; Dellink et al. 2013; Nordhaus 2010);

¹² An exponent of 4 equals the 90% percentile of Nordhaus’s (1994) subjective cumulative probability function in the explorative stage of his sensitivity analysis.

correspondingly, we assume that there is common uncertainty about θ_2 to all regions. This is implemented by simply drawing a single θ_2 in every Monte-Carlo run for all regions, thereby, leaving the heterogeneity of regions unaffected.

4. Results

We studied the effect of the four specified uncertainties on coalition performance and stability separately. To this end, we conducted numerical experiments for uncertainty in abatement costs, global damages, regional damages and the curvature of the damage function (for an overview see Table 3). For each scenario (and coalition) we ran the model 500 times. We reserved one model run for the deterministic default parameter values, leaving four samples each with 499 runs for the Monte-Carlo analysis.

Our choice of the number of Monte-Carlo runs was guided by the trade-off between the computational costs of the model runs and the precision of the estimated mean and variance of the Monte-Carlo ensemble (for the computational burden, see Footnote 3). For precision, we need the number of Monte-Carlo runs to be large enough for robust estimates of sample mean and sample variance. Higher moments are more demanding, so we consider the standard error of the *uniform minimum variance unbiased* estimator for the sample standard deviation. For large sample sizes n (here, more than 50 can be considered large), we have $\frac{\sigma_s}{\sigma} = \sqrt{2(n-1)}^{-1}$. Applying this equation, for $n = 500$ the standard error of the estimated standard deviation is just over 3 percent (compared to the estimated standard deviation). For $n = 100$, the error is 7 percent, and doubling n (i.e. $n = 1000$) only reduces the error to 2.2 percent.

In this section the results are reported successively for coalition performance (Section 4.1), coalition stability without transfers (Section 4.2) and coalition stability with transfers (Section 4.3). Histograms for all specified distributions can be found in the Appendix.

Table 3: Overview on modelled uncertainties

Scenario	Parameter	Type of uncertainty	pdf
abatement costs	b_1	Independent	normal (truncated)
global damages	θ_1	perfectly correlated	normal (truncated)
regional damages	θ_1	independent	Lognormal
damage curvature	θ_2	perfectly correlated	lognormal (truncated)

4.1 Coalition performance

In general, economic output, e.g. discounted consumption, the endogenous discount rate and welfare, are little affected by the imposed uncertainties. For instance, global discounted consumption in the cases where no regions or all regions cooperate shows very small coefficients of variation and the uncertainty for each region is negligible compared to interregional differences.

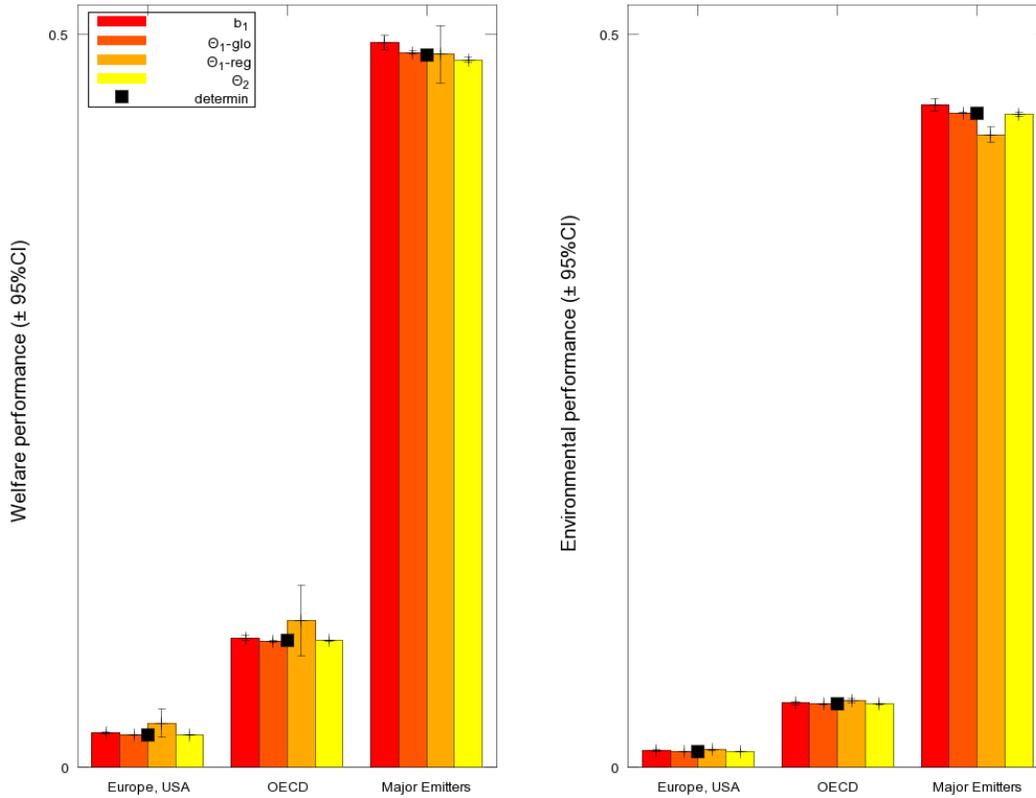


Figure 1: Mean welfare performance (WP) and environmental performance (EP) and corresponding 95% confidence intervals (whiskers) for uncertainty in abatement costs (\mathbf{b}_1), global damages (θ_1 -glo), regional damages (θ_1 -reg) and curvature of the damage function (θ_2). The figure shows results for the coalitions Europe and USA (EU), OECD and ‘major emitters’. Black boxes indicate welfare performance and environmental performance for the deterministic case.

This is in line with the findings of many IAMs that aggregated climate change damages in monetary terms makes up only a small fraction of global economic output (Tol 2013). The environmental variables temperature, cumulative emissions and carbon price are more sensitive to the uncertainty in abatement and damages costs, as one would expect given the more direct link to the uncertain variables.

The performance of all studied coalitions in both welfare and environmental respects follow a similar pattern. In both cases, the major emitter coalition achieves around half of the social optimum, due to its size, while all smaller coalitions achieve little. Generally, performance tends to increase with the size of coalitions.

Figure 1 displays the performance of three selected coalitions to facilitate global welfare and limit the temperature increase by 2195. The applied uncertainties do not lead to substantially different mean values of the coalition performance indicators, which stay relatively constant across all four scenarios. The induced uncertainty, as displayed by the confidence intervals, is also low and shows only minor differences among the four modelled uncertainties. The only exception is the uncertainty in regional damages which substantially increases uncertainty about welfare.

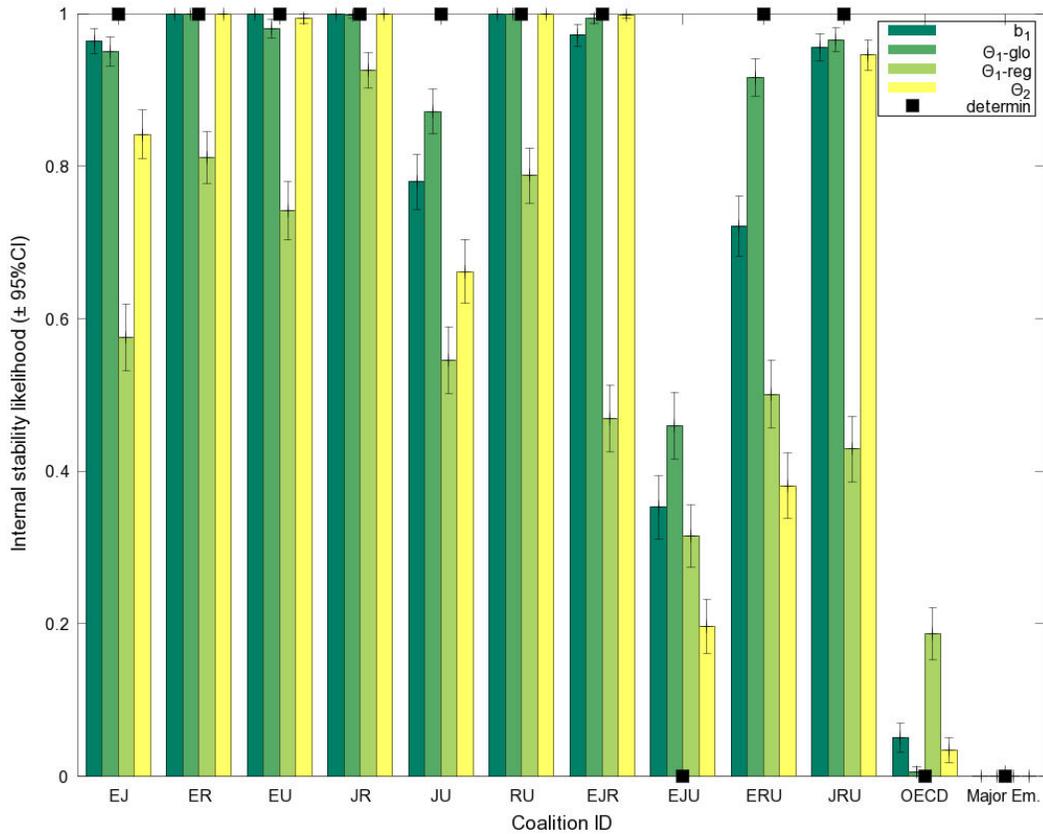


Figure 2: Internal stability likelihood and corresponding confidence intervals (whiskers) by coalitions for uncertainty in abatement costs (b_1), global damages ($\theta_1\text{-glo}$), regional damages ($\theta_1\text{-reg}$) and curvature of the damage function (θ_2). Black boxes indicate the internal stability likelihood for the deterministic case

Furthermore, mean welfare and environmental performance under uncertainty are hardly different from the estimates without uncertainty, i.e. in the deterministic case. Thus, coalition performance generally appears to be only little affected by uncertainty in damages and abatement costs.

4.2. Coalition stability without transfers

In this section we discuss the effect of uncertainty on coalition stability when no transfers are available, as measured by internal stability likelihood (ISL) indicator. Figure 2 displays the relative effect of the four simulated uncertainties on ISL. It shows that uncertainty affects the internal stability of all coalitions except the major emitter coalition, but that differences between scenarios are considerable. For the major emitter coalition incentives for free-riding are so high – due to its size – that we did not find a scenario for which internal stability holds.

The relative effect of the uncertainties depends on the coalition. However, all the coalitions that are internally stable in the deterministic setting become unstable in many runs in the scenario with uncertainty in the regional distribution of damages. Compared to the binary result for stability in the deterministic set-

ting, the internal stability likelihood reveals much more mixed results. Coalitions that were internally stable in the deterministic setting are less often internally stable in the scenario with uncertainty in the regional distribution of damages; conversely, some coalitions that were not internally stable are sometimes internally stable when uncertainty is concerned. In other words, parameter variation within the ranges of uncertainty changes the incentive to participate from positive to negative and vice versa. The only exception is the major emitter coalition, for which the incentives to free-ride are so large that it is not affected by the imposed uncertainty at all.

It is particularly interesting to compare the effect of uncertainty in regional and global damages, as only the former affects the heterogeneity of damages among signatories. For all coalitions that were internally stable in the deterministic setting, the internal stability likelihood is lower in the scenario with uncertainty in the regional distribution of damages, compared with uncertainty in global damages. However, for coalitions that were not internally stable in the deterministic setting results are less clear cut: while the internal stability likelihood of the OECD coalition is higher for uncertainty in regional damages than for uncertainty in global damages, the reverse is true for the EJU coalition. Overall, this suggests that without transfers the internal stability of coalitions is lost by changes in the heterogeneity of damages. When heterogeneity, i.e. the regional pattern of damages, changes, the incentives of single regions frequently flip from positive to negative. Transfers could (at least in part) even out these changes. But in their absence these changes affect the internal stability of the coalition more frequently than in the scenario with uncertainty in global damages.

Furthermore, it is instructive to compare the effects of uncertainties in abatement and regional damages, as both scenarios lead to changes in heterogeneity among signatories. The results of our numerical experiments suggest that uncertainty in damages matters more to signatories than uncertainties in abatement costs, indicated by the more pronounced changes in internal stability likelihood values compared to the deterministic setting. These results reflect the greater degree of uncertainty associated with the estimation of damages compared to abatement costs. Moreover, uncertainty about the curvature of the damages function, i.e. the sensitivity of damages to the magnitude of temperature increase, affects internal stability more than uncertainty about the level of global damages.

For insights into the underlying reasons for these observations, we investigate what drives internal stability for the example of the OECD coalition, consisting of the European Union, Japan, ‘rest of the world’ and the USA. The conditions under which the OECD becomes internally stable are elicited by comparing parameter values in runs that are internally stable with their mean values across all Monte-Carlo runs (Table 4). This allows us to study parameter values for which the OECD becomes internally stable. For mean parameter values greater than one (smaller than one), uncertainty increases stability in cases where there is an increase (decrease) in the respective parameter relative to the whole sample mean. For instance for b_{1i} a value larger than unity indicates that for region i the abatement costs have to increase on average relative to whole sample mean in order for the OECD to become internally stable.

Table 4: Summary statistics (mean [min; max]) for uncertain parameters in runs in which the OECD coalition is internally stable relative to means across all Monte-Carlo runs.

	‘rest of the world’	Japan	USA	European Union
b_1	1.293 [0.307; 2.243]	1.092 [0.212; 2.206]	0.249 [0.005; 1.931]	0.821 [0.009; 1.966]
θ_1-glo	0.013 [0.008; 0.022]			
θ_1-reg	1.013 [0.009; 11.74]	0.84 [0.062; 6.77]	0.296 [0.011; 1.324]	0.237 [0.003; 1.463]
θ_2	1.868 [1.71; 2.081]			

Within the OECD coalition, the ‘rest of the world’ region has a high abatement burden, as a result of its relatively cheap mitigation options¹³, which however mainly benefits the other signatories (when no transfers are available). As the ‘rest of the world’ suffers relatively little from climate change, cooperation is unattractive for the region when the deterministic case is considered. The simulated uncertainties display that in order to give the ‘rest of the world’ an incentive to remain in the OECD, its abatement burden needs to drop, while the European Union and especially the USA have to increase abatement efforts. This is reflected in higher European Union and US abatement costs and lower ‘rest of the world’ abatement costs in the IS runs compared to the whole sample means. Or alternatively, the relative economic impact of climate change to the ‘rest of the world’ has to increase to give it sufficiently high benefits from cooperation. Note that for the scenarios of uncertainty about abatement costs b_1 and especially for uncertainty about regional damages θ_1-reg the OECD can become stable for a wide range of parameter values of each signatory, suggesting that stability is most affected by uncertainty about the distributional effects of climate coalitions, especially relating to damages.

Comparison of the effects of uncertainty in global damages and uncertainty in the curvature of the damage function reveals the importance of functional uncertainty for coalition stability. Regarding the former, the OECD becomes only internally stable if there are hardly any damages, which is the trivial case of no climate change externality. In contrast the likelihood that the OECD is internally stable increases with uncertainty in the curvature of the damage function. Specifically, the OECD becomes internally stable when the sensitivity of damages to the magnitude of temperature increase is high – more than 1.7 times the mean. This suggests that cooperation is induced by common interest of avoiding a temperature threshold above which severe damage will occur.

4.3. Coalition stability with transfers

To study how coalition stability is affected by uncertainty when transfers between signatories are available, we now analyse the relative effect of different uncertainties on the potential internal stability likelihood. Almost all studied coalitions are potentially internally stable (PIS) in the deterministic setting (Figure 3), as the coalition payoff is high enough to compensate signatories with a negative incentive. However, the large coalition of all major emitters is not internally stable, even with transfers, as free-rider incentives are so high that not all signatories can be kept within the coalition.

¹³ Abatement costs are measured by an abatement cost index, which is defined as the reciprocal of the cumulative abatement over the 21st century by each region in the coalition of all regions compared to the all singletons scenario. The rationale for this approach is that within the grand coalition abatement is done where it is less costly and if marginal abatement costs are well-behaved, abatement costs are inversely related to the abatement done in the grand coalition. This approach was developed by Lessmann et al. (2015).

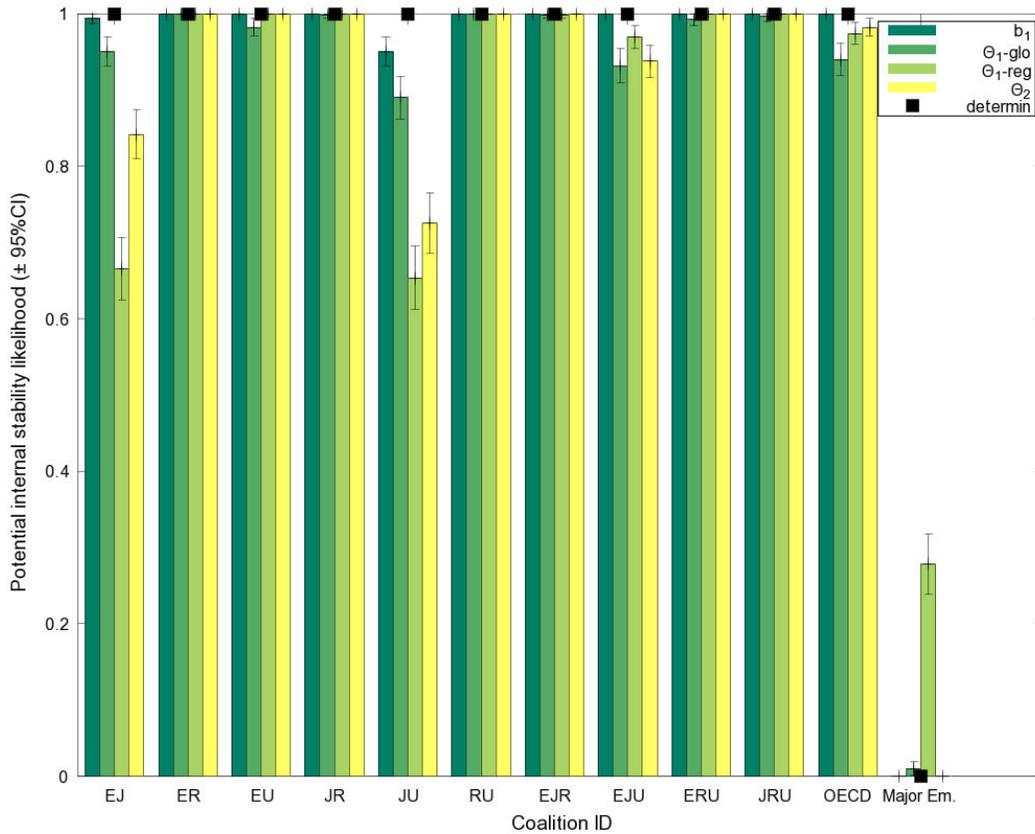


Figure 3: Potential internal stability likelihood and corresponding confidence intervals (whiskers) by coalitions for uncertainty in abatement costs (b_1), global damages (θ_{1-glo}), regional damages (θ_{1-reg}) and curvature of the damage function (θ_2). Black boxes indicate the potential internal stability likelihood for the deterministic case.

Transfers increase the stability of coalitions under uncertainty. When transfers are possible the internal stability of coalitions is less sensitive to the applied uncertainties compared to a situation with no transfers, measured as the deviation from the deterministic case. However, some coalitions, especially the coalition between the European Union and Japan (EJ) and Japan and USA (JU), are unstable in the potential internal stability sense in many runs. Remarkably, the major emitter coalition, the only coalition which is not potentially internally stable in the deterministic case, has a considerable likelihood to be potentially internally stable under uncertainty.

The differences between the internal stability likelihood (Figure 2) and potential internal stability likelihood (Figure 3) indicate that the effect of heterogeneity on stability depends on the availability of transfers. In particular, the higher potential internal stability likelihood of the major emitter coalition under uncertainty in the regional distribution of damages is in line with the theoretical result that any coalition can become potentially internally stable if the heterogeneity between players is sufficiently high (Finus and Pintassilogo 2013; Weikard 2009). However, for coalitions that are already potentially internally sta-

ble in the deterministic case the PIS likelihood indicator does not reveal the increasing incentives that may result from uncertainty for signatories to stay in the coalition.

A different indicator is necessary to gain a better understanding of how uncertainty influences the incentive to stay within a coalition when transfers are available but the value of the PIS likelihood is already one. To this end, we studied the amount of consumption that each player can pass on while still having an incentive to stay inside the coalition, denoted as the surplus (Carraro et al. (2006) introduce a similar concept for transferable utility models).¹⁴ We find that the surplus, used as a continuous indicator of potential internal stability, yields more detailed and more positive conclusions than the PIS likelihood about the effect of uncertainty on the stability of (already stable) coalitions.

Figure 4 shows the change in the surplus under uncertainty relative to the absolute estimate in the deterministic case (ratio). When negative, the surplus illustrates how far off a coalition is from being potentially internally stable; when positive, the surplus is a measure of the extent to which the coalition is generating a positive incentive for its members to participate. For coalitions that are always potentially internally stable, a higher surplus indicates that under uncertainty the regions within the coalition have a stronger incentive to stay after redistribution, as the higher surplus is shared among the signatories.

The surpluses shown in Figure 4 indicate that even coalitions where the potential internal stability under uncertainty drops below one have on average greater incentives to stay in the coalition, compared to the deterministic case, since they can expect to gain substantial benefits from cooperation. A remarkable exception is the coalition of major emitters. The negative incentives for signatories to stay in this coalition, due to its size, are reinforced under uncertainty in abatement costs, global damages and the shape of the damage function. Only uncertainty about regional damages causes the incentive for signatories to stay to increase. This finding is in line with the theoretical result that when transfers are available any coalition can become stable if heterogeneity between signatory is sufficiently high (Finus and Pintassilogo 2013; Weikard 2009).

¹⁴ Kornek et al. (2014) define a measure for the surplus in the non-transferable utility framework and develop an algorithm to compute it, which we describe here. For a coalition S consider the transfer scheme τ that redistributes consumption so that the payoff of each signatory k is at least at its free-rider level. The surplus can then be defined as the maximal consumption per coalition member every signatory j could lose, still having a positive incentive to stay, discounted at rate r_t over time t_0 to t_m :

$$lo(S, d) = \max_{\tau_{k,t}, \Delta C_t} \left(\sum_{t=t_0}^{t_n} \frac{1}{1 + r_t} \Delta C_t(S, d) \right)$$

subject to $\pi_k(c_j(S, d) + \tau_{k,t}(S, d) - \Delta C_t(S, d)) \geq \pi_k(S \setminus \{j\}, d), \forall k \in S.$

The surplus aggregates the consumption streams of all members of the coalition that are available for arbitrary redistribution when the coalition is internally stable after transfers. To gain an indicator of how much additional transfers each member could receive, we divide the sum of consumption by the number of members. If the surplus is positive, each member has an incentive to stay inside the coalition after redistribution and the magnitude of the surplus indicates how strong the incentives to stay are. If the surplus is negative, the coalition is not PIS and the negative magnitude of the surplus indicates how much more consumption would be necessary inside the coalition in order for the incentive for members to stay to become positive.

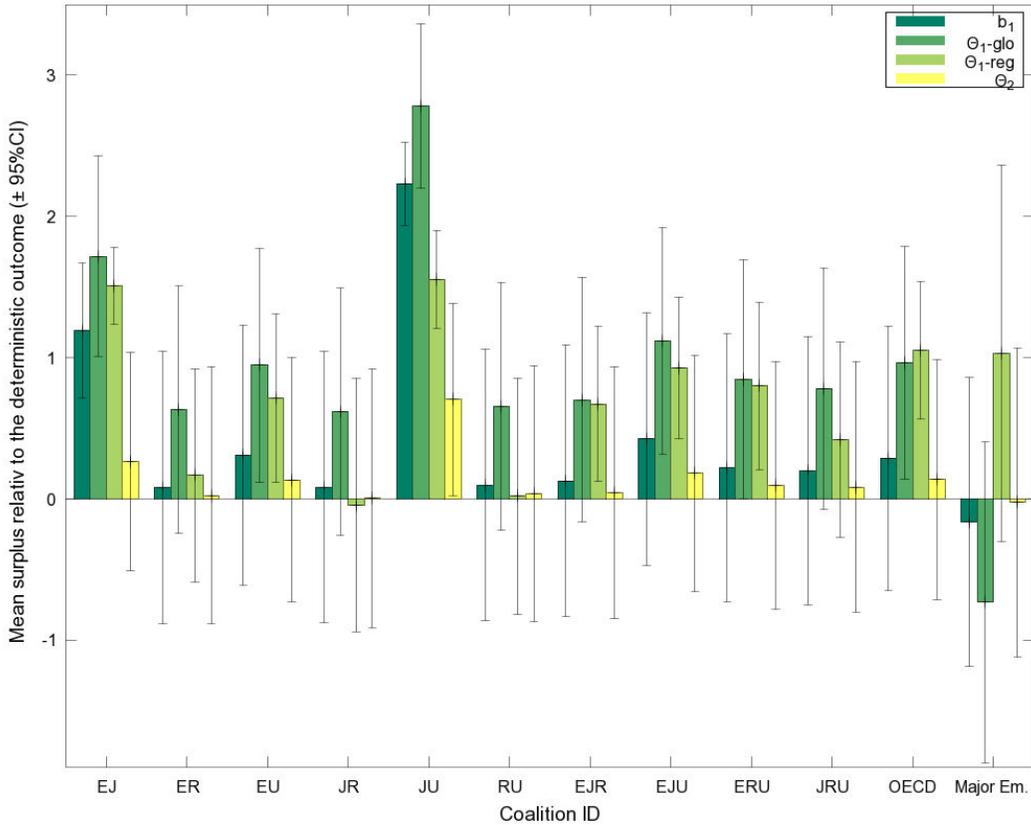


Figure 4: Mean surplus relative to the absolute estimate in the deterministic case (ratio) and corresponding confidence intervals (whiskers) for uncertainty in abatement costs (b_1), global damages (θ_{1-glo}), regional damages (θ_{1-reg}) and damage curvature (θ_2)

A more detailed view of how the incentive to stay within a coalition with transfers is affected by uncertainty can be gained by looking at the distribution of the surpluses (Table 7 in the Appendix). The uncertainty in the surplus – measured by the standard deviation – is in most cases highest for uncertainty in the regional distribution of damages. For the majority of the studied coalitions the mean of the surplus is higher than the median and hence the probability density of the surplus is right-skewed. The only exception is again the major emitter coalition, which has a left-skewed distribution of the incentive to stay in all uncertainty scenarios except in the case of uncertainty about regional damages.

Uncertainty in the regional distribution of damages produces the most skewed distributions of the incentive to stay within the coalitions, as the difference between the median and the mean surplus is always highest in this scenario. Furthermore, for uncertainty in regional damages the likelihood of having a strong incentive to stay goes up for all studied coalitions, including even the major emitter coalition. Thus, when transfer schemes are available, uncertainty about regional damages increases the internal stability of coalitions.

tions. However, these observations have an explorative character due to the limited sample of coalitions in this study.

A principal result of the above analysis is that uncertainty in damage and abatement costs mainly affects coalition stability through its effect on the heterogeneity of signatories and that the direction of the effect depends on the availability of transfers. To test the link between the heterogeneity of signatories and the stability of coalitions without and with transfers we study correlations between the coalition stability and the distribution of damages and abatement costs (Table 5), an approach employed by Eyckmans and Bréchet (2012). For each coalition, heterogeneity is measured as the coefficient of variation (CV) across all signatories in abatement, CV_A , and damage cost, CV_Ω , respectively. Table 5 compares the two measures of heterogeneity against internal stability of coalitions without and with transfers, using ‘surplus’ as a measure of stability when transfers are available for the reasons explained above.

Regarding abatement costs, the relationship between heterogeneity and coalition stability is not clear cut. By contrast, for damages we find that heterogeneity is negatively correlated with coalition stability when countries have no possibility to transfer consumption. However, the relationship is reversed when transfers schemes are available. In this case heterogeneity in damages is always significantly positively correlated with the internal stability of coalitions. This means that when transfers are available, uncertainty in the regional distribution of damages is likely to increase the stability of coalitions. Thus, our experiments add numerical evidence to the analytical results regarding the relationship between heterogeneity and stability of coalitions (Weikard 2009) and link these to the issue of uncertainty in coalition formation (Finus und Pintassilgo 2013).

Table 5: Correlations between the heterogeneity in abatement costs (CV_A) or damages (CV_Ω) and coalition stability without transfers (IS) and with transfers (surplus)

Coalition	$r(CV_A, IS)$	$r(CV_A, surplus)$	$r(CV_\Omega, IS)$	$r(CV_\Omega, surplus)$
EJ	0.689**	-0.438**	-0.063	0.332**
ER	-	0.557**	-0.396**	0.183**
EU	-	0.562**	-0.388**	0.254**
JR	-	0.037	-0.371**	0.192**
JU	0.475**	-0.456**	0.021	0.371**
RU	-	0.549**	-0.369**	0.197**
EJR	0.742**	-0.151**	-0.334**	0.216**
EJU	0.339**	-0.061	-0.051	0.248**
ERU	-0.320**	0.477**	-0.368**	0.305**
JRU	0.812**	-0.133*	-0.279**	0.212**
OECD	0.091	-0.011	-0.178**	0.276**
Major Emitters	-	0.014	-	0.188**

Correlation coefficients (Pearson's R) for n = 500; (*) (**) indicate significance at the $p < 0.01$ and $p < 0.001$ level, respectively; coalitions with missing values are either always or never internally stable.

5. Conclusion

We have studied how uncertainties about climate change damages and abatement costs affect the stability of coalitions in a numerical model of coalition formation (MICA). While we do not explicitly consider decision making under uncertainty, our numerical Monte-Carlo analysis allowed us to quantify the probability that the claim that certain coalitions are stable is correct. To this end, we have translated estimates of uncertainty about abatement costs and climate change damages from empirical studies into four probability distributions for parameters in our model, and propagated these distributions through the model using Monte-Carlo analysis.

We found, firstly, that stability is very sensitive to uncertainty. As parameter variations within the bounds of uncertainty frequently change the stability of coalitions, whether or not a coalition is stable often appears to be a knife-edge case. Finding a specific coalition to be stable (or not) should therefore not be mistaken for a prediction. This underlines the value of more comprehensive indicators of the incentives to cooperate, such as stability likelihood, which reveals the robustness of such incentives.

Second, uncertainty about damages matters more than uncertainty about abatement costs. This is, of course, partly due to the larger empirical uncertainty in the case of damages, and highlights once again the pressing need for more research on the economic impacts of climate change. Furthermore, uncertainty in the regional distribution of damages has a stronger impact on stability than uncertainty in global damages. This, too, highlights the need for more accurate information on climate change impacts. The strong effect of uncertainty about regional damages implies that heterogeneity in regional damages, i.e. how the burden of climate change impacts is shared, matters more for cooperation than the magnitude of the global burden. It is therefore important to understand impacts on the regional level.

Third, we find evidence that heterogeneity increases coalition stability when transfers are available and decreases coalition stability when they are not. This result has already been demonstrated in the theoretical literature (Finus and Pintassilgo 2013; Weikard 2009); our calculations suggest that it also applies to more complex models incorporating non-linear relationships. As such, we view this as encouraging for climate change negotiations. While real-world heterogeneity may initially present an obstacle to cooperation, as it inevitably creates losers (along with the winners), there is a rich set of transfer and compensation schemes that can potentially turn heterogeneity from an obstacle into an opportunity, e.g. by burden sharing schemes (Lessmann et al. 2015) or introducing trade in emission permits (Jakob et al. 2014).

The analysis presented here shows that results regarding the stability of coalitions derived from models of international climate cooperation should be interpreted with some caution. Interestingly, for some coalitions, our analysis indicates that outcomes may be more positive than those suggested by deterministic analyses. Under uncertainty the likelihood of important coalitions like the OECD being internally stable might increase, and even large coalitions have a considerable chance to become internally stable when appropriate transfer schemes are in place.

The methods employed in this study limit the generalizability of our results. Results are derived from analysis of a small set of coalitions and should be tested in a wider sample of all possible coalitions. We chose to focus on a limited set of coalitions to identify the major effects and to examine the underlying mechanics, and while we would not expect qualitatively different results, an extension to a larger set of coalitions would strengthen our argument. Finally, while uncertainty propagation by Monte-Carlo analysis allows us to study the effect of real-world uncertainty on model results, uncertainty does not enter the

strategic decision making on whether or not to cooperate. This requires a modelling approach which incorporates decision making under uncertainty, e.g. stochastic programming. While our study reveals which uncertainties drive the stability of coalitions, future research could explicitly incorporate decision making under uncertainty in order to gain an in-depth understanding of how uncertainty influences their formation.

References

- Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers*, 46, 878-894.
- Barrett, S. (2003). *Environment and Statecraft: The Strategy of Environmental Treaty-Making*. Oxford: Oxford University Press.
- Barrett, S. (2005). The theory of international environmental agreements. In K.-G. Mäler, J. R. Vincent, (Ed.), *Handbook of Environmental Economics* (pp. 1457-1516). Amsterdam: Elsevier.
- Barrett, S. (2013). Climate treaties and approaching catastrophes. *Journal of Environmental Economics and Management*, 66(2), 235-250.
- Bosetti, V., Carraro, C., De Cian, E., Massetti, E., & Tavoni, M. (2013). Incentives and stability of international climate coalitions: An integrated assessment. *Energy Policy*, 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*, Special Issue, Hybrid modelling of energy –environmental policies: reconciling bottom-up and top-down, 27, 13-37.
- Bréchet, T., Gerard, F. & Tulkens, H. (2011). Efficiency vs. stability in climate coalitions: A conceptual and computational appraisal. *The Energy Journal*, 31(1), 49-75.
- Bréchet, T., Thénier, J., Zeimes, T. & Zuber, S. (2012). The benefits of cooperation under uncertainty: the case of climate change. *Environmental Modelling and Assessment*, 17(1), 149-162.
- Carraro, C., Eyckmans, J. & Finus, M. (2006). Optimal transfers and participation decisions in international environmental agreements. *The Review of International Organizations*, 1(4), 379-396.
- Carraro, C. & Siniscalco, D. (1993). Strategies for the international protection of the environment. *Journal of Public Economics*, 52(3), 309-328.
- Clarke, L., Edmonds, J., Krey, V., Richels, R., Rose, S. & Tavoni, M. (2009). International climate policy architectures: Overview of the EMF 22 international scenarios. *Energy Economics*, 31, 564-581.
- Crost, B., & Traeger, C. P. (2011). Risk and aversion in the integrated assessment of climate change. *CUDARE Working Papers*. Department of Agricultural & Resource Economics, University of California, Berkeley.
- Dellink, R., Dekker, T. & Ketterer, J. (2013). The fatter the tail, the fatter the climate agreement. Simulating the influence of fat tails in climate change damages on the success of international climate negotiations. *Environmental and Resource Economics*, 56(2), 277-305.
- Dellink, R. (2011). Drivers of stability of climate coalitions in the STACO model. *Climate Change Economics*, 2(2), 105-128.
- Dellink, R., Altamirano-Cabrera, J. C., Finus, M., von Ireland, E. & Weikard, H. (2004). Empirical background paper of the STACO model. <http://www.wageningenur.nl/web/file?uuid=9115e403-7832-48d6-ac1c-eb61f67edbec&owner=d1bd6906-08fd-4139-956f-b8004307a16e>. Accessed 01 August 2013.
- Dellink, R. & Finus, M. (2012). Uncertainty and climate treaties: Does ignorance pay? *Resource and Energy Economics*, 34(4), 565-584.
- Dellink, R., Finus, M. & Olieman, N. (2008). The stability likelihood of an international climate agreement. *Environmental and Resource Economics*, 39(4), 537-377.
- Dietz, S. (2011). High impact, low probability? An empirical analysis of risk in the economics of climate change. *Climatic Change*, 108(3), 519-541.
- Eyckmans, J. & Bréchet, T. (2012). Coalitions in the 18 region stochastic CWS model, presented at the Workshop on Modeling Climate Coalitions, Potsdam Institute for Climate Impact Research, February 8-9, 2012.
- 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.
- Hoel, M. (1992). International environmental conventions: the case of uniform reductions of emissions. *Environmental and Resource Economics*, 2(2), 141-159.
- Hwang, I. C., Reynès, F. & Tol, R. S. J. (2013). Climate policy under fat-tailed risk: an application of dice. *Environmental and Resource Economics*, 56(3), 415-436.

- IPCC (2001). *Climate Change 2001: Impacts, Adaptation, and Vulnerability*. Cambridge: Cambridge University Press.
- Jakob, M., Lessmann, K. & Wildgrube, T. (2014). The role of emissions trading and permit allocation in international climate agreements with asymmetric countries. *Strategic Behavior and the Environment*, 4(4), 361-392.
- Kolstad, C. D. (2007). Systematic uncertainty in self-enforcing international environmental agreements. *Journal of Environmental Economics and Management*, 53, 88-79.
- Kolstad, C. D. & Ulph A. (2008). Learning and international environmental agreements. *Climatic Change*, 89, 125-141.
- Kolstad, C. D. & Ulph A. (2011). Uncertainty, learning and heterogeneity in international environmental agreements. *Environmental and Resource Economics*, 50, 389 – 403.
- Kornek U., Lessmann K. & Tulkens H. (2014). Transferable- and non-transferable utility implementations of coalitional stability in integrated assessment models. *CORE Discussion Paper*, No. 35.
- Kornek, U., Steckel, J., Lessmann, K. & Edenhofer, O. (2017). The climate rent curse: new challenges for burden sharing. *Int Environ Agreements*, DOI:10.1007/s10784-017-9352-2
- Leimbach, M., Bauer, N., Baumstark, L. & Edenhofer, O. (2010). Mitigation costs in a globalized world: climate policy analysis with REMIND-R. *Environmental Modelling and Assessment*, 15(3), 155-173.
- Lessmann, K., Kornek, U., Bosetti, V., Dellink, R., Emmerling, J., Eyckmans, J., Nagashima, M., Weikard, H. & Yang, Z. (2015). The stability and effectiveness of climate coalitions: A comparative analysis of multiple integrated assessment models. *Environmental and Resource Economics*, 62(4), 811-836.
- 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.
- Na, S. & Shin, H. S. (1998). International environmental agreements under uncertainty. *Oxford Economic Papers*, 50(2), 173-1785.
- Nordhaus, W. D. & Yang, Z. (1996). A regional dynamic general-equilibrium model of alternative climate-change-strategies. *The American Economic Review*, 86(4), 741-765.
- Nordhaus, W. D. (1994). *Managing the Global Commons. The Economics of Climate Change*. London: MIT Press.
- Nordhaus, W. D. (2008). *A question of balance: weighting the options on global warming*. New Heaven: Yale University Press.
- Nordhaus, W. D. (2010). Economic aspects of global warming in a post-Copenhagen environment. *PNAS*, 107(26), 11721-11726.
- Nordhaus, W. D. (2011). The economics of tail events with an application to climate change. *Review of Environmental Economics and Policy*, 5(2), 240-257.
- Nordhaus, W. D. (2012). Economic Policy in the face of severe tail events. *Journal of Public Economic Theory*, 14(2), 197–219.
- Olieman, N. J. & Hendrix, E. M. T. (2006). Stability likelihood of coalitions in a two-stage cartel game: An estimation method. *European Journal of Operational Research*, 174(1), 333-348.
- Tavoni, M. & Tol, R. S. J. (2010). Counting only the hits? The risk of underestimating the cost of stringent climate policy. A letter. *Climatic Change*, 100, 769-778.
- Tol, R. S. T. (1995). The damage costs of climate change toward more comprehensive calculations. *Environmental and Resource Economics*, 5(4), 353-374.
- Tol, R. S. J. (2012). On the uncertainty about the total economic impact of climate change. *Environmental and Resource Economics*, 53, 97-116.
- Tol, R. S. J. (2013). Targets for global climate policy: An overview. *Journal of Economic Dynamics & Control*, 37, 911-928.
- Tol, R. S. J. (2014). Correction and update: The economic effects of climate change. *The Journal of Economic Perspectives*, 28(2): 221-225.
- Weikard, H. P. (2009). Cartel stability under optimal sharing rule. *The Manchester School*, 77, 575-593.
- Weitzman, M. L. (2007). A review of the stern review on the economics of climate change. *Journal of Economic Literature*, 45(3), 703-724.

- Weitzman, M. L. (2009). On modelling and interpreting the economics of catastrophic climate change. *The Review of Economics and Statistics*, 91(1), 1-19.
- Weitzman, M. L. (2010). What is the “damage function” for global warming – and what difference might it make? *Climate Change Economics*, 1(1), 57-69.
- Weitzman, M. L. (2012). GHG targets as insurance against catastrophic climate damages. *Journal of Public Economic Theory*, 14(2), 221-244.

Appendix

Table 6: Makeup of regions studied in the Model of International Climate Agreements

Model region	Countries
AFR	Sub-Saharan Africa without South Africa
CHN	China
EUR	EU27 countries
IND	India
JPN	Japan
LAM	All American countries but Canada and the US
MEA	North Africa, Middle Eastern and Arab Gulf Countries, Resource exporting countries of former Soviet Union (FSU), Pakistan
OAS	South East Asia, both Koreas, Mongolia, Nepal, Afghanistan
ROW	Non-EU27 European states without Russia, Australia, Canada, New Zealand and South Africa
RUS	Russia
USA	USA

Table 7: Summary statistics of the surplus distributions - by coalition and scenario

Coalition	b_1			$\theta_1\text{-glo}$			$\theta_1\text{-reg}$			θ_2			deter- ministic
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	
EJ	0.000014	0.000008	0.000038	0.000018	0.000012	0.000022	0.000016	0.000003	0.000054	0.000008	0.000003	0.000017	0.000007
ER	0.000334	0.000304	0.000127	0.000505	0.000422	0.000435	0.000360	0.000189	0.000871	0.000316	0.000211	0.000321	0.000309
EU	0.000223	0.000182	0.000155	0.000333	0.000253	0.000332	0.000293	0.000118	0.000785	0.000194	0.000104	0.000263	0.000171
JR	0.000094	0.000086	0.000038	0.000141	0.000118	0.000120	0.000083	0.000053	0.000099	0.000088	0.000060	0.000083	0.000087
JU	0.000024	0.000011	0.000060	0.000028	0.000018	0.000036	0.000019	0.000004	0.000056	0.000013	0.000003	0.000027	0.000008
RU	0.000462	0.000420	0.000175	0.000696	0.000579	0.000604	0.000428	0.000246	0.000786	0.000437	0.000289	0.000455	0.000421
EJR	0.000099	0.000088	0.000042	0.000149	0.000120	0.000135	0.000146	0.000065	0.000449	0.000091	0.000056	0.000107	0.000087
EJU	0.000132	0.000101	0.000109	0.000196	0.000142	0.000210	0.000179	0.000072	0.000526	0.000110	0.000050	0.000174	0.000093
ERU	0.000249	0.000223	0.000118	0.000378	0.000294	0.000362	0.000369	0.000160	0.000948	0.000225	0.000127	0.000288	0.000205
JRU	0.000144	0.000126	0.000067	0.000214	0.000171	0.000197	0.000171	0.000087	0.000420	0.000130	0.000078	0.000156	0.000120
OECD	0.000155	0.000135	0.000087	0.000237	0.000178	0.000240	0.000247	0.000105	0.000707	0.000137	0.000068	0.000201	0.000121
Major Em.	-0.001066	-0.001013	0.000238	-0.001591	-0.001285	0.001447	0.000026	-0.000202	0.003472	-0.000940	-0.000616	0.000977	-0.000918

Supplementary online material

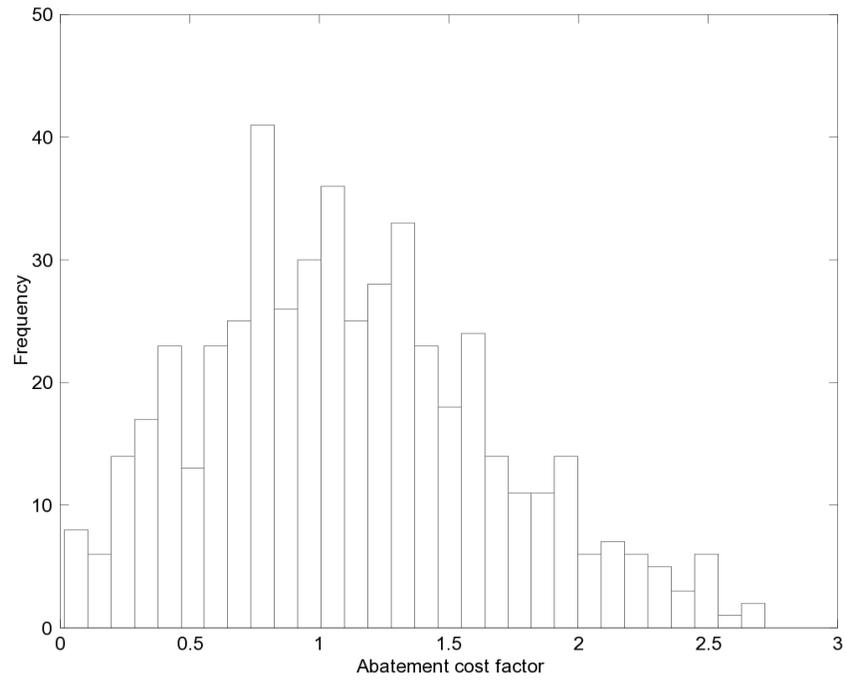


Figure 5: Histogram of the realization of the abatement cost factor k

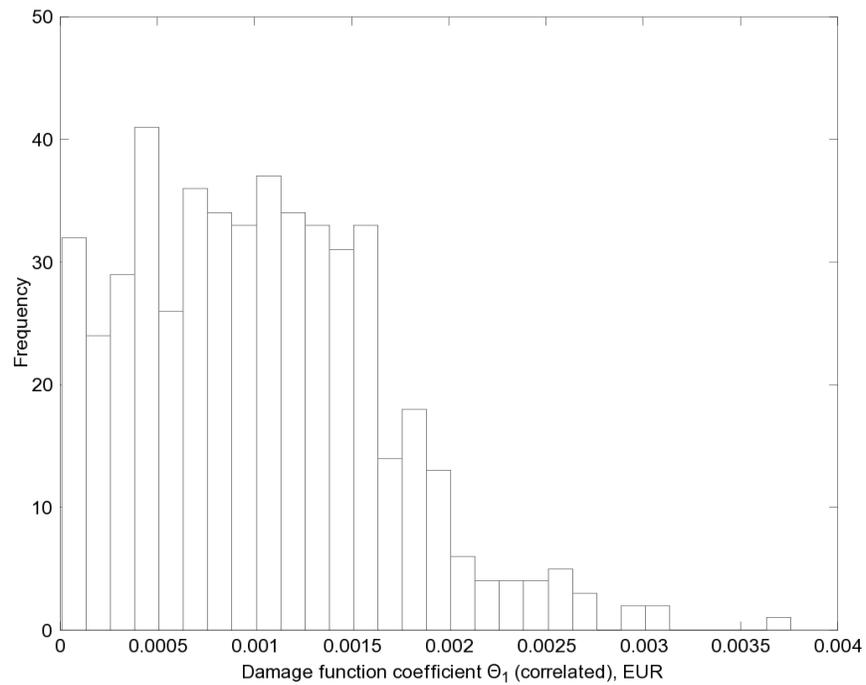


Figure 6: Histogram of the realization of Θ_1 in case of uncertainty in global damages for EUR

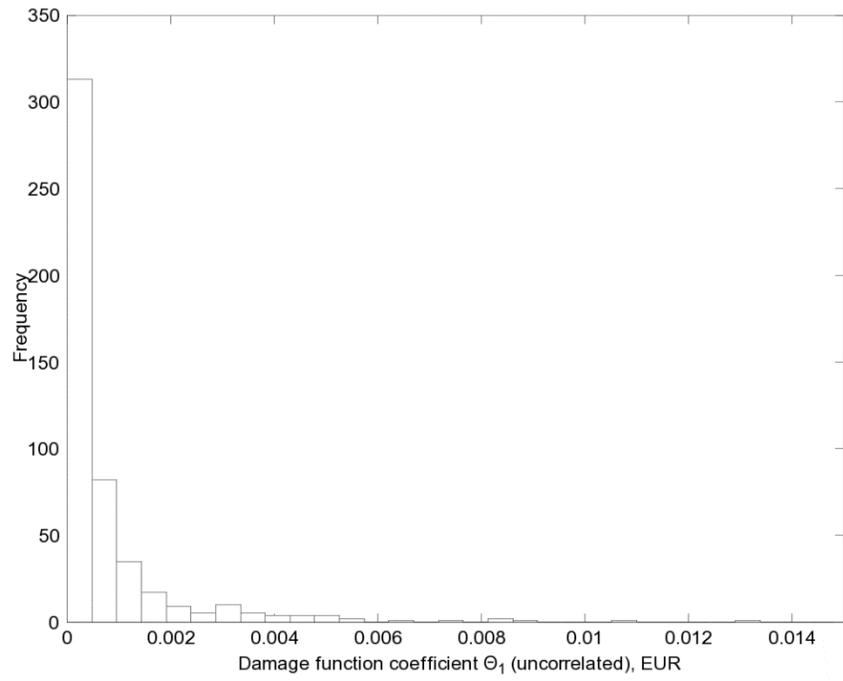


Figure 7: Histogram of the realization of Θ_1 in case of uncertainty in regional damages for EUR

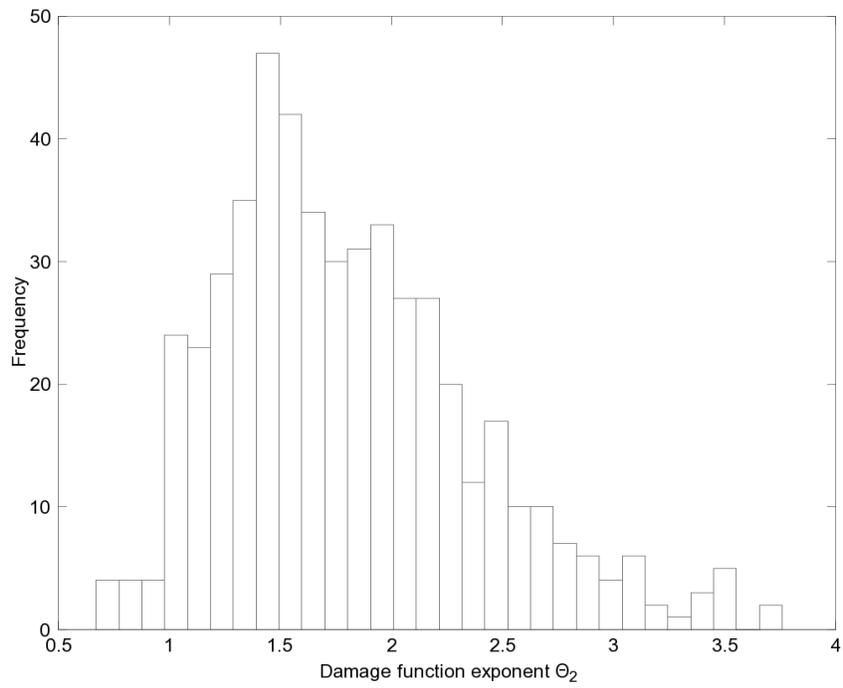


Figure 8: Histogram of the realization of Θ_2