Climatic Change (2022) 174:25
https://doi.org/10.1007/s10584-022-03436-6

Optimal climate policies under fairness preferences

Marco Rogna[1] · **Carla J. Vogt**[1]

Received: 8 October 2021 / Accepted: 9 September 2022 / Published online: 20 October 2022

Abstract

Integrated assessment models are tools largely used to investigate the benefit of reducing polluting emissions and limiting the anthropogenic mean temperature rise. However, they have been often criticized for their underlying assumptions, often leading to low levels of abatement. Countries and regions that are generally the actors in these models are usually depicted as having standard concave utility functions in consumption. This, however, disregards a potentially important aspect of environmental negotiations, namely its distributive implications. The present paper tries to fill this gap assuming that countries/regions have Fehr and Schmidt (*The Quarterly Journal of Economics* 114(3):817–868, 1999) (F&S) utility functions, specifically tailored for including inequality aversion. By adopting the RICE model, we compare its standard results with the ones obtained introducing F&S utility functions, showing that, under optimal cooperation, the level of temperature rise is lower in the last scenario. In particular, the peak temperature, reached in 2155 under standard assumptions and one or two decades later under F&S preferences with, respectively, heterogeneous and homogeneous F&S inequality aversion parameters (α and β), is of 2.86 ˚C in the former scenario and of 2.65 ˚C and 2.67 ˚C in the latter two. Furthermore, it is shown that potentially stable coalitions are easier to be achieved when F&S preferences are assumed. However, potentially stable coalitions are far from reaching environmental targets close to limiting the mean temperature rise below 2 ˚C despite the adoption of F&S utility functions. The 2 ˚C target is almost achieved in all scenarios when the payoffs in the F&S utility function are given by the per-capita consumption rather than utility of consumption, with F&S preferences and heterogeneous F&S inequality aversion parameters leading to a peak temperature rise of 2.04 ˚C.

Keywords Abatement · Climate policy · Inequality aversion · Paris agreement · RICE model

✉ Marco Rogna
marco.rogna@hs-bochum.de

Carla J. Vogt
carla.vogt@hs-bochum.de

[1] Bochum University of Applied Sciences, Am Hochschulcampus 1, 44801 Bochum, Germany

1 Introduction

In the current debate about climate change, there are two numbers that have become very prominent: 1.5 and 2. These are the two thresholds, expressed in degrees Celsius, representing the increase in mean temperature above pre-industrial era, that should not be passed. The former threshold is more environmentally conservative and generally preferred by ecologists (Knutti et al. 2016), whereas the latter is sometimes judged as more realistic given the prompt and firm actions, and the related costs, required to meet the 1.5° target (Jewell and Cherp 2020). Despite the debate on which of the two temperature targets should shape climate policies, it is now almost given for granted that no other target levels of peak temperature increase will be taken into serious consideration (Knopf et al. 2012).

However, before the widespread convergence on these two numbers, there has been a long debate on the optimal temperature target. Economic–environmental models, either in the form of integrated assessment models (IAMs) or computable general equilibrium models (CGEs), have been largely used to investigate the topic. Among IAMs, the DICE/RICE[1] model of Nordhaus (1992) has played a central role having been one of the first to include detailed equations to depict the relations between economic activities, emissions, temperature rise and climate damages.

According to this model, even when considering the fully cooperative scenario, the derived optimal mean temperature increase for the end of the twenty-first century is well above 2.5 °C (Nordhaus and Yang 1996). In the subsequent revisions of the model, from one side, the inclusion of more serious environmental damages induced a decrease of optimal mean temperature rise, but on the other, a more sensitive relation between emissions and climate change pushed in the opposite direction, obtaining a substantial parity. In the 2010 version of RICE, in fact, the optimal mean temperature rise under full cooperation is slightly below 3 °C (Nordhaus 2010).

Several reasons have been indicated for this substantial difference between the DICE-/RICE-derived prescriptions and the one provided by environmentalists and climatologists. Many critiques have been directed to the damage function that, even after the updates, has been judged to underestimate catastrophic and immaterial damages (Weitzman 2012; Howard and Sterner 2017) and not to account for uncertainty (Roughgarden and Schneider 1999; Diaz and Moore 2017). Another crucial aspect that plays a central role in determining optimal emissions is the inter-temporal distribution. A high value of the discount rate, as many judge the one adopted in DICE/RICE (Dietz et al. 2018), naturally leads to a low evaluation of future climate damages, thus calling for lower levels of abatement. Alternative configurations of inter-temporal distribution in the DICE model have proven to be determinant in sensibly decreasing the derived optimal level of temperature rise: e.g. Dennig et al. (2015), Botzen et al. (2018), Hänsel and Quaas (2018), Glanemann et al. (2020), Hänsel et al. (2020), and Gazzotti et al. (2021).

With climate damages and inter-temporal discounting having been the two most controversial aspects over which the academic debate has flourished, there is another important theme that has been partially shadowed by the mentioned debate: equity. Besides inter-generational equity represented by discounting, the sharp difference in economic attainment and emissions level of the 12 (originally 6) countries/regions in the RICE model calls for a

[1] The DICE model (dynamic integrated model of climate and the economy), firstly presented in Nordhaus (1992), considers the world as a whole, whereas the RICE model (regional integrated model of climate and the economy), firstly presented in Nordhaus and Yang (1996), decomposes it into countries/regions.

thorough consideration of this aspect. The RICE model attempts to consider it by including the elasticity of marginal utility (EMU) of consumption into the countries/regions utility function. Actors with larger levels of per-capita consumption will enjoy lower levels of utility increase for additional units of consumption. However, we argue that this standard formulation of a concave utility function in per-capita consumption may not be adequate to fully capture the disutility caused by inequality.

In particular, in the present paper, we propose a more systematic inclusion of equity concerns based on insights from behavioural economics in the form of Fehr and Schmidt (1999) (F&S) utility functions. In fact, these capture the phenomenon that people compare themselves to others and possibly derive disutility if their payoff is below or above other players' payoffs. This utility function is in line with numerous observations made in experimental economics and it has proven to be successful in explaining observed behaviours in bargaining and cooperation games (see Fehr and Schmidt (2006) for a review).

As mentioned earlier, the original RICE model already has a component that, at least indirectly, accounts for inequality aversion, namely the curvature of the utility function, governed by the parameter EMU (η in the model). This component, however, represents the effect of diminishing marginal utility of consumption, that is a very standard assumption in economics. The consideration for inequality is only indirect, and stems from the fact that an additional unit of consumption has a higher value for lower levels of wealth than for higher ones. This is therefore the inequality aversion of a hypothetical social planner. But this is the same assumption adopted for the utility function characterizing participants to games in experimental labs that generally fail to explain the results of games such as the Ultimatum or the Dictator game (Fehr and Schmidt 1999). The utility function proposed in Fehr and Schmidt (1999) is an attempt to overcome the shortcomings of a standard utility function in predicting the results of the mentioned classes of games, performing rather well in this objective. In other words, this utility function represents inequality aversion from the point of view of individuals. It is therefore interesting to apply it to a context such as the one of climate negotiations where fairness considerations seem to naturally play an important role.

One possible objection could be that countries are different from individuals and results from economic experiments are almost always based on the latter. To counter this objection, we refer to the median voter argument. Governments interested in being re-elected have to obey, at least to a certain extent, to the according preferences of their pivotal voter, which may be approximated by the median voter in democratic political systems. Thus, the median voter's fairness preferences determine, in the end, which burden sharing in a climate treaty is acceptable for a country's delegation. Collective preferences are therefore related to individual preferences. It is the F&S utility function of median voters that, ultimately, determines the outcome of international climate negotiations.

In our simulations, we use as base values for the inequality aversion parameters present in the F&S utility function the median values reported in Fehr and Schmidt (2006). We further test more extreme values in order to cover the whole spectrum shown in the mentioned paper and take into consideration the work of Dannenberg et al. (2010) as a source for additional values to be tested.

Our results show that, even keeping the original inter-temporal discounting and climate damages as in RICE v2013, the adoption of F&S utility functions sensibly reduces the peak global level of emissions in the fully cooperative case. The peak temperature increase, in fact, is approximately 2.67 ˚C in 2175 under F&S preferences with homogeneous inequality aversion parameters — derived from Fehr and Schmidt (1999) — compared to 2.86 ˚C in the standard model run, reached in 2155. When considering F&S preferences with heterogeneous inequality aversion parameters — derived from Dannenberg et al. (2010) — peak temperature is further decreased to 2.65 °C, reached in 2165. We further show that cooperation, in the form of stable coalitions, is significantly enhanced by the adoption of F&S utility functions. However, self-sustaining coalitions, even when F&S preferences are assumed, cannot grant a peak temperature increase lower than 4.5 °C.

2 Literature review

Integrated assessment models have been one of the main instruments to investigate the costs and benefits of taking actions to counteract human–induced climate warming. The large use of these modeling tools in the Intergovernmental Panel on Climate Change (IPCC) reports well testifies this fact (Rosen 2015; Hansson et al. 2021). Several models, with different underlying assumptions, focus and databases, have been proposed, among which MIRAGE (Easter et al. 2004), WITCH (Bosetti et al. 2006), MAgPIE (Dietrich et al. 2019) and POLES (Keramidas et al. 2017) are prominent examples. As mentioned earlier, the DICE/RICE model plays a central role in this list being among the first attempts to link the whole world economy to the earth's climate system and depicting the influence that each of the two has on the other.

From its original formulation in 1992 (Nordhaus 1992), the DICE model has been subject to several major revisions along time (Nordhaus 2018b). The decomposition of the world into 6 countries/regions in 1996 with the introduction of the RICE model (Nordhaus and Yang 1996) has been one of the major changes. To this, several updates have followed in 2000 (Nordhaus and Boyer 2000), in 2007 (Nordhaus 2007), in 2010 (Nordhaus 2010), in 2013 (Nordhaus 2013) and in 2016 (Nordhaus 2018a), increasing the number of countries/regions considered in RICE from 6 to 12, refining the economic and damage equations and changing the underlying database.

Despite the model updates in its economic side, the optimal emissions of CO_2, and the associated temperature rise, has remained far above the targets of 1.5 °C/2 °C settled in the Paris Agreement. Note that this is not a peculiarity of DICE/RICE, since several other models provide an estimate of the SCC leading to carbon prices that are too low to meet the 1.5 °C/2 °C targets, as shown in Tol (2019) and in Ackerman and Munitz (2016), despite there are notable exceptions to this trend such as Dennig et al. (2015) and Hänsel et al. (2020) and Budolfson et al. (2021). This has generated a wave of critiques directed towards IAMs. As mentioned in the introduction, the underestimation of catastrophic and immaterial damages (Weitzman 2012; Howard and Sterner 2017), the lack of account for uncertainty (Roughgarden and Schneider 1999; Diaz and Moore 2017) and a high inter-temporal discounting (Dietz et al. 2018) are the main targets of the mentioned critiques.

Several attempts have been made to overcome these perceived shortcomings. De Bruin et al. (2009) separate the mitigation and the adaptation costs in the DICE model, while Michaelis and Wirths (2020) consider the rate of temperature rise in addition to its level of increase, showing that ignoring the former aspect may substantially underestimate climate damages. Still adopting DICE, Tol (1994) proposes a different method for incorporating intangible damages, Botzen and van den Bergh (2012) assume an alternative specification of the damage function and Ackerman et al. (2010) attempt to model catastrophic damages and their distribution. Finally, Dietz and Asheim (2012) and Botzen et al. (2018) investigate different forms of inter-temporal discounting. Generally, these modifications lead to a lower level of temperature rise under optimality and earlier and tighter efforts for decarbonizing the economy.

On the game–theoretic side of climate change analysis, there are a number of papers that have tested the possibility of going beyond the standard assumption of pure self-interest, embracing insights derived from behavioural and experimental economics. Lange and Vogt (2003) assume preferences á la Bolton and Ockenfels (2000), van der Pol et al. (2012) add a component of pure altruism while Vogt (2016) and Rogna and Vogt (2020) consider a utility function based on F&S preferences. A common finding of these papers is that, despite a general increase in stability once abandoning standard preferences, cooperation is not dramatically enhanced without transfers. Being game–theoretic papers, however, they all portray a very stylized and scarcely realistic representation of both the economic and the environmental side.

The present paper aims at including non-standard preferences into a dynamic and more complex model framework such as the RICE model, thus filling a current gap of the literature. The choice is for F&S preferences whose consideration of aversion for both advantageous (altruism) and disadvantageous (envy) inequality has proven to be able to capture several deviations from standard economic theory observed in laboratory experiments (Fehr and Schmidt 2006). In particular, we are interested in observing which is the effect of this alternative specification of the utility function both on the optimal level of emissions abatement, and, therefore, on the temperature rise, and on the stability of climate coalitions.

3 The RICE model with Fehr and Schmidt preferences

Our starting model, also used as benchmark, is RICE v2013. A synthetic description of all its variables (endogenous and exogenous) and parameters can be found in the Supplementary Materials, Section A1. Its basic equations, instead, can be found in Section A2, in the Supplementary Materials. Compared to the original RICE v2013, one modification has been introduced. It is the reduction of the number of control variables to one, namely the level of proportional abatement ($\mu_{i,t}$ in the model), while the original model has the saving–investment rate as an additional control variable. In particular, we treat the saving–investment rate — $\sigma^I_{i,t}$ in the model, equation (A4) — as exogenous, deriving its value from running the original model in the non-cooperative scenario. The reason for this choice is to simplify the model given that the introduction of F&S preferences adds a considerable computational burden. Furthermore, leaving a single control variable, abatement, simplifies and renders more explicit the interpretation of results.

Except for the modification just explained, the set of equations in Section A2 in the Supplementary Materials faithfully reproduces the original RICE model v2013, with population size used as weighting factor but without Negishi weights. This will be used as our benchmark scenario, without adding any exogenous environmental target or any price for CO_2. As mentioned earlier, the modification proposed in this paper is the introduction of F&S utility functions. The following is their mathematical definition:

$$U_i = \pi_i - \frac{\alpha_i}{n-1}\sum_{j\in I^+}(\pi_j - \pi_i) - \frac{\beta_i}{n-1}\sum_{k\in I^-}(\pi_i - \pi_k), \quad (1)$$

where i is a generic player of set N, whose cardinality is represented by n, π is the payoff of a player, I^+ and I^- are the sets of players having, respectively, a payoff higher and lower than player i and, finally, α_i and β_i are the parameters representing the aversion for disadvantageous, the former, and for advantageous, the latter, inequality. The expression $\frac{\alpha_i}{n-1}\sum_{j\in I^+}(\pi_j - \pi_i)$, where the component inside the round brackets is always positive since $\pi_j > \pi_i$ by definition, represents the disutility suffered by player i for having a payoff lower than all players j (envy). Similarly, the expression following β_i, necessarily positive by definition as well, represents the disutility for advantageous inequality (altruism). Despite suffering disutility for having more than others may seem less intuitive than for having less, the β component is necessary to explain the results of Dictator games such as donations to players that have no strategic power. Finally, note that the payoffs referred to in (1) are utility payoffs (as will be seen in (2)) and not monetary payoffs as in the standard F&S utility function. For this reason, it would be more precise to call the utility function in (1) an F&S–style rather than a proper F&S utility function.

Willing to adopt the F&S utility function in the RICE v2013 model, equation (A2) in the Supplementary Materials must be substituted by the following two equations:

$$\begin{aligned} \pi_{i,t} &= \frac{1}{1-\eta}\left(\frac{C_{i,t}}{L_{i,t}}\right)^{1-\eta} + 1, & \forall i \in N, \forall t \in T, \\ U_{i,t} &= \pi_{i,t} - \frac{\alpha_i}{n-1}\sum_{j\in I^+}(\pi_{j,t} - \pi_{i,t}) - \frac{\beta_i}{n-1}\sum_{k\in I^-}(\pi_{i,t} - \pi_{k,t}), & \forall i \in N, \forall t \in T, \end{aligned} \tag{2}$$

where $\pi_{i,t}$ is simply defined as the value of per-capita consumption. The objective of the optimization problem of each player, instead, remains the same in the standard and in the F&S cases:

$$W_i = \sum_{t\in T}\frac{U_{i,t}}{(1+\rho)^{ts\times t}}, \qquad \forall i \in N \setminus C; \tag{3}$$

$$W_i = \sum_{i\in C}\sum_{t\in T}\frac{POP_i}{\sum_{j\in N} POP_j}\frac{U_{i,t}}{(1+\rho)^{ts\times t}}, \qquad \forall i \in C. \tag{4}$$

The former represents the maximization objective of a player non-member of any coalition, whereas the latter the one of a coalition member, that will then maximize the welfare of the whole coalition rather than solely its own. For the parameters η (coefficient of marginal utility of consumption) and ρ (coefficient of inter-temporal discounting), the following values, taken from Nordhaus and Sztorc (2013), have been, respectively, used: 1.5 and 0.015. Note that, in the cooperative objective function, the utility of each country is weighted by its population size (*POP*), so to reflect their bargaining power and their per-capita utility. As mentioned earlier, only population weights are used as weighting factors, while most of the implementations of the RICE model further use Negishi weights (Nordhaus and Boyer 2000; Nordhaus and Sztorc 2013). While all the parameters of the model can be retrieved from the documentation of RICE, the addition of the new utility function brings the burden of setting values for α and β. The next sub-section is dedicated to describe our choice for these values.

3.1 The choice of *α* and β

Since the values of α and β represent the intensity with which the disutility from disadvantageous and advantageous inequality is felt, they are of crucial importance in the present paper. Several works have tried to estimate them. In particular, the original work of Fehr and Schmidt (1999) provides some estimates of these two parameters, retrieved by a sort

Table 1 Mean values of α and β from Dannenberg et al. (2010)

	α	β	Countries/regions
EU	0.388	0.525	EU
G8	0.275	0.503	USA, JAP, RUS, OHI
G77	0.472	0.603	EUR, CHI, IND, MEST, AFR, LAM, OTH

Partial reproduction of Table 3 in Dannenberg et al. (2010)

of backward induction, as to say by finding that values of α and β capable of explaining the deviations from standard theory reported in experimental economics papers. For α, Fehr and Schmidt (1999) report an interval of [0,4.5], with 0.833 as median value, whereas for β the interval is [0,0.6], with 0.288 as median. Subsequent studies, such as Dannenberg et al. (2010), Blanco et al. (2011) and Ponti and Rodriguez-Lara (2015), have tried, through ad hoc experiments, to determine the values of α and β, roughly confirming the intervals provided in Fehr and Schmidt (1999). In Rogna and Vogt (2020), the median values of α and β as provided in Fehr and Schmidt (1999) are used as base case and countries/regions are assumed to be homogeneous with respect to both parameters. In the present paper, the same approach is used, due to the relevance of these two values in the literature. Therefore, our starting point will be setting $\alpha = 0.833$ and $\beta = 0.288$. However, we will conduct a thorough sensitivity analysis letting the values of α and β vary in the whole range provided in Fehr and Schmidt (1999): $\alpha \in [0,4.5], \beta \in [0,0.6]$. We will also test the mean values experimentally obtained by Dannenberg et al. (2010). These have the peculiarity to be regionally specific (EU, G8, G77), so it is possible to add a degree of heterogeneity in the α and β parameters of the RICE countries/regions. Table 1 reports the mean values of α and β as reported in Dannenberg et al. (2010). The last column of the table shows the countries/regions of the RICE model to which the respective values of α and β have been imputed. A caveat must be remembered. As mentioned earlier, the payoff functions inside the F&S utility function are utility rather than monetary payoffs. The estimated values of α and β parameters just reported have been retrieved through monetary payoffs; therefore, they may not suit perfectly the present case. However, lacking estimates for the inequality aversion parameters with utility payoffs, we have to rely on these values.

4 Simulations and results

In our basic simulations, three scenarios are benchmarked, the original RICE v2013 model with population rather than with Negishi[2] weights, the modified version with F&S preferences and homogeneous α and β parameters — with values, respectively, of 0.833 and of 0.288 — and the version with F&S preferences, heterogeneous α s and βs, with values given in Table 1. The simulation is run from 2015 to 2205[3], with a 10-year time step. In all three cases, all possible coalitions have been examined, meaning 4095 coalitions since

[2] A comparison with Negishi weights would surely be interesting, but the number of cases to compare would increase rendering the analysis of results more confusing.

[3] Actually, we report the results until the year 2205, but we have run the simulation until 2305 to reduce the last periods drop in abatement consequent to the "end of time" effect.

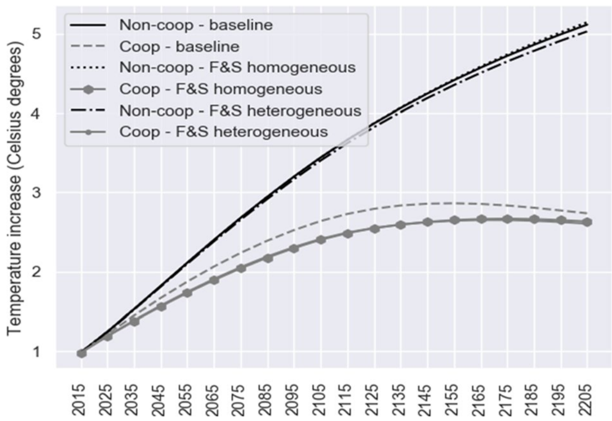

Fig. 1 Temperature rise under different preference specifications. Baseline: RICE v2013 model with standard utility function. The temperature trajectories for "Coop – F&S homogenous" and "Coop – F&S heterogenous" are almost identical and can therefore not be clearly distinguished in the figure

the model features 12 countries/regions. The way of solving the model follows the original algorithm described in Nordhaus and Yang (1996), while http://www.pyomo.org/Pyomo, a Python package for modeling optimization problems, has been used for the computation.[4]

The first thing to be examined is the temperature increase above the pre-industrial level obtained in the different scenarios. Each of the three cases mentioned above is further subdivided into two: the cooperative case (grand coalition in game–theoretic jargon) and the non-cooperative case, where no multi-countries coalition is formed. Basically, in the non-cooperative case, each player's maximization objective is represented by (3), with $C = \varnothing$, whereas in the cooperative case, each player's objective function is represented by (4), with $C = N$. From Fig. 1, it is possible to observe significant differences among the various scenarios. Clearly, the three cooperative cases lead to a lower level of final temperature rise than their non-cooperative counterparts. However, it is interesting to note as, with the standard utility functions of RICE, even under full cooperation, the peak rise in temperature, reached in 2155, is of 2.86 ˚C circa, whereas with F&S preferences and heterogeneous α s and βs, the rise is far more modest, approximately 2.65 ˚C, attained in 2165.

Further note that the cooperative case with homogeneous α s and βs is in between, with a peak temperature rise in 2175 equal to 2.67 ˚C circa. In the non-cooperative scenario, instead, the cases of F&S preferences with homogeneous inequality aversion and the baseline scenario (standard utility functions) lead to a similar temperature rise in 2205: 5.14 ˚C the former and 5.11 ˚C the latter. The case of F&S preferences and heterogeneous inequality aversion parameters confirms to have the lowest temperature rise, namely 5.03˚C in 2205. Note that, in the non-cooperative scenarios, temperatures keep rising until the simulation end period, namely 2305, with peak temperatures rise slightly above 6 ˚C in all cases. Global level of emissions, expressed in GTC, can be seen in Figure A31, while per country/region emission in Figure A32, in Supplementary Materials. Taken together, the introduction of F&S preferences does not lead to significant differences in the non-cooperative case, while the differences in the fully cooperative scenario are more marked.

With this being the level of temperature rise, it is also interesting to see how much abatement is undertaken by each country/region in each scenario. This is represented in Fig. 2, where all

4 The model and the data used for the simulation are available on a GitHub repository at this https://github.com/white-heomoi/RICE13_pyomolink.

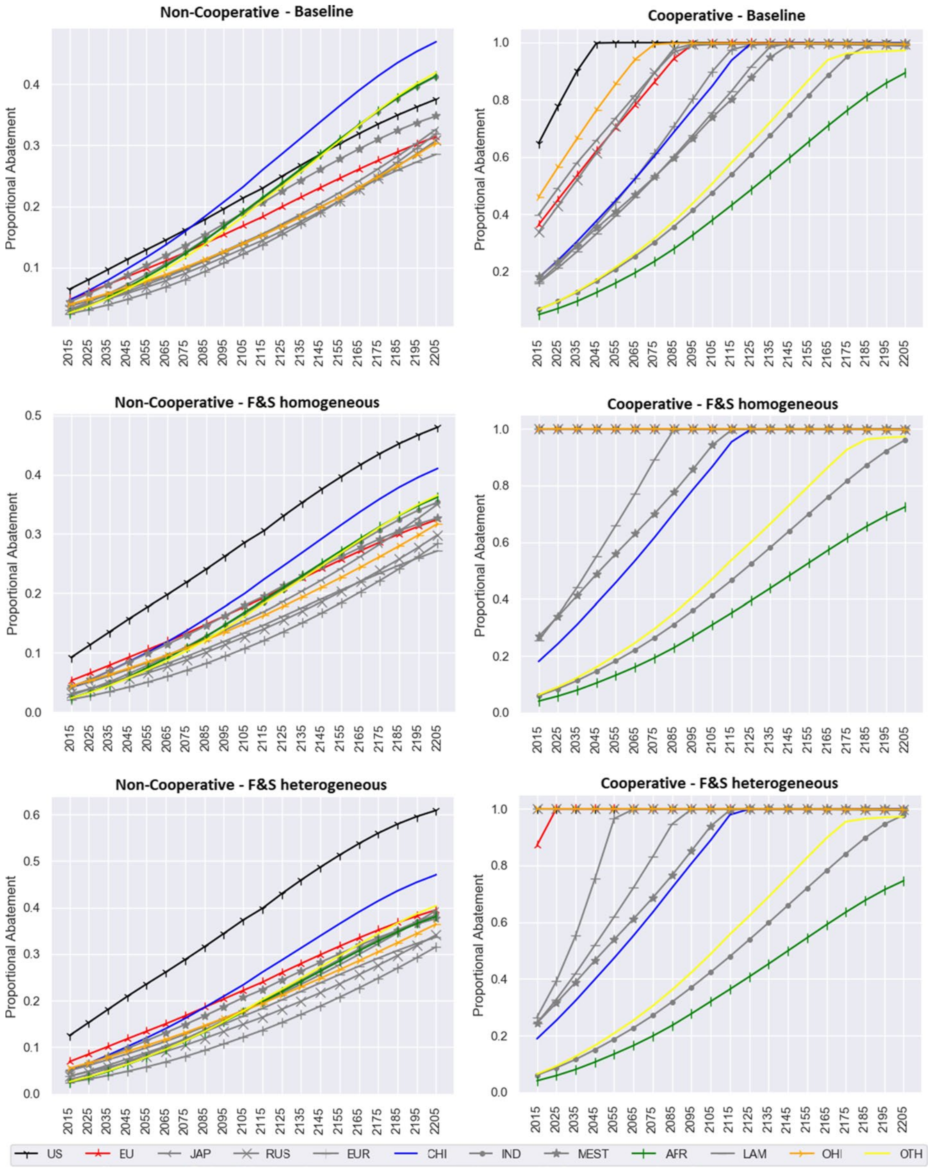

Fig. 2 Abatement share under different preference specifications

six scenarios are reported. Abatement is represented in proportional terms, so that 1 means a total abatement of polluting emissions. It is interesting to notice that there are only slight differences among the two scenarios with F&S preferences, with several countries/regions fully abating from the first time period, while this does not happen for any player in the baseline. Note how in the F&S heterogeneous case both Europe and Eurasia fully abate at a later stage compared with the F&S homogeneous case. The lower temperature peak in the former case, therefore, is due to the slightly stronger abatement efforts of China (CHN), other Asian countries

(OTH), India (IND) and Africa (AFR) that totally compensate the lower efforts of Europe (EU) and Eurasia (EUR). China (CHN), in particular, gives a substantial contribution to achieve this outcome, confirming to be a key player in the global emissions game.

The lower level of α s in the heterogeneous compared to the homogeneous case, together with higher levels of β, particularly for poorer countries/regions, in the former case, lead to lower efforts for some richer regions such as Europe and Eurasia, but tighter ones for relatively poorer regions: India, Africa, other Asian countries and China. Despite the α value for the G77 regions (the poorest ones) is the highest under heterogeneous inequality aversion parameters, thus reducing the environmental efforts of weak players in face of strong ones, the fact that it is almost half the value retrieved by Fehr and Schmidt (1999) lead to stronger abatement of weak players in the F&S heterogeneous scenario. In the non-cooperative cases, instead, the level of abatement is dramatically lower, hardly reaching 50%.

It has to be remembered that the present model does not feature any assumption about future very cheap and clean technologies for energy production or the presence of a technology allowing to remove CO_2 from the atmosphere. These assumptions were actually present in the original RICE model (v2013) (Nordhaus and Sztorc 2013), but we preferred to exclude them since potentially controversial. Under such assumptions, it may be possible to reach, particularly in the heterogeneous F&S scenario, a peak temperature rise below 2 ˚C.

4.1 Stability of coalitions

After having examined the temperature and abatement trajectories in the two extreme cases, namely no cooperation and grand coalition, the focus will now be placed on the stability of coalitions. We will therefore consider all the intermediate possibilities of aggregation of countries/regions and look at their stability using the well known criterion of d'Aspremont et al. (1983): a coalition is stable when both internal — no member of the coalition has an incentive to leave — and external — no outsider has an incentive to join the coalition — stabilities are satisfied. The condition of potential internal stability (PIS) is further considered, being it a weaker version of internal stability according to which a coalition satisfies PIS if it can be internally stabilized via transfers among its members. Finally, the case in which PIS is attained together with external stability is further shown. Here is the mathematical definition of the stability concepts:

$$
\begin{aligned}
&W_i(C) \geq W_i(C \setminus i), \quad \forall i \in C. && \text{Internal Stability}\\
&W_i(C) \geq W_i(C \cup i), \quad \forall i \in N \setminus C. && \text{External Stability}\\
&\sum_{i \in C} W_i(C) \geq \sum_{i \in C} W_i(C \setminus i), \quad \forall i \in C. && \text{PIS}
\end{aligned}
$$

Table 2 shows the number of stable coalitions for each scenario divided into coalition size and for the considered stability condition. The grand coalition is excluded from the table but it is neither internally stable nor potentially internally stable in any of the three considered scenarios. As largely shown in the game–theoretic literature on environmental coalitions formation (e.g. Carraro and Siniscalco 1993; Barrett 1994), the number of stable coalitions is very low, particularly due to the difficulty of reaching internal stability. Furthermore, stable coalitions have a very modest size. However, it is interesting to note the difference existing between the various scenarios, with the baseline case having just two stable coalitions, whereas when F&S preferences and heterogeneous α s and βs are considered, the number increases to 4.

Table 2 Number of stable coalitions under different scenarios

Coalition size	2	3	4	5	6	7	8	9	10	11	Total
Baseline scenario (RICE v2013 with standard utility functions)											
Fully stable	1	1	0	0	0	0	0	0	0	0	2
Internally stable	4	2	0	0	0	0	0	0	0	0	6
Externally stable	7	21	48	175	302	316	199	75	51	10	1204
PIS	13	23	41	11	0	0	0	0	0	0	88
PIS & externally st.	1	3	9	0	0	0	0	0	0	0	13
F&S preferences (α =0.833,β =0.288)											
Fully stable	3	0	0	0	0	0	0	0	0	0	3
Internally stable	3	1	0	0	0	0	0	0	0	0	4
Externally stable	4	21	98	112	174	180	130	89	36	8	852
PIS	55	89	146	171	157	46	0	0	0	0	664
PIS & externally st.	1	12	32	25	28	5	0	0	0	0	103
F&S preferences with heterogeneous α and β values											
Fully stable	2	2	0	0	0	0	0	0	0	0	4
Internally stable	2	1	1	0	0	0	0	0	0	0	4
Externally stable	1	12	37	92	161	201	164	96	38	9	811
PIS	28	200	328	285	142	32	3	0	0	0	1018
PIS & externally st.	0	8	15	21	31	12	0	0	0	0	87

First of all, we can notice that the two scenarios having F&S preferences, with homogeneous and heterogeneous inequality aversion parameters, are very similar in terms of stable coalitions, for most of the stability criteria. The heterogeneous case has more PIS stable coalitions, but a slightly lower number of externally stable coalitions and a more significant decrease in coalitions that are PIS + externally stable. When compared with the baseline, we can notice that F&S preferences lead to significantly higher numbers of PIS stable coalitions, but to less externally stable coalitions. However, the number of PIS + externally stable coalitions is also significantly higher under F&S preferences. When considering full and internal stability, instead, the difference among the various scenarios is trivial since the number of coalitions satisfying these conditions is very limited in all of them.

Besides the total number of stable coalitions, it is also important to consider the maximal size of coalitions that are stable under the various criteria. Once again, no noticeable differences are present when full or internal stabilities are considered, except that F&S preferences with heterogeneous α s and βs lead to one internally stable 4 players coalition, whereas the maximal size in the other scenarios is of 3 players. When considering external stability, all scenarios have most of the 11 players coalitions satisfying this condition. A significant difference is observable when the last two conditions, PIS and PIS plus external stability, are taken into account. The maximal size of coalitions respecting these two conditions is 5, for the former condition, and 4, for the latter, in the baseline scenario. This number increases to 7 for both conditions under F&S preferences and homogeneous inequality aversion parameters. With heterogeneous α s and β, instead, we have three 8 players coalitions that are PIS stable and twelve 7 players coalitions that are PIS and externally stable. It is important to notice that this last condition is very important since it implies that a coalition could be fully stabilized through internal transfers.

4.2 Discussion of coalition stability

Vogt (2016) provides a condition to obtain internal stability under F&S preferences. Remembering that internal stability is defined as $U_i(C) \geq U_i(C \setminus i)$, the condition reads as:

$$\frac{\alpha_i}{n-1}\left(\sum_{j\in I^+}\left[\pi_j(C\setminus i)-\pi_i(C\setminus i)\right]-\sum_{j\in I^+}\left[\pi_j(C)-\pi_i(C)\right]\right)\geq \pi_i(C\setminus i)-\pi_i(C)+$$
$$\frac{\beta}{n-1}\left(\sum_{k\in I^-}\left[\pi_i(C)-\pi_k(C)\right]-\sum_{k\in I^-}\left[\pi_i(C\setminus i)-\pi_k(C\setminus i)\right]\right).$$

If this inequality is satisfied for each member of coalition C, then C is internally stable. Our numerical simulation has fully confirmed this result, that, besides providing a method to check for coalition stability, is also useful to understand the reasons inducing to internal stability.

As mentioned in Vogt (2016), the bracket term on the LHS of this inequality reflects the development of disadvantageous inequality for player i after leaving coalition C, with the sign of this term being unknown a priori: disadvantageous inequality may be either decreased or increased by i's exit. The RHS, instead, includes the material gain obtained by player i in exiting, plus the development of advantageous inequality. By leaving a coalition, and in absence of transfers, a player can generally expect to improve her absolute as well as her relative payoff. This incentive increases with increasing values of α s. However, this effect can be contrasted, and even reversed, by introducing well tailored transfer schemes as in Rogna and Vogt (2020). In particular, these transfers are effective only if they are capable of sufficiently reducing disadvantageous inequality for a certain set of players. Note that increasing levels of α's for decreasing per-capita income[5] increases the incentive to leave for poorer countries. These countries, however, are also the ones that are easier to satisfy through transfers since they have a higher marginal utility of consumption and, therefore, even modest monetary transfers are able to generate substantial changes in utility. This is most likely the explanation why the number of PIS coalitions is strongly increased when introducing inequality aversion.

One thing to note is that the required transfers to stabilize a coalition are here measured in utils. However, countries are likely bounded to make monetary transfers. Since a monetary transfer impacts directly the utility of a country modifying its per-capita income, but also its α and β components together with the ones of other players, it is necessary to envisage a rather complex transfer scheme in order to have monetary transfers that are equivalent to direct utility transfers. Such a mapping is beyond the scope of the present paper.

It seems not possible, instead, to individuate a clear pattern of PIS and PIS & Ext. stable coalitions. In all scenarios, the coalitions that are at least potentially stable are very heterogeneous in terms of composition. Some are composed by members with similar characteristics, such as per-capita income and emissions, whereas others are very diverse. Furthermore, there does not seem to be a clear prevalence of PIS coalitions neither of industrialized nor of developing regions. It is therefore not possible to relate stability with some specific characteristics such as vulnerability to climate damages or emissions

[5] This is the case of F&S preferences with heterogeneous α s and βs, since the level of α for G77 countries is the highest.

intensity. One possibility to detect some patterns could be to perform a series of systematic and small variations in the parameters, a task that is left for a future work.

4.3 Temperature change under stable coalitions

Besides leading to a lower temperature rise in the cooperative case, that is, however, not stable, the F&S scenarios allow for a significant increase in PIS stable coalitions and in the number of coalitions that can be fully stabilized through transfers. However, the simple number of stable coalitions may be considered as scarcely informative, since it does not tell much about the outcomes. In particular, it is interesting to see, in the various scenarios, which is the best and realistically achievable environmental target. On this regard, we have to anticipate that the results are rather grim, since no stable or potentially stable coalition is able to prevent a continuous temperature rise until the year 2305. This is true for the seven players PIS & Ext. stable coalitions and also for the eight players coalition that is only PIS stable. Therefore, no PIS or PIS plus externally stable coalition is able to impede a final temperature rise well above 3°.

We will then compare the various scenarios taking 2155 as the reference year. Remembering that no breakthrough technological improvements have been assumed in our model, one might expect that the simulation after a certain period in the future becomes overly pessimistic due to the under-evaluation of the presence of cheap and clean technologies. The year 2155 should then be considered as a threshold, rather conservative, around which the appearance of a breakthrough technology may be supposed as extremely likely. Not surprisingly, the three PIS stable coalitions of eight countries/regions reached under F&S preferences with heterogeneous α s and βs are the ones achieving the lowest temperature rise. Among them, the coalition between the USA, Europe, Japan, China, Russia, India, Middle East and other Asian countries is the one performing better with a temperature rise of 3.33 °C. If we restrict the attention to potentially stable coalitions — namely, PIS + Ext. stability — the best performance is reached by the seven players coalition under F&S preferences and heterogeneous α s and βs among all industrialized countries except Japan plus China, Eurasia, India and Middle East. The temperature rise in 2155 is of 3.55 °C.

As mentioned, the result is rather pessimistic since it is more than double the 1.5 °C of temperature rise objective of the current international negotiations. However, these coalitions lead to a temperature rise in 2155 that is more than 1° lower than the one reached in the non-cooperative scenario with F&S preferences. If the lowest temperature rise has been reached under F&S preferences and heterogeneous α and β parameters, there is not a very significant difference with the other scenario. Among the PIS and PIS & Ext. stable coalitions obtained with F&S preferences and homogeneous inequality aversion parameters, in fact, the most environmentally successful one is a seven players coalition leading to a temperature rise of 3.59 °C in the reference period. This is the coalition among the USA, China, Africa, India and other industrialized and other Asian nations. In the baseline scenario, instead, the best result is achieved by a PIS stable coalition of five players, namely Russia, Eurasia, China, India and other industrialized countries, with a temperature rise of 4.15 °C. All the fully stable coalitions, given their limited size, do not achieve significant environmental advantages compared to the non-cooperative case.

Also in this case, it is difficult to relate particular characteristics of countries/regions to the effect of coalitions. Except for the obvious result that larger coalitions perform environmentally better, it is possible to observe that the inclusion of richer countries in potentially stable coalitions is helpful in achieving better environmental results. But also this observation is

Table 3 Sensitivity analysis: variants of α and β values

Cases	α	β	Temperature rise in 2105	Peak temperature rise	Year of peak temperature
1	Het.	0.1	2.460	2.711	2165
2	Het.	0.288	2.442	2.689	2165
3	Het.	0.6	2.415	2.655	2165
4	0.1	0.1	2.557	2.787	2155
5	0.833	0.288	2.402	2.668	2175
6	0.833	0.6	2.374	2.640	2175
7	1.5	0.1	2.365	2.668	2185
8	1.5	0.288	2.349	2.650	2195
9	1.5	0.6	2.323	2.632	2185
10	2.5	0.6	2.263	2.627	2195
11	3.5	0.6	2.255	2.647	2205
12	4.5	0.6	2.258	2.669	2205

rather intuitive, given the positive correlation of wealth and emissions. One interesting thing to note, instead, is the strong sensitivity of temperatures rise to free–riders. Despite the reduction obtained with the PIS stable coalitions of eight players under F&S preferences and heterogeneous α s and βs, temperatures never decline in the whole time span of the simulation. This is due to free–riders that do not implement any serious abatement and, even if they are very limited in number, keep the temperature rise steady in all time periods.

4.4 Sensitivity analysis

In this last section, the sensitivity of the results to variations in the α and β parameters is tested. We focus on the case of full cooperation because, as seen, non-cooperation always results in perpetually increasing temperatures rise far above any acceptable level, whose slight differences will be of scarce interest. The most of the tested variants, reported in Table 3, are based on the assumption of homogeneous α s and βs, with the values of these parameters being taken from the range provided in Fehr and Schmidt (1999): $\alpha \in [0,4.5]$,β∈ (0,0.6),α >β. Exceptions are the three first cases, with case 1 being identical to the scenario reported in Fig. 1, where the value of β is made homogeneous and progressively increased while holding the values of α as in Table 1. In the first three cases, it is possible to observe that an increasing value of β leads to a mild decline in both peak and in the 2105 temperature rise.

The positive effect of increasing levels of altruism can be observed also looking at cases 5 and 6 and then at cases 7, 8 and 9. In both occurrences, for constant levels of α, equal to 0.833 in the former group and to 1.5 in the latter, both peak and temperature rise in 2105 are decreasing in levels of altruism. The same seems to apply to the level of aversion to disadvantageous inequality, apart for one exception. In fact, if we compare cases 6, 9, 10, 11 and 12, where β is always equal to 0.6, temperature rise is inversely correlated with the value of α, except for the last case, namely α =4.5, where temperatures are the highest of the mentioned group. The absolute peak, however, is reached when both the values of α and β are at their minimum, 0.1, with a peak temperature approaching the one obtained in the baseline simulation. Two additional observations are worth to be made. The first is that the lowest level of peak temperature rise is obtained with the maximum value of β (0.6) and the third highest value of α, namely 2.5. We have already mentioned that higher levels of β are conducive of lower peak temperatures and this is so because both more and less economically prosperous countries are incentivized to abate more in order to favour the less advantaged. Only the last country in the ranking, namely Africa, does not have this incentive, but the amount of pollution produced by Africa is rather modest. Increasing levels of α seem to also have a much beneficial impact in lowering temperature rise. Furthermore, this effect seems more pronounced than the one of β, although the range of the two values is different, so that the higher impact of α may be simply due to larger variations in this parameter. However, after a certain threshold, increasing levels of α are deleterious for peak temperature rise. This leads to our second observation. Higher values of α have the effect of increasing the efforts of rich countries/regions, but also the counter–effect of reducing the efforts of poorer nations. Since rich countries generally produce higher amounts of CO_2, the net effect is a reduction in the global level of CO_2 for increasing values of α. After a certain threshold, however, when the possibility of rich countries to further reduce emissions is limited (in Fig. 2, for example, several high income regions start to completely abate since the first time period), the reduction in environmental efforts of poor regions induced by higher levels of α lead to an increase in peak temperatures. The relation between peak temperature reduction and α values may be described by a quadratic function.

Finally, we have done some additional simulations to assess the importance of disadvantageous (α) versus advantageous (β) inequality aversion. In the first simulation, we have set $\alpha = 0$ and $\beta = 0.1$, while we have reversed the values in the second, holding the assumption of homogeneity. In the former, the peak temperature rise, reached in 2155, is of 2.84 ˚C, whereas in the latter it is of 2.80 ˚C, still in 2155. In the other two simulations, we have kept the values of one of the inequality aversion parameters to zero, while holding the values of the other as in Table 1. When α is set to zero, the peak temperature is of 2.75 ˚C, whereas when it is β to be set to zero, the peak is 2.72 ˚C. It seems therefore possible to conclude that disadvantageous inequality aversion has a mildly stronger impact on temperatures rise than advantageous inequality aversion.

4.5 Variations in the value of η

In the analysis conducted so far, we have used as payoff in the F&S utility function, the utility function of the original Nordhaus (2013) model, namely:

$$\pi_{i,t} = \frac{1}{1-\eta}\left(\frac{C_{i,t}}{L_{i,t}}\right)^{1-\eta} + 1, \qquad \forall i \in N, \forall t \in T.$$

The marginal utility of consumption (whose elasticity is η) is used to evaluate per-capita consumption. This should not constitute a theoretical problem since declining marginal utility of consumption is not in contrast with inequality aversion, modeled through the F&S utility function. However, in the economic experiments used to retrieve the values of α and β, players receive pure monetary sums. This implies that the estimated values of α and β, used in this paper with utility payoffs rather than pure monetary payoffs, may be biased. Unfortunately, it does not seem possible to retrieve a conversion factor or a mapping function to duly modify the values of α and β for the present case. This is due to the fact that all the experimental attempts to estimate α and β parameters are based on purely monetary payoffs. Despite it is not sure this causes a significant misrepresentation of α s and βs when used in conjunction with the marginal utility of consumption, we have also tested different levels of η including setting

Table 4 Temperatures rise for different values of η and different scenarios

Scenario	η Value	Peak temperature rise	Year of peak temperature
Baseline	0	2.104	2115
F&S homogeneous		2.172	2145
F&S heterogeneous		2.04	2135
Baseline	0.5	2.461	2125
F&S homogeneous		2.38	2155
F&S heterogeneous		2.31	2145
Baseline	1.5	2.863	2155
F&S homogeneous		2.67	2175
F&S heterogeneous		2.654	2165
Baseline	2.5	3.253	2205
F&S homogeneous		2.97	2205
F&S heterogeneous		2.993	2205

it equal to zero. In this last case, the payoff is equal to per-capita consumption and, therefore, it is a purely monetary amount.

Table 4 shows the peak temperature rise, with the year of achievement, for our three standard scenarios — baseline, F&S with homogeneous and with heterogeneous inequality aversion parameters — for different values of η. The most interesting case is $\eta = 0$, where the utility of the standard RICE model collapses to per-capita consumption. The peak temperature is achieved earlier and it is substantially lower than both the case examined until here ($\eta = 1.5$) and the other cases ($\eta = 0.5$ and $\eta = 2.5$). Actually, the peak temperature is very close to the 2 ˚C threshold in all scenarios. It is interesting to note that, when $\eta = 0$, the baseline scenario has a lower peak temperature rise (2.1 ˚C) than the F&S scenario with homogeneous α s and βs (2.17 ˚C). However, the scenario with F&S preferences and heterogeneous α s and βs is the one leading to the lowest peak temperature, namely, 2.04 ˚C. Therefore, when per-capita consumption substitutes the utility of consumption, besides obtaining significantly lower peak temperatures, we further have a partial confirmation that F&S preferences contribute to increase environmental efforts. The confirmation is only partial since this is not the case for homogeneous α and β parameters.

Table 4 further shows that peak temperatures are positively correlated with the magnitude of η. In particular, for $\eta = 2.5$, in all scenarios, we obtain steadily increasing temperatures throughout the entire model horizon. It is also interesting to note that, starting from the η value of 0.5, the baseline scenario is always the one with the highest peak temperature rise.

5 Conclusions

The economic–environmental models that have been used to investigate the opportunity–costs of limiting the temperature rise caused by greenhouse gases often suggest low levels of CO_2 emissions reduction. A very well-known example is one of the first adopted models to undertake this type of analysis, the DICE/RICE model, whose predicted optimal path leads to a peak temperature increase of 2.86 ˚C over the pre-industrial mean temperature. Several other models reach similar conclusions. The critiques have targeted the computation of climatic damages, judged too mild, the excessive discounting of future payoffs and the lack of consideration for irreversibility and catastrophic risks. The inclusion of these elements generally leads to larger abatement efforts under optimality.

Instead of focusing on one of the mentioned elements, the present paper examines the role of other regarding preferences, assuming that the actors (countries and regions) involved in the negotiation of abatement efforts have Fehr and Schmidt (1999) preferences. Compared to the standard utility functions adopted in integrated assessment models, where the concavity of the utility function described by the elasticity of marginal utility (EMU) of consumption is the only consideration paid to income differences, the F&S function has inequality aversion as its central focus. By adopting the RICE model v2013 and leaving proportional abatement as the only control variable, the paper compares the results obtained, over a time span of 200 years, in three different scenarios: the standard, unaltered, model, the one with F&S preferences and homogeneous parameters of inequality aversion and the one with F&S preferences and heterogeneous α s and βs, adopting for them the mean values retrieved by Dannenberg et al. (2010).

The results show that, when considering F&S utility functions in conjunction with a cooperative behaviour by the players (grand coalition), there is a significant decrease in the peak level of temperature rise compared to the standard RICE model. This is much more pronounced in the case of heterogeneous α s and βs, where the peak temperature rise in 2165 is of 2.65 ˚C. Also when non-cooperation is considered, F&S utility with heterogeneous α s and βs leads to lower levels of peak temperatures. In any case, the effect of F&S preferences in the non-cooperative Nash equilibrium is rather modest.

When the value of the elasticity of marginal utility of consumption (η) is set to zero, all examined scenarios reach a peak temperature rise under full cooperation that is close to the publicly debated 2 ˚C threshold. The lowest temperature peak is still achieved with F&S preferences and heterogeneous inequality aversion parameters (2.04 ˚C), but the baseline scenario leads to a lower peak temperature rise than F&S preferences and homogeneous α s and βs.

In all three scenarios, the grand coalition is not stable and, as evidenced in the environmental game–theoretic literature, cooperation is hard to reach. However, it is important to note that under F&S preferences with heterogeneous α s and βs, the number and size of potentially stable coalitions is higher than in the other cases. However, even considering the largest PIS stable coalitions of eight players obtained in this scenario, the strong effect of free–riding on carbon emissions results in rising temperatures throughout the entire model horizon. By considering modification to the damage function or other corrective mechanisms leading to stricter levels of abatement in conjunction with F&S utility functions, it may still be possible to find self-sustaining coalitions leading to a temperature rise close to the Paris Agreement threshold. This is left for future studies.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s10584-022-03436-6.

Acknowledgements Funding by the German Ministry of Science and Education is gratefully acknowledged.

Author contribution Prof. Carla Vogt: research planning and study design, data analysis, article draft and revision. Dr. Marco Rogna: modeling and numerical simulation, data analysis; article draft and revision.

Funding Open Access funding enabled and organized by Projekt DEAL. Prof. Carla Vogt: research planning and study design, data analysis, article draft and revision. Dr. Marco Rogna: modeling and numerical simulation, data analysis, article draft and revision.

Data availability The model and the underlying data are publicly available on the GitHub repository accessible at this https://github.com/white-heomoi/RICE13_pyomo.

Declarations

Ethics approval Not applicable.

Consent for publication All data used for the present study are in the public domain.

Consent to participate Not applicable.

Competing interests The authors declare no competing interests.

References

Ackerman F, Munitz C (2016) A critique of climate damage modeling: carbon fertilization, adaptation, and the limits of FUND. Energy Res Soc Sci 12:62–67

Ackerman F, Stanton EA, Bueno R (2010) Fat tails, exponents, extreme uncertainty: simulating catastrophe in DICE. Ecol Econ 69(8):1657–1665

Barrett S (1994) Self-enforcing International Environmental Agreements. Oxf Econ Pap 46(1):878–894

Blanco M, Engelmann D, Normann HT (2011) A within-subject analysis of other-regarding preferences. Games Econom Behav 2(72):321–338

Bolton GE, Ockenfels A (2000) ERC: a theory of equity, reciprocity, and competition. Am Econ Rev 90(1):166–193

Bosetti V, Carraro C, Galeotti M, Massetti E, Tavoni M (2006) A world induced technical change hybrid model. Energy J (Special Issue#2)

Botzen WW, van den Bergh JC (2012) How sensitive is Nordhaus to Weitzman? Climate policy in DICE with an alternative damage function. Econ Lett 117(1):372–374

Botzen WW, Van Den Bergh JC, Chichilnisky G (2018) Climate policy without intertemporal dictatorship: Chichilnisky criterion versus classical utilitarianism in DICE. Clim Chang Econ 9(02):1850002

Budolfson MB, Anthoff D, Dennig F, Errickson F, Kuruc K, Spears D, Dubash NK (2021) Utilitarian benchmarks for emissions and pledges promote equity, climate and development. Nat Clim Chang 11(10):827–833

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

Dannenberg A, Sturm B, Vogt C (2010) Do equity preferences matter for climate negotiators? An experimental investigation. Environ Resour Econ 47(1):91–109

d'Aspremont C, Jacquemin A, Gabszewicz JJ, Weymark JA (1983) On the stability of collusive price leadership. Can J Econ 16(1):17–25

De Bruin KC, Dellink RB, Tol RS (2009) AD-DICE: an implementation of adaptation in the DICE model. Clim Chang 95(1):63–81

Dennig F, Budolfson MB, Fleurbaey M, Siebert A, Socolow RH (2015) Inequality, climate impacts on the future poor, and carbon prices. Proc Natl Acad Sci 112(52):15827–15832

Diaz D, Moore F (2017) Quantifying the economic risks of climate change. Nat Clim Chang 7(11):774–782

Dietrich JP, Bodirsky BL, Humpenöder F, Weindl I, Stevanović M, Karstens K, Kreidenweis U, Wang X, Mishra A, Klein D et al (2019) MAGPIE 4–a modular open-source framework for modeling global land systems. Geosci Model Dev 12(4):1299–1317

Dietz S, Asheim GB (2012) Climate policy under sustainable discounted utilitarianism. J Environ Econ Manag 63(3):321–335

Dietz S, Gollier C, Kessler L (2018) The climate beta. J Environ Econ Manag 87:258–274

Easter RC, Ghan SJ, Zhang Y, Saylor RD, Chapman EG, Laulainen NS, Abdul-Razzak H, Leung LR, Bian X, Zaveri RA (2004) MIRAGE: Model description and evaluation of aerosols and trace gases. J Geophys Res Atmos 109(20):1–46

Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. Q J Econ 114(3):817–868
Fehr E, Schmidt KM (2006) The economics of fairness, reciprocity and altruism–experimental evidence and new theories. Handb Econ Giv Altruism Reciproc 1:615–691
Gazzotti P, Emmerling J, Marangoni G, Castelletti A, van der Wijst K-I, Hof A, Tavoni M (2021) Persistent inequality in economically optimal climate policies. Nat Commun 12(1):1–10
Glanemann N, Willner SN, Levermann A (2020) Paris Climate Agreement passes the cost-benefit test. Nat Commun 11(1):1–11
Hänsel MC, Drupp MA, Johansson DJ, Nesje F, Azar C, Freeman MC, Groom B, Sterner T (2020) Climate economics support for the UN climate targets. Nat Clim Chang 10(8):781–789
Hänsel MC, Quaas MF (2018) Intertemporal distribution, sufficiency, and the social cost of carbon. Ecol Econ 146:520–535
Hansson A, Fridahl M, Anshelm J, Haikola S et al (2021) Boundary work and interpretations in the IPCC review process of the role of bioenergy with carbon capture and storage (BECCS) in limiting global warming to 1.5◦C. Front Clim 3:34
Howard PH, Sterner T (2017) Few and not so far between: a meta-analysis of climate damage estimates. Environ Resour Econ 68(1):197–225
Jewell J, Cherp A (2020) On the political feasibility of climate change mitigation pathways: is it too late to keep warming below 1.5◦C? Wiley Interdiscip Rev Clim Chang 11(1):e621
Keramidas K, Kitous A, Després J, Schmitz A (2017) POLES-JRC model documentation. Publ Off Eur Union 10:225347
Knopf B, Kowarsch M, Flachsland C, Edenhofer O (2012) The 2◦C target reconsidered. Climate change, justice and sustainability. Springer, Berlin, pp 121–137
Knutti R, Rogelj J, Sedláček J, Fischer EM (2016) A scientific critique of the two-degree climate change target. Nat Geosci 9(1):13–18
Lange A, Vogt C (2003) Cooperation in international environmental negotiations due to a preference for equity. J Public Econ 87(9):2049–2067
Michaelis P, Wirths H (2020) DICE-RD: an implementation of rate-related damages in the DICE model. Environ Econ Policy Stud 22(4):555–584
Nordhaus W (2007) Accompanying notes and documentation on development of dice-2007 model: notes on DICE-2007. delta. v8 as of september 21 2007. Miscellaneous publication, Yale University, New Haven, NE, USA
Nordhaus W (2018) Projections and uncertainties about climate change in an era of minimal climate policies. Am Econ J Econ Pol 10(3):333–60
Nordhaus W (2018) Evolution of modeling of the economics of global warming: changes in the DICE model, 1992–2017. Clim Chang 148(4):623–640
Nordhaus W, Sztorc P (2013) Dice 2013r: introduction and user's manual. Yale University and the National Bureau of Economic Research, USA
Nordhaus WD (1992) An optimal transition path for controlling greenhouse gases. Science 258(5086):1315–1319
Nordhaus WD (2010) Economic aspects of global warming in a post-Copenhagen environment. Proc Natl Acad Sci 107(26):11721–11726
Nordhaus WD (2013) The climate casino. Yale University Press, New Haven
Nordhaus WD, Boyer J (2000) Warming the world: economic models of global warming. MIT press, Cambridge
Nordhaus WD, Yang Z (1996) A regional dynamic general-equilibrium model of alternative climate-change strategies. Am Econ Rev 86(4):741–765
Ponti G, Rodriguez-Lara I (2015) Social preferences and cognitive reflection: evidence from a dictator game experiment. Front Behav Neurosci 9:146
Rogna M, Vogt C (2020) Coalition formation with optimal transfers when players are heterogeneous and inequality averse. Number 865. Ruhr Economic Papers
Rosen RA (2015) IAMS and peer review. Nat Clim Chang 5(5):390–390
Roughgarden T, Schneider SH (1999) Climate change policy: quantifying uncertainties for damages and optimal carbon taxes. Energy Policy 27(7):415–429
Tol RS (1994) The damage costs of climate change: a note on tangibles and intangibles, applied to DICE. Energy Policy 22(5):436–438
Tol RS (2019) A social cost of carbon for (almost) every country. Energy Econ 83:555–566
van der Pol T, Weikard H-P, van Ierland E (2012) Can altruism stabilise international climate agreements? Ecol Econ 81:112–120
Vogt C (2016) Climate coalition formation when players are heterogeneous and inequality averse. Environ Resour Econ 65(1):33–59

Weitzman ML (2012) GHG Targets as insurance against catastrophic climate damages. J Public Econ Theory 14(2):221–244