FAERE
French Association
of Environmental and Resource Economists

Working papers

Intergenerational equity under catastrophic climate change

Aurélie Méjean - Antonin Pottier - Stéphane Zuber - Marc Fleurbaey

WP 2017.24

Suggested citation:

A. Méjean, A. Pottier, S. Zuber, M. Fleurbaey *(2017).* Intergenerational equity under catastrophic climate change. *FAERE Working Paper, 2017.24.*

ISSN number: **2274-5556**

www.faere.fr

Intergenerational equity under catastrophic climate change

Aurélie Méjean*[1], Antonin Pottier[2], Stéphane Zuber[3] and Marc Fleurbaey[4]

[1]CIRED - CNRS, Centre International de Recherche sur l'Environnement et le Développement (CNRS, Agro ParisTech, Ponts ParisTech, EHESS, CIRAD)

[2]Centre d'Economie de la Sorbonne

[3]Paris School of Economics - CNRS

[4]Woodrow Wilson School of Public and International Affairs, Princeton University

Abstract

Climate change raises the issue of intergenerational equity. As climate change threatens irreversible and dangerous impacts, possibly leading to extinction, the most relevant trade-off may not be between present and future consumption, but between present consumption and the mere existence of future generations. To investigate this trade-off, we build an integrated assessment model that explicitly accounts for the risk of extinction of future generations. We compare different climate policies, which change the probability of catastrophic outcomes yielding an early extinction, within the class of variable population utilitarian social welfare functions. We show that the risk of extinction is the main driver of the preferred policy over climate damages. We analyze the role of inequality aversion and population ethics. Usually a preference for large populations and a low inequality aversion favour the most ambitious climate policy, although there are cases where the effect of inequality aversion is reversed.

Keywords: Climate change; Catastrophic risk; Equity; Population; Climate-economy model

JEL Classification: D63 ; Q01 ; Q54 ; Q56 ; Q5.

*Corresponding author: mejean@centre-cired.fr - CIRED, 45 bis avenue de la Belle Gabrielle, 94736 Nogent-sur-Marne, France

1 Introduction

Since the seminal work of Cline (1992) and Nordhaus (1994), the economics literature has mainly considered climate change as a problem of intertemporal consumption trade-off, where the costs of climate change mitigation measures lower consumption today, but increase consumption in the future as some damages due to climate change are avoided. This approach assumes that climate change occurs at a relatively slow pace - a pace which allows ecosystems and societies to adapt - and has reversible impacts.

The literature on climate science has since then pointed out the possibility of tipping points, where the climate and ecosystems may change in abrupt and irreversible ways (Lenton et al., 2008; Scheffer et al., 2001), possibly bringing catastrophic outcomes. Tipping points may include the shutoff of the Atlantic thermohaline circulation, the collapse of the West Antarctic ice sheet or the dieback of the Amazon rainforest. Abrupt climate change may have indirect impacts, for instance through increased migration and conflicts (Reuveny, 2007; Hsiang et al., 2013).

In the environmental economics litterature, catastrophic outcomes are translated into the irreversible reduction of society's level of consumption or welfare to zero (Cropper, 1976; Clarke and Reed, 1994), or into a discontinuous decline in welfare (Tsur and Zemel, 1996), which is partially reversible[1]. In these models, catastrophic events are triggered by the level of pollution, whether exogenous or endogenous to the model. Integrated assessment models (IAMs) have been used to account for the effect of stochastic climate catastrophes on the optimal climate policy (Peck and Teisberg, 1995; Gjerde et al., 1999; Lontzek et al., 2015; Lemoine and Traeger, 2016). These studies conclude that stochastic tipping points justify more ambitious emission reductions than in the deterministic case. The risk of climate catastrophes is sometimes modelled as a sudden jump in climate damages when the temperature rises above a given temperature threshold (Ambrosi et al., 2003; Pottier et al., 2015), or through a power law damage function with a large exponent (Ackerman et al., 2010; Weitzman, 2012).

In fact, the possibility that social welfare may drop to zero due to climate change can be interpreted as human extinction. The trade-off is then not only between present and future consumption, but also between present consumption and the possible future extinction of civilization due to climate change (Weitzman, 2009). While the economics literature has mainly

[1]The issue of global catastrophic risk has gathered a lot of interest well beyond the issue of climate change (Posner, 2004). In the economics literature, rare disasters (including economic disasters such as The Great Depression, but also natural disasters and epidemics) are also modelled as a drop in consumption (Barro, 2006; Martin, 2008; Barro and Jin, 2011).

focused on the first trade-off, this paper explores the second one. Few papers have approached this issue. Bommier et al. (2015) show that the representation of preferences in terms of risk aversion greatly matters for the appropriate level of mitigation when the risk of catastrophic collapse is accounted for. In their setting, the catastrophe depends on the pollution stock and can be construed as human extinction. Martin and Pindyck (2017) examine the impact of deadly disasters (e.g. disease outbreaks such as the 1918 Spanish flu pandemic), and treat death as a welfare-equivalent reduction in consumption, relying on estimates of the value of a statistical life (VSL).

However, these works do not explicitly address the issue of population ethics, that is the collective attitudes towards population size. The fact that climate change and climate policies could affect the size of the earth's population raises the issue of evaluating policies with varying population size (Broome, 2012). The question of how to value population change, for instance when accounting for the possibility of climate catastrophes, has been largely ignored in the literature and should be treated in an explicit way to inform climate policy analysis (Millner, 2013; Kolstad et al., 2014).

This paper aims at filling this gap, and at identifying public policies that can strike an acceptable compromise between present and future generations when the potential impact of catastrophic climate change on population is accounted for. As we account for the risk of catastrophic climate change, total population over all generations can vary, which raises the issue of the weight given to total population size in the evaluation. We use an IAM, which provides a simple representation of the interaction between climate and the economy, and allows us to evaluate climate policies. We follow Cropper (1976) and assume that the catastrophe is irreversible and is akin to truncating the planning horizon. We depart from the standard optimization framework, and instead consider various climate policies that are ordered according to their performance in terms of welfare. We explore the impact of inequality aversion, of attitudes towards population size and of the risk of extinction on the preferred climate policy.

The paper is structured as follows. Section 2 presents the analytical framework, the model, and the numerical experiment. The analytical results demonstrate that we cannot predict the impact of changes of ethical parameters (inequality aversion and the value of population size) on policy decision. This is because the preferred climate policy depends on the relative impact of these ethical parameters on the welfare gained due to a lower hazard rate and the welfare lost due to a lower consumption stream. In the numerical analysis presented in section 3, we find cases where increasing inequality aversion favours the most ambitious climate policy. This rather unusual result is explained by the relative effect of inequality aversion on the risk and consumption components of the welfare difference. We also find that the risk of extinction is the main driver of the preferred policy over climate damages. Section 4 concludes.

2 Analytical framework

This section describes the analytical framework used in the paper. We introduce the model and social wefare function used (section 2.1), provide tools to analyse policy change in the marginal and general cases (section 2.2), summarize the effects at play (section 2.3), and present the integrated assessment model used in the numerical analysis (section 2.4).

2.1 Model and social welfare functions

To focus on the question of population size, we introduce a model with a risk of extinction. Successive generations are indexed by $t \in \mathbb{N}$. Provided generation t comes into existence, its size is n_t, which is exogenous. We denote $N_t = \sum_{\tau=0}^{t} n_\tau$ the total population up to generation t. This is total population when the world goes extinct just after generation t. At each period t, aggregate consumption $C_t = c_t \cdot n_t$ is deterministic and equally distributed among people of generation t, provided extinction has not yet happened.

In the absence of extinction risk, aggregate welfare U depends on the stream of consumption per capita c and on the number of individuals that will exist. More specifically, we follow most of the existing literature and assume that aggregate welfare has the following (generalized) utilitarian form.[2]

Definition 1 (Variable population utilitarian social welfare functions) *For a finite horizon T, a social welfare function is a variable population utilitarian social welfare function if there exist real numbers $\beta \in [0, 1]$, $\bar{c} \in \mathbb{R}_{++}$ and $\eta \in \mathbb{R}_+$ such that:*

$$U(c) = N_T^{\beta-1} \left\{ \sum_{\tau=0}^{T} n_\tau \left[\frac{c_\tau^{1-\eta}}{1-\eta} - \frac{\bar{c}^{1-\eta}}{1-\eta} \right] \right\}. \quad (1)$$

[2] There exists a limited literature in population ethics proposing alternatives to this generalized utilitarian formula, for instance equally-distributed equivalent criteria (Fleurbaey and Zuber, 2015), egalitarian criteria (Blackorby et al., 1996) or rank-dependent criteria (Asheim and Zuber, 2014, 2016).

The variable population utilitarian social welfare function depends on three important ethical parameters.

1. Parameter η is inequality aversion. It determines the social marginal utility of consumption by each individual in the present and the future. A higher value of η implies that the utility function of individuals is more concave in consumption, so that it brings less marginal utility at high consumption levels relative to marginal utility at low consumption levels. A more inequality averse social welfare function means that we are willing to sacrifice more to equalize consumption levels across individuals.

2. Parameter β determines how valuable larger populations are. Indeed, Eq. (1) embeds well-known approaches to utilitarianism when population size may vary. *Total utilitarianism* is the view that we should consider the total sum of welfare, which happens when $\beta = 1$. *Average utilitarianism*, on the contrary, is the view that we should consider the average welfare, which happens when $\beta = 0$. Values of β between 0 and 1 allow us to span cases between the total and the average views. This is related to a general version of utilitarianism, named "Number-dampened utilitarianism", that was proposed by Ng (1989). Boucekkine et al. (2014) use a similar formulation for within generation population ethics.

3. Parameter $\bar{c}$ is a consumption threshold, which is a level of per capita consumption. When $\beta = 1$ (the standard total utilitarian criterion), the *critical level of consumption* (i.e. the level of consumption such that, if it is enjoyed by an additional individual, total welfare is left unchanged when that individual is added to the population) is equal to $\bar{c}$. This parameter plays a key role in critical-level generalized utilitarianism (Broome, 2004; Blackorby et al., 2005). When $\beta \neq 1$ however, the critical level is not constant and may depend on population size and average welfare in the existing population. Parameter $\bar{c}$ still plays a key role in measuring the value of aggregate welfare and thus the value of changing population size (and the risk on population size).

Now we introduce the risk of extinction: in period t, there is a probability $(1 - p_t)$ of surviving to the next period, and thus a probability p_t of extinction. We name p_t the hazard rate in period t, which can be affected by policy. The planning horizon probability $P_t = p_t \prod_{\tau=0}^{t-1}(1 - p_\tau)$ is the probability that there exists exactly t generations.

In the presence of extinction risk, aggregate welfare W depends both on the stream of consumption per capita c and on the stream of hazard rate p. Indeed, we assume that aggregate welfare

W is the expected value of a variable population utilitarian social welfare function.

Definition 2 (Expected variable population utilitarian social welfare functions) *For a finite horizon T, a social welfare function is an expected variable population utilitarian social welfare function if there exist real numbers $\beta \in [0,1]$, $\bar{c} \in \mathbb{R}_{++}$ and $\eta \in \mathbb{R}_{+}$ such that:*

$$W(c,p) = \mathbb{E}\Big[U(c)\Big] = \sum_{T=0}^{\infty} P_T \left(N_T^{\beta-1} \left\{ \sum_{\tau=0}^{T} n_\tau \left[\frac{c_\tau^{1-\eta}}{1-\eta} - \frac{\bar{c}^{1-\eta}}{1-\eta} \right] \right\} \right). \tag{2}$$

One can write Eq. (2) as follows:

$$W(c,p) = \mathbb{E}\Big[U(c)\Big] \quad = \quad \sum_{\tau=0}^{\infty} \left(\underbrace{\sum_{t=\tau}^{\infty} P_t N_t^{\beta-1}}_{\theta_\tau} \right) n_\tau \left[\frac{c_\tau^{1-\eta}}{1-\eta} - \frac{\bar{c}^{1-\eta}}{1-\eta} \right]. \tag{3}$$

In Eq. (3), the scalar θ_τ is like a discount factor on the wellbeing of generation τ. This discount factor arises only from the uncertainty about the planning horizon: contrary to a standard approach, there is no 'pure' discounting of the utility of future generations. We thus endorse a normative approach that rules out treating generations in an unfair (viz. not symmetric) way (see Ramsey, 1928; Stern, 2007, for arguments in favor of treating generations fairly). Discounting is thus endogenous and depends on: a) the probabilities of extinction; b) the attitudes towards population size as embodied in parameter β. Table 1 summarizes the main notations of our model.

W	welfare
c_τ	consumption per capita at date τ
p_t	hazard rate
N_t	total population up to date t
n_τ	size of generation τ
$\bar{c}$	consumption threshold parameter
η	inequality aversion parameter
β	population ethics parameter
P_t	probability that there exists exactly t generations

Table 1: Main notations

2.2 Evaluating policy change

In this paper, we want to compare different policies proposed in the debate around climate change, taking into account how these policies may change the probability of catastrophic outcomes yielding an early extinction. We also want to better understand the role and impact of ethical parameters of the social welfare function, in particular the inequality aversion parameter η and the population ethics parameter β. We will thus consider policies that affect both the consumption path (climate policy may reduce consumption now and in the short run, but increase consumption in the long run by avoiding climate damages) and the risk profile (climate policy may reduce the probability of extinction in all future periods).

2.2.1 Marginal welfare change

In this section, we introduce additional notation and decompose welfare change in the case of a marginal change in climate policy. Let us denote:

$$AW_T(c) = \sum_{\tau=0}^{T} \frac{n_\tau}{N_T} \left[\frac{c_\tau^{1-\eta}}{1-\eta} - \frac{\bar{c}^{1-\eta}}{1-\eta} \right],$$

the average welfare when there are exactly T generations. Total welfare (Eq. (1)) is thus the product of average welfare and a population weight N_T^β. A first key quantity to analyze climate policy is the well-known social discount rate.

Definition 3 (Social discount rate) *The social discount rate from generation* 0 *to generation* t*, denoted* ρ_t*, is:*

$$\rho_t = \left(\frac{\frac{\partial W}{\partial c_0}}{\frac{\partial W}{\partial c_t}} \right)^{\frac{1}{t}} - 1 = \left(\frac{c_t}{c_0} \right)^{\frac{\eta}{t}} \left(\frac{\sum_{T=0}^{\infty} P_T N_T^{\beta-1}}{\sum_{T=t}^{\infty} P_T N_T^{\beta-1}} \right)^{\frac{1}{t}} - 1. \tag{4}$$

The formula of the social discount rate indicates that increasing η (when $c_t \geq c_0$) increases the discounting of future benefits and may thus reduce the value of the policy. On the other hand, increasing β decreases the social discount rate, because future generations becomes more valuable as they increase total population size (see proof in Appendix A.1).

Let g_t be the growth rate of consumption between 0 and t, hence $c_t = c_0 \cdot (1+g_t)^t$. If we introduce a pseudo pure time preference rate δ_t such that $(1+\delta_t)^t = \theta_0/\theta_t$, with $\theta_t = \sum_{T=t}^{\infty} P_T N_T^{\beta-1}$, we can write Eq. (4) as:

$$1 + \rho_t = (1 + \delta_t)(1 + g_t)^{\eta} \tag{5}$$

Eq. (5) is the Ramsey formula in discrete time. Said differently, introducing a risk of extinction is equivalent to introducing an endogenous pure time preference rate.

Another important quantity describes how much the society is willing a generation to pay to avoid extinction before the next period.

Definition 4 (Social value of catastrophic risk reduction) *The social value of catastrophic risk reduction in period t, denoted ξ_t, is:*

$$\xi_t = -\frac{\frac{\partial W}{\partial p_t}}{\frac{\partial W}{\partial c_t}} = -\frac{\sum_{T=0}^{\infty} \frac{\partial P_T}{\partial p_t} N_T^{\beta} AW_T(c)}{c_t^{-\eta} \sum_{T=t}^{\infty} P_T N_T^{\beta-1}}. \tag{6}$$

The social value of catastrophic reduction ξ was introduced in Bommier et al. (2015), and is related to 'the value of statistical civilization' as introduced by Weitzman (2009). The social value of catastrophic risk extinction bears resemblance to the value of a statistical life that plays a key role in analyzing individual attitudes towards mortality risk. It is formally very close as it measures a risk-consumption trade-off. However, the social value of catastrophic risk extinction has to do with the willingness to add people to a population rather than extending the life of existing individuals. The impact of parameters η and β on ξ is unclear.

With $P_{\geq t} = \prod_{\tau=0}^{t-1}(1 - p_\tau)$ the probability of generation t existing and $P_T^{|t} = \frac{P_T}{P_{\geq t}}$ the probability of the world existing for exactly $T \geq t$ periods, conditional on generation t existing, we have (see Appendix A.2):

$$\xi_t = \frac{\sum_{T=t}^{\infty} P_T^{|t} N_T^{\beta} AW_T(c) - N_t^{\beta} AW_t(c)}{(1 - p_t)(c_t)^{-\eta} \sum_{T=t}^{\infty} P_T^{|t} N_T^{\beta-1}}. \tag{7}$$

The numerator is the expected gain from living longer than for just t generations (conditional on the t first generations existing). The denominator involves the risk of survival at t, the marginal social value of consumption at t and another conditional expectation.

Let us show that the social discount rate and the social value of catastrophic risk reduction are essential to analyze climate policy. Consider a marginal policy that reduces consumption in period 0 by a small amount dc_0 and induces gains in the future in terms of consumption increase (dC_t: reduction of climate damages) and probability reduction of the hazard rate ($-dp_t$). The total welfare gain is:

$$dW = -dc_0 \frac{\partial W}{\partial c_0} + \sum_{T=1}^{\infty} dc_T \frac{\partial W}{\partial c_T} - \sum_{T=1}^{\infty} dp_T \frac{\partial W}{\partial p_T}. \tag{8}$$

This can be written:

$$\begin{aligned} dW &= dc_0 \frac{\partial W}{\partial c_0} \left(-1 + \sum_{T=1}^{\infty} \left(\frac{\frac{\partial W}{\partial c_0}}{\frac{\partial W}{\partial c_T}} \right)^{-1} \frac{dc_T}{dc_0} + \sum_{T=1}^{\infty} \left(\frac{\frac{\partial W}{\partial c_0}}{\frac{\partial W}{\partial c_T}} \right)^{-1} \left(-\frac{\frac{\partial W}{\partial p_T}}{\frac{\partial W}{\partial c_T}} \right) \frac{dp_T}{dc_0} \right) \\ &= dc_0 \frac{\partial W}{\partial c_0} \left(-1 + \sum_{T=1}^{\infty} \frac{1}{(1+\rho_T)^T} \left(\frac{dc_T}{dc_0} + \xi_T \frac{dp_T}{dc_0} \right) \right) \end{aligned}$$

The above expression clearly disentangles both effects of the policy, by separating the impact on consumption and the impact on the risk profile. However, it holds only for marginal policies, which are not the kind of policies we want to consider.

2.2.2 Non-marginal welfare change

We thus develop a more general procedure for non marginal impacts. We want to compare two policies i and j, with paths of per capita consumption c^i and c^j and paths of hazard rates p^i and p^j. The social welfare of each policy is computed from the consumption and hazard rate paths: $W(c^i, p^i)$ and $W(c^j, p^j)$, respectively. We examine the sign of the welfare difference between these policies: $\Delta W = W(c^j, p^j) - W(c^i, p^i)$: when this quantity is positive, policy j is preferred; when it is negative, policy i is preferred. We want to understand how ΔW depends on η and β. In the general case, we can write[3]:

$$\begin{aligned} \Delta W &= W(c^j, p^j) - W(c^i, p^i) \\ &= (W(c^j, p^j) - W(c^j, p^i)) - (W(c^i, p^i) - W(c^j, p^i)) \\ &= \Delta_p W - \Delta_c W \end{aligned} \tag{9}$$

The first term $\Delta_p W = W(c^j, p^j) - W(c^j, p^i)$ is the part of the welfare difference that is explained by the variation of *risk*. We have $\Delta_p W = \sum_t N_t^\beta \cdot (P_t^j - P_t^i) \cdot AW_t^j(c)$. As $P_t = p_t \cdot \prod_{s \le t-1} (1 - p_s)$, the contribution of a lower hazard rate at time t has two opposite effects: there are fewer states of the world with exactly t generations, which reduces total welfare, but there are more states

[3] We could also write $\Delta W = (W(c^j, p^j) - W(c^i, p^j)) - (W(c^i, p^i) - W(c^i, p^j))$ where the first term is the part of the welfare difference that is explained by the variation of *consumption*, while the second term is explained by the variation of *risk*. Using one decomposition or the other gives very similar results.

of the world with strictly more than t generations, which increases total welfare[4].

Reducing the hazard rate comes down to swapping extinction in early periods for extinction in later periods (where the welfare of the state of the world where there are exactly t generations, $N_t^\beta \cdot AW_t(c)$, is higher). If $AW_t(c)$ is above the subsistence level $\overline{c}$ and if $AW_t(c)$ is an increasing sequence, a lower hazard rate at time t increases social welfare. Then, when $p_t^i \geq p_t^j$ for all periods t, $W(c^j, p^j) - W(c^j, p^i)$ is positive.

The second term $\Delta_c W = W(c^i, p^i) - W(c^j, p^i)$ is the part of the welfare difference that is explained by the variation of *consumption*. $\Delta_c W$ reflects a situation where only consumption changes. We have $\Delta_c W = \sum_t N_t^\beta P_t \left(AW_t^i(c) - AW_t^j(c)\right)$. So, when $c_t^i \geq c_t^j$ in all periods t, $W(c^i, p^i) - W(c^j, p^i)$ is positive.

To get a better sense of how parameters β and η may affect welfare change, let us consider the special case where there are no climate damages.

2.2.3 Analytical results when there are no climate damages

Assume that policy j has lower emissions than policy i and, assuming no climate damages, that for each time step we have: $c_t^i \geq c_t^j$ (consumption is always higher in scenario i as less resources are devoted to mitigation), and $p_t^i \geq p_t^j$ (less mitigation in scenario i leads to a higher hazard rate at each time step).

We know that:

$$\Delta_c W = \sum_{t=0}^{\infty} P_t^i N_t^\beta \cdot \left(AW_t(c^i) - AW_t(c^j)\right).$$

Given that $c_t^i \geq c^j t$, $AW_t(c^i) - AW_t(c^j) \geq 0$ for any time horizon t. Hence, $\Delta_c W$ increases in β, as N_t^β increases in β as long as $N_t > 1$.

On the other hand:

$$AW_t(c^i) - AW_t(c^j) = \sum_{\tau=0}^{t} \frac{n_\tau}{N_t} \left[\frac{\left(c^i \tau\right)^{1-\eta}}{1-\eta} - \frac{\left(c_\tau^j\right)^{1-\eta}}{1-\eta} \right].$$

But, when $c_\tau^i \geq c_\tau^j$, $\frac{\left(c_\tau^i\right)^{1-\eta}}{1-\eta} - \frac{\left(c_\tau^j\right)^{1-\eta}}{1-\eta}$ is decreasing in η, provided consumption per capita is larger than $\bar{c}$, which is itself larger than 1.[5]

[4]Provided that these generations are above the subsistence level $\overline{c}$.

[5]Indeed, let a, b two positive real numbers such that $a \geq b$, and let f the function $f(\eta) = \frac{a^{1-\eta} - b^{1-\eta}}{1-\eta}$. We have $f'(\eta) = \frac{(-\ln(a)a^{1-\eta} + \ln(b)b^{1-\eta})(1-\eta) + a^{1-\eta} - b^{1-\eta}}{(1-\eta)^2} = \frac{(1-(1-\eta)\ln(a))a^{1-\eta} - (1-(1-\eta)\ln(b))b^{1-\eta}}{(1-\eta)^2}$. But the function

Similarly, we know that:

$$\Delta_p W = \sum_{t=0}^{\infty} N_t^{\beta} \cdot (P_t^j - P_t^i) \cdot AW_t(c^j).$$

We show in Appendix A.3 that $\Delta_p W$ is increasing with β and decreasing in η. The intuition behind this result follows. As η increases, the concavity of the utility of consumption increases, bringing the utility from the consumption of an individual closer to the utility of the threshold consumption level, $\bar{c}$. Said differently, when inequality aversion increases, the welfare gain of increasing consumption well above the critical level is lower (because the increase in consumption is mainly beneficial at very low income levels), and therefore a population of individuals with sufficiently large consumption levels has less value. Therefore, as η increases, the value in terms of welfare of an additional individual is reduced, i.e. the welfare gained due to a lower risk profile is lower than in the low η case. The following proposition summarizes our results.

Proposition 1 *Assume that policy j and policy i are such that:*

1. *c_t^i and c_t^j are larger than $\bar{c}$, increasing in t and such that $c_t^i \geq c_t^j$ for all t;*
2. *$p_t^i \geq p_t^j$ for all t;*

then both $\Delta_c W$ and $\Delta_p W$ are increasing with β and decreasing with η.

Hence the total impact on ΔW of changes in the ethical parameters β and η is ambiguous, because they have opposite impacts on the cost of a climate policy (reduction of consumption) and the benefit of the policy (better risk profile). We thus need to use numerical examples to see how the parameters may affect policy decision in realistic cases.

2.2.4 Contributions

In the numerical experiment, we propose a method to assess which of the stream of consumption c or the stream of hazard rate p determines the sign of the welfare difference. This method will allow us to check the validity of our results and will prove particularly useful when we account for climate damages, as in this case one consumption path is not necessarily inferior or superior to the other over the whole period and we cannot use the previous analysis to understand the results. The difference in welfare $\Delta W = W(c^j, p^j) - W(c^i, p^i)$ can either be explained by a

$g(x) = (1 - (1-\eta)\ln(x))x^{1-\eta}$ is decreasing in x provided $x > 1$, so that $f'(\eta) < 0$.

difference in consumption, a difference in hazard rate, or both. We examine the product of the variation of welfare ΔW compared to the variation of welfare when only the stream of consumption or the stream of hazard rate changes ($\Delta_c W$ or $\Delta_p W$). If both variations have the same sign, one can attribute the initial welfare variation to consumption or to the hazard rate. Table 2 gives the possible cases as a function of the sign of the quantity at the top of each column[6].

product of welfare differences		*diagnostic*
$\Delta_c W \cdot \Delta W$	$\Delta_p W \cdot \Delta W$	
+	+	Δc_t and Δp_t cause ΔW
+	0	Δc_t causes ΔW, p_t streams play no role
+	-	Δc_t causes ΔW, Δp_t counteracts
-	+	Δp_t causes ΔW, Δc_t counteracts

Table 2: Which of the hazard rate or consumption explains the difference in welfare?

2.3 Summary of the possible effects at play

Several effects may play a role in determining which policy is preferred in the model.

- The first effect is the intertemporal consumption trade-off. As future generations are assumed to be richer (as the model uses an exogenous rate of technical change), a high inequality aversion can give preference to present consumption. This could thus lead to favour no abatement in order to preserve the consumption of the present, poorer generation.

- The second effect is the trade-off between consumption today and the probability of future generations existing. If the risk of extinction depends on the temperature increase compared to pre-industrial levels, climate policy can delay extinction due to climate change. In that case, short-term abatement could be favoured, which would then translate into lower consumption of the present generation, as abatement comes at a cost. Note that the risk on population size is implemented as 'all or nothing', i.e. there is no gradual decrease of population due to climate change. The risk is therefore on cumulative population. Also note that the model accounts for climate damages, but these are certain, i.e. there is no risk on consumption in the model.

[6] We only show one decomposition and the cases that occur in the results presented in section 3.

- The risk of extinction discounts future welfare and thus has an impact on the intertemporal consumption trade-off. Indeed, the contribution of the welfare of future generations can become negligible with a high probability of extinction (whether the probability of extinction is purely exogenous or depends on temperature).

2.4 The model and numerical experiment

2.4.1 The climate-economy model

Our numerical exercise is performed using a climate-economy model named RESPONSE. RESPONSE is a dynamic optimization model (Ambrosi et al., 2003), which belongs to the tradition of compact integrated assessment models such as DICE (Nordhaus, 1994), PAGE (Hope et al., 1993) or FUND (Tol, 1997). Response combines a simple representation of the economy and a climate module. The model can be used to determine the optimal climate objective by comparing mitigation costs and avoided climate damages. The economic module is a Ramsey-like growth model with capital accumulation and population growth. Population growth is considered to be exogenous and welfare is evaluated at the aggregate level at a given period in time. The model includes climate mitigation costs that account for the inertia of technical systems. The climate module describes the evolution of the global temperature and radiative forcing. The model includes a climate damage function. The inter-temporal social welfare is obtained by aggregating individual utilities over time. Key parameters of the model include technical progress[7], climate sensitivity and the functional form of climate damages. A thorough description of the model and its equations can be found in Dumas et al. (2012) and Pottier et al. (2015).

The model is used as a simulation tool to compute consumption and the hazard rate for a given policy. The associated welfare is then computed and the various climate policies are ordered accordingly. We account for the risk of extinction due to climate change by assuming that the hazard rate p depends on the temperature increase compared to pre-industrial levels, noted T. We assume that the hazard rate p is a linear function of temperature increase T, above a temperature threshold T_0. Note that this hazard rate will be referred to as $p_t = p(T_t)$, to be associated to a given date.

[7]In this paper, we assume technical progress decreases over time due to the very long time horizon considered. Indeed, assuming constant labour productivity growth would bring unrealistically high levels of consumption per capita at the time scales considered. This issue is usually not discussed in the literature due to the shorter time horizon used in models (typically a few hundred years).

$$p(T) = \begin{cases} p_0 + b \cdot (T - T_0), & \text{if } T \geq T_0 \\ 1, & \text{if } T \geq T_0 + \frac{1-p_0}{b} \\ p_0, & \text{otherwise} \end{cases} \tag{10}$$

$p(T)$	hazard rate
p_0	minimum hazard rate, set at 10^{-3} per annum
T	temperature increase compared to pre-industrial levels (°C)
T_0	temperature increase above which the risk of extinction starts rising with temperature, set at 1 °C
b	marginal hazard rate per °C above T_0

2.4.2 Parameter ranges and climate policies

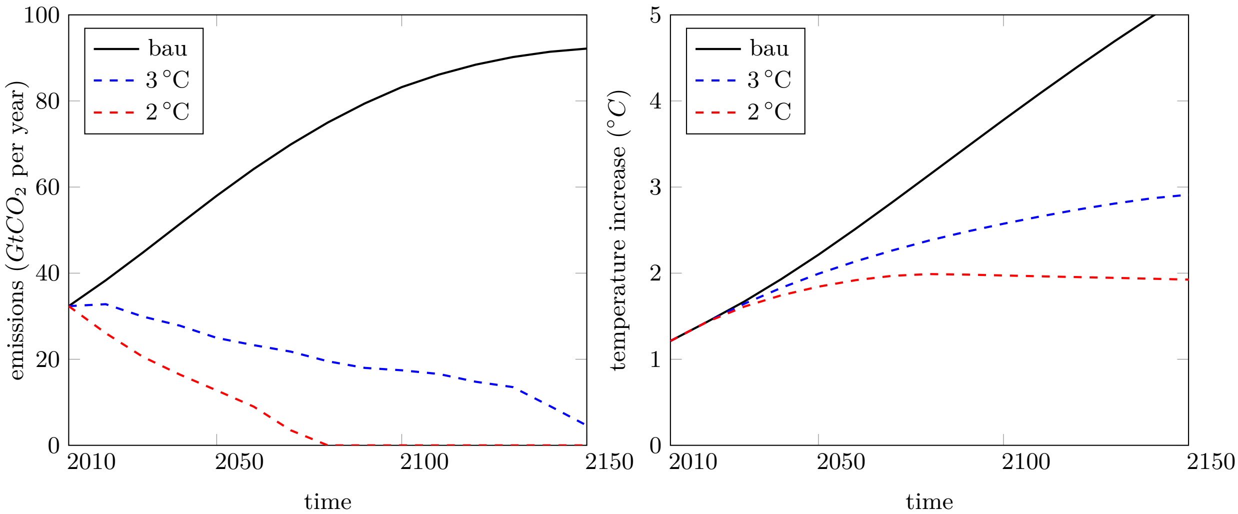
We set p_0 at 10^{-3}, following the Stern Review (2007). Indeed it treats generations in a symetric way and retains a non-zero utility discount rate, set at 10^{-3} to account for the possible extinction of humanity. As there is little, if any, empirical basis to calibrate the minimum hazard rate, we follow this approach. The marginal hazard rate b, i.e. the additional hazard rate per °C above T_0, should be compared with p_0. To grasp what it means, consider that, at the minimum hazard rate $p_0 = 10^{-3}$, the probability of survival after a hundred years is 90%. On top of that, if we assume that the temperature anomaly T is constant at 2 °C (1 °C above the threshold) for a hundred years, the probability of survival after a hundred years would be 89% for $b = 10^{-4}$ per °C, 82% for $b = 10^{-3}$ per °C and 33% for $b = 10^{-2}$ per °C.

We introduce the parameters ranges and climate policies used in the numerical experiment. We choose $0 \leq \beta \leq 1$, which means there is at least a weak preference for large populations. The case $\beta = 0$ corresponds to average utilitarianism, while the case $\beta = 1$ corresponds to total utilitarianism (note that it is then indifferent to add new individuals at the critical level $\bar{c}$). Table 3 shows a summary of the parameter values considered.

parameter	*description*	*value*
η	inequality aversion parameter	from 0.5 to 5.0
β	population parameter	from 0 to 1
b	marginal hazard rate	10^{-7} to 10^{-2} per °C

Table 3: Value ranges of scenario parameters

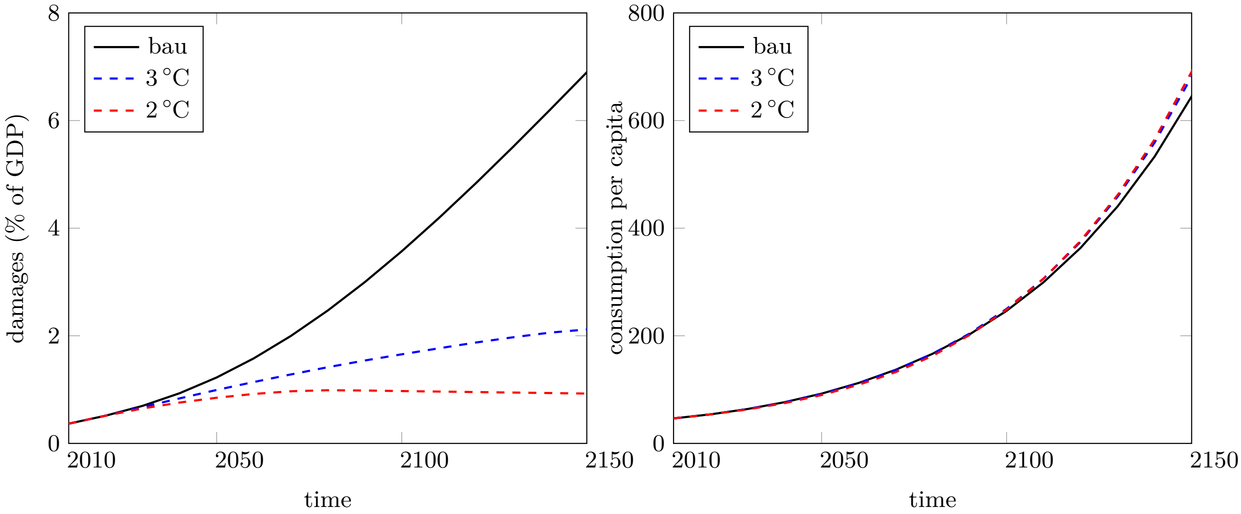
The model is used as a simulation tool to evaluate and order various climate policies. We consider three emission paths: one business-as-usual scenario, i.e. continuing past trends of emissions, calibrated on the middle of the range of business-as-usual scenarios from the database used for the fifth assessment report of the IPCC, see (IIASA, 2014); and two abatement scenarios which are expected to limit the increase of the global temperature above the pre-idustrial level to 3 °C and 2 °C. These policies are defined in terms of emission reduction over time compared to the baseline. The social welfare criteria are used to evaluate pairs of emission paths over time: the results show which policy is preferred according to a given social welfare function, i.e. for given values of the risk and inequality aversion parameters (η), given values of the population parameter (β) and various values of the marginal hazard rate (b). Figure 1 illustrates the policies considered in terms of emissions and temperature. Figure 2 illustrates the associated damages and consumption streams (for a specific set of technical parameters). Figure 3 shows the evolution of the hazard rate p for the three policies considered, for two different values of the marginal hazard rate b.



(a) Emissions

(b) Temperature increase

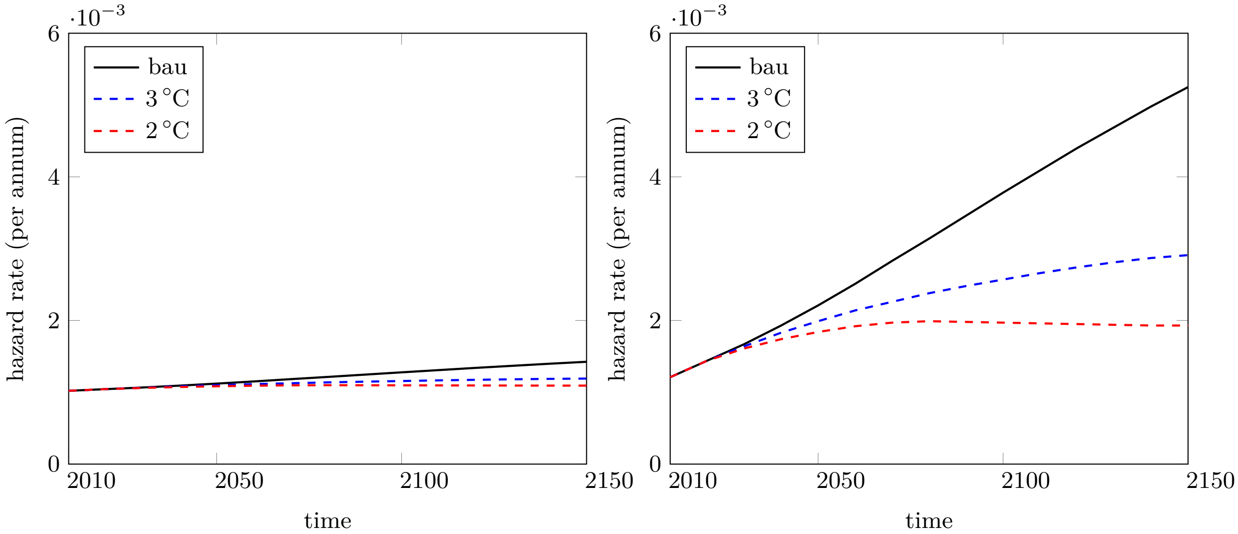
Figure 1: Emissions and temperature increase over time (bau, 3 °C, 2 °C)



(a) Damages

(b) Consumption per capita

Figure 2: Damages and consumption over time (bau, 3 °C, 2 °C)



(a) $b = 10^{-4}$ per °C

(b) $b = 10^{-3}$ per °C

Figure 3: Hazard rate (bau, 3 °C, 2 °C)

2.4.3 The issue of the saving rate

In an optimization framework, integrated assessment models are used to determine the optimal policy in terms of greenhouse gas abatement and savings over time. Abatement and savings are the two decision variables of the model, and are both usually allowed to vary over time. Here, we use the integrated assessment model as a tool to order policies according to their performance given a social welfare function. The policies are thus defined exogenously in terms of both abatement and savings. The difficulty here is that, in principle, each parametrization of the social welfare function admits a different path for optimal savings over time. Two approaches can be adopted here. The first one is a normative approach where one adopts the socially optimal savings rate for each scenario that is examined with a given social welfare function. This means that it may be difficult to compare scenarios along the values of the ethical parameters that define the welfare function, as the underlying saving rate differs. The second approach is positive, as one assumes a certain path of savings over time, independently of the social welfare function considered. This way, scenarios with different ethical parameters are more readily comparable, as they assume the same exogenous evolution of the saving rate. We opt for the positive approach here, as we impose a constant saving rate in all scenarios, like (Golosov et al., 2014). Following Dennig et al. (2015), we choose the value of 25.8%, which is consistent with the observed world average gross saving rate (The World Bank, 2017).

3 Results

We examine the role of the risk of extinction (section 3.1) and the role of ethical parameters (section 3.2) in determining the preferred climate policy.

3.1 The role of the risk of extinction

This section examines the effect of the risk of extinction on the preferred policy between the business-as-usual scenario and a 3 °C scenario. We also stress the importance of the time horizon when computing welfare. In a first approach, we do not account for climate damages, in order to primarily focus on the trade-off between consumption and the risk of extinction. In doing so, we do not consider the intertemporal consumption trade-off determined by the balance between abatement costs and climate damages. Abatement reduces both the stream of consumption and the stream of hazard rate over the whole period[8]. The most ambitious abatement will therefore be favoured because of a reduction in the risk of extinction, while the least ambitious policy will be favoured because of an increase in consumption at early periods. We later examine the impact of climate change damages on the preferred policy.

3.1.1 Consumption vs. risk of extinction

We examine the preferred policy option between BAU and 3 °C as a function of the probability of extinction for a given degree of inequality aversion ($\eta = 2.0$) and for a given value of the weight on population size ($\beta = 1.0$, i.e. the case of total utilitarianism). The results presented in table 4 show that only a zero marginal hazard rate b (i.e. the case of a purely exogenous hazard rate) leads to favour the BAU scenario (i.e. no abatement)[9]. With a marginal hazard rate equal to zero, mitigation has no impact on extinction and only reduces consumption. So climate policy necessarily reduces welfare given that we exclude climate damages for now. With a marginal hazard rate superior to zero, there is a chance that climate action may avoid extinction, and the 3 °C scenario is favoured over the BAU. Further tests show that for a given set of ethical parameters (η and β), the preferred policy is influenced by the relative order of magnitude of the exogenous hazard rate (p_0) and of the marginal hazard rate (b).

b (per $^\circ C$)						
0	10^{-7}	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}
BAU	3°C	3°C	3°C	3°C	3°C	3°C

(light)	Δc_t causes ΔW, Δp_t plays no role
(grey)	Δp_t causes ΔW, Δc_t counteracts

Note that the preferred policy is written inside each cell, while the color of the cell indicates what determines the welfare difference (see section 2.2.4)

Table 4: BAU vs. 3 °C ($\beta = 1.0$, $\eta = 2.0$)

The results show that the difference in the hazard rate always explains the preference for the 3 °C scenario, while the difference in consumption counteracts (in grey). Conversely, the difference

[8] The stream of consumption is lower over the whole period due to the fact that we do not account for climate damages. This is not be the case if climate damages were accounted for, cf. Figure 2b.

[9] Further tests show that for a very high marginal hazard rate (above 0.5 per °C, i.e. if future generations are unlikely to exist) it is not worth abating emissions today. The result that a rational social planner may voluntarily choose not to abate emissions in a case where it appears too difficult to avoid catastrophic climate damages has been shown in (Perrissin-Fabert et al., 2014). A high marginal hazard rate (b) at low temperatures favours the BAU scenario, as the state of the climate is then considered hopeless, and one might as well favour present consumption if future generations are unlikely to exist.

in consumption explains the preferrence for the BAU scenario when the hazard rate is purely exogenous (i.e. for a marginal hazard rate $b = 0.0$ per °C), as hazard rate streams are identical in the 3 °C scenario and the BAU scenario. This means that when the 3 °C scenario is preferred to the BAU, this is due to its effect on reducing the hazard rate, while when the BAU scenario is preferred to the 3 °C scenario, this is due to its effect on consumption streams.

3.1.2 Welfare and the role of the time horizon

The welfare is calculated by adding the successive contributions of various generations to welfare. In other words, the welfare is calculated by adding welfare over states of the world, not over time periods. In practice, welfare is calculated by truncating Eq. (2) at a chosen time horizon. We find that the choice of the time horizon of the model is crucial to correctly interpret the results. Indeed, as the hazard rate p depends on the emissions path, the time horizon should be chosen so that the probability of survival at the end of the period is close to zero for all the emission paths considered, in order to ensure that long term welfare is not overlooked when calculating the aggregated welfare (used to determine which policy should be preferred). If the time horizon is too short, i.e. if the probability of survival at the end of the time horizon is still significant, the long term benefits of a given policy are cut out of the assessment, which for instance may lead to the wrong conclusion that a BAU scenario should be preferred to a 3 °C scenario. With a minimum hazard rate p_0 set at 10^{-3} per annum, the minimum time horizon is $10,000$ years when the most ambitious climate policy considered is a 3 °C scenario (over $10,000$ years the probability of survival is of the order of 4.10^{-5}).

3.1.3 The role of damages

So far, the analysis has not accounted for the intertemporal consumption trade-off between abatement costs and climate damages. We now test the sensitivity of the results to the inclusion of climate damages in the model. Climate damages occur due to temperature increase, and are subtracted from production, thus reducing consumption. As before, they can be mitigated thanks to abatement, which comes at a cost. The results show that the preferred policies are unchanged whether or not climate damages are accounted for, with the exception of $b = 0$ per °C , i.e. for a purely exogenous hazard rate. This result means that the risk of extinction due to climate change is the main driver of the policy choice over the effect of climate damages on consumption.

	b (per $^\circ C$)						
	0	10^{-7}	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}
w/o damages	BAU	3°C	3°C	3°C	3°C	3°C	3°C
w damages	3°C	3°C	3°C	3°C	3°C	3°C	3°C

Δc_t causes ΔW, Δp_t counteracts or plays no role

Δp_t causes ΔW, Δc_t counteracts or plays no role

Δc_t and Δp_t cause ΔW

Table 5: BAU vs. 3 °C with and without damages ($\beta = 1.0$, $\eta = 2.0$)

The table shows that the 3 °C scenario is preferred due to both the risk and the consumption effects (again with the exception of $b = 0$ per °C). This contrasts with the case without climate damages, where the 3 °C scenario was preferred due to the risk effect alone, while the consumption effect played in favour of the BAU. The effect on consumption is indeed expected to favour the 3 °C scenario, as the reduced consumption of future generations due to climate damages can impact total welfare in a significant way. The results also show that when damages are accounted for and when there is no risk of extinction ($b = 0$ per °C), the 3 °C scenario is preferred to the BAU scenario. This result is in accordance with the results of the Stern Review (2007), which, assuming a pure time preference of 10^{-3}, concludes that the avoided damages of a 3 °C scenario outweigh its abatement costs, and that such a policy should be pursued. It is interesting to note that, with damages, Δc_t no longer causes ΔW when b increases from 10^{-3} to 10^{-2} per °C. Indeed, the 3 °C scenario is then preferred due to the risk of extinction only, as the intertemporal consumption trade-off no longer plays in favour of the 3 °C scenario for higher values of b. This is due to the fact that with a higher marginal hazard rate, the benefits of abatement on future consumption through avoided climate damages have less weight in total welfare, as future generations are less likely to exist. This is an example of how the hazard rate and climate damages interact in the intertemporal consumption trade-off.

3.2 The role of ethical parameters

We now examine the role of ethical parameters in determining the preferred policy in a case with climate damages in addition to the catastrophic risk of extinction due to climate change. We consider in turn the role of the weight on population size (β), and the role of the inequality aversion (η).

3.2.1 The role of population ethics

In table 6, we examine preferred policy options[10] (between the BAU, 3 °C and 2 °C scenarios) as a function of the marginal hazard rate b for various weights on population size (various β) and for given values of inequality aversion ($\eta = 1.5$ and $\eta = 2.5$). The contributions of the difference in hazard rate and of the difference in consumption to the difference in welfare is calculated when comparing the preferred policy and the next best policy (see section 2.2.4).

For a given marginal hazard rate b, a higher weight on population size favours the most ambitious climate policy. This result is consistent with intuition, as ambitious climate policies allow for larger cumulative population. Above a certain value of the marginal hazard rate, the value of β plays no role on the preferred policy. We show that the difference in hazard rate always explains the preference for the most ambitious climate scenario, while the difference in consumption counteracts. Conversely, the difference in consumption always explains the preference for the least ambitious climate scenario while the difference in hazard rate counteracts, except when the hazard rate is purely exogenous (i.e. for a marginal hazard rate $b = 0$).

[10]See Appendix B.2 for the comparison between the 3 °C and BAU scenarios, and between the 3 °C and 2 °C scenarios.

β = 0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	b (per $^\circ C$)
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[10^{-5}; 10^{-2}]$
3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[4.10^{-6}; 9.10^{-6}]$
3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[2.10^{-6}; 3.10^{-6}]$
3°C	3°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-6}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	2°C	2°C	10^{-7}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

(a) $\eta = 1.5$

β = 0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	b (per $^\circ C$)
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[10^{-4}; 10^{-2}]$
3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[3.10^{-6}; 10^{-5}]$
3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[10^{-6}; 2.10^{-6}]$
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	2°C	2°C	10^{-7}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

(b) $\eta = 2.5$

Light grey: Δc_t causes ΔW, Δp_t counteracts (or plays no role)

Dark grey: Δp_t causes ΔW, Δc_t counteracts (or plays no role)

Table 6: BAU vs. 3 °C vs. 2 °C as a function of β and b (with climate damages)

3.2.2 The role of inequality aversion

We examine preferred policy options as a function of the marginal hazard rate b for various degrees of inequality aversion (η) and for a given of the weight on population size ($\beta = 0, 0.1, 1$). We examine the preferred policy between the BAU, 3 °C and 2 °C scenarios in table 7 (see Appendix C.2 for the comparison between the 3 °C and BAU scenarios, and between the 3 °C and 2 °C scenarios).

In the case of average utilitarianism ($\beta = 0$, table 7a), at a given b, a higher inequality aversion favours the least ambitious climate policy. As the marginal hazard rate b decreases, the minimum level of inequality aversion that justifies the least ambitious scenario is reduced: richer generations are added, which enhances inequalities between generations. At high values of the marginal hazard rate b, the most ambitious climate scenario is favoured for all values of inequality aversion.

As climate damages are accounted for, the 2 °C scenario is preferred due to both the differences in hazard rate and consumption streams for relatively low values of inequality aversion ($\eta \leq 1.0$) in the case of average utilitariansim (i.e. $\beta = 0$), see table 7a. This contrasts with the case without climate damages, where the most ambitious scenario is preferred due to the difference in hazard rate alone, while the consumption effect played in favour of the BAU (cf. Appendix C.1). The difference in consumption streams is indeed expected to favour the most ambitious climate policy for low inequality aversion, as climate damages reduce the consumption of future generations more than that of present ones. As future generations are richer than present ones in the baseline scenario due to technical change, this effect only occurs for low inequality aversion.

Note that, as before, the contributions of the difference in hazard rate and of the difference in consumption to the difference in welfare is calculated when comparing the preferred policy and the next best policy. When the 3 °C policy is preferred, the next best policy is either the BAU or the 2 °C. When the 3 °C scenario is preferred (see for instance table 7a), when the next best policy is the 2 °C, the corresponding cell is white (i.e. the 3 °C scenario is preferred to the 2 °C due to the difference in consumption, while the difference in hazard rate counteracts); when the next best policy is the BAU, the conrresponding cell is grey (i.e. the 3 °C scenario is preferred to the BAU due to the difference in hazard rate, while the difference in consumption counteracts).

The pattern is different for $\beta = 0.1$ (table 7b). Increasing the inequality aversion η still favours the least ambitious climate policy for low values of η, but the effect is reversed for $\eta > 2.0$. As shown in Prop. 1, there are two opposite effects of increasing inequality aversion: it decreases the value of reducing the extinction but it also decreases the cost of implementing the policy by making the first (poor) generations pay. It turns out that the decrease in the value of risk reduction is first faster and then slower that the decrease in the cost of the policy (see Appendix C.3).

The choice of β significantly changes the preferred policy option for low values of the marginal hazard rate (b), but plays no role for higher values of b (top lines of the tables). In the case of average utilitarianism ($\beta = 0.0$), inequality aversion plays a role on the preferred policy option for a wider range of b values than in the case of total utilitarianism ($\beta = 1.0$). The choice of β does not influence the preferred policy when the hazard rate is exogenous.

η										b (per $^\circ C$)
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-2}
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-3}
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-4}
2°C	2°C	2°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-5}
2°C	2°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[8.10^{-6}; 9.10^{-6}]$
2°C	2°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[4.10^{-6}; 7.10^{-6}]$
2°C	2°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3.10^{-6}
2°C	2°C	3°C	3°C	3°C	3°C	3°C	BAU	BAU	BAU	2.10^{-6}
2°C	2°C	3°C	3°C	3°C	BAU	BAU	BAU	BAU	BAU	10^{-6}
2°C	2°C	3°C	3°C	3°C	BAU	BAU	BAU	BAU	BAU	10^{-7}
2°C	2°C	3°C	3°C	3°C	BAU	BAU	BAU	BAU	BAU	0

(a) $\beta = 0$

η										b (per $^\circ C$)
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-2}
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[4.10^{-6}; 10^{-3}]$
2°C	2°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	3.10^{-6}
2°C	2°C	3°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	2.10^{-6}
2°C	2°C	3°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	10^{-6}
2°C	2°C	3°C	3°C	3°C	3 °C	3 °C	3°C	2°C	2°C	10^{-7}
2°C	2°C	3°C	3°C	3°C	BAU	BAU	BAU	BAU	BAU	0

(b) $\beta = 0.1$

η										b (per $^\circ C$)
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-2}
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[10^{-6}; 10^{-3}]$
2°C	2°C	3°C	3°C	2°C	2°C	3°C	2°C	2°C	2°C	10^{-7}
2°C	2°C	3°C	3°C	3°C	BAU	BAU	BAU	BAU	BAU	0

(c) $\beta = 1.0$

- (light) Δc_t causes ΔW, Δp_t counteracts (or plays no role)
- (grey) Δp_t causes ΔW, Δc_t counteracts (or plays no role)
- (black) Δp_t and Δc_t cause ΔW

Table 7: BAU vs. 3 °C vs. 2 °C as a function of η and b (with climate damages)

4 Conclusion

With a probability of extinction that depends on temperature increase compared to pre-industrial levels, two effects are competing. On the one hand, future generations are assumed to be richer, and a high inequality aversion thus gives preference to present consumption. This plays in favour of the least ambitious climate policy as a way to preserve the consumption of the present, poorer generation. On the other hand, emission reductions can prevent extinction, which can favour ambitious climate policies. The main results are summarised below.

- We cannot predict the impact of changes in the ethical parameters β (value of population size) and η (inequality aversion) on the preferred policy, even in the case without climate damages. The preferred policy depends on the relative effect of η and β on the welfare gained due to a lower hazard rate and the welfare lost due to a lower consumption stream. A high η tends to lower the welfare gained due to higher consumption stream. It also reduces the value of postponing extinction. A high β increases both the welfare lost due to a lower consumption stream, and the welfare gained as the size of the cumulative population increases due to a lower hazard rate.
- In the numerical analysis, a large β always favours the most ambitious climate scenario. This result is consistent with our first intuition: a higher β gives as a higher weight to the welfare of future generations. Low inequality aversion (η) also usually favours the most ambitious scenario, as future generations are assumed to be richer. This is consistent with results from the literature. More precisely, the results confirm that total utilitarians who are little inequality averse would tend to favour ambitious climate policies.
- However, we find cases where increasing inequality aversion favours the most ambitious climate policy. For instance, the 2°C policy is preferred to the 3°C policy as η increases for the combination of parameters ($\eta \geq 1.5$; $\beta = 0.1$, $b \leq 5.10^{-4}$). This pattern is also observed for higher values of β. This result is explained by the relative effect of inequality aversion on the risk and consumption components of the welfare difference. This analysis may thus help identifying new spaces of compromise between various ethical stances to set the ambition of climate policies. Indeed, a coalition could be formed between opposite sides of the ethical spectrum regarding inequality aversion.
- Finally, except for very low hazard rates ($b \leq 3.10^{-4}$, i.e. a probability of survival in

2100 of 97% if the temperature increase compared to the industrial level reaches 2°C), the risk of extinction is the main driver of the preferred policy over climate damages, as the preferred policy is unchanged whether or not climate damages are accounted for.

The effect of ethical parameters on the preferred policy should be further explored. Indeed, the values of the ethical parameters may play a significant role on the optimal climate policy, which was not the focus of this exercise. The next step would thus be to examine the influence of inequality aversion and of the weight given to population size on the optimal policy. A further step would be to consider more general welfare functions that disentangle risk aversion and inequality aversion, to explore whether one of these parameters has more influence on the results. Finally, it would be interesting to include the possibility of partial extinction, i.e. to allow for a variable population size as a function of the severity of climate change instead of the 'all-or-nothing' set-up used in this paper.

Acknowledgements

This research has been supported by the Chair on Welfare Economics and Social Justice at the Institute for Global Studies (FMSH - Paris), the Franco-Swedish Program on Economics and Philosophy (FMSH - Risksbankens Jubileumsfond), the ANR project EquiRisk, Grant # ANR-12-INEG-0006-01, the ANR project FairClimpop, Grant # ANR-16-CE03-0001-01 and Princeton University's Program in Science, Technology and Environmental Policy (STEP). We thank an anonymous referee from the French Association of Environmental and Resource Economists (FAERE).

References

Ackerman, F., Stanton, E. A., and Bueno, R. 2010. Fat tails, exponents, extreme uncertainty: Simulating catastrophe in DICE. *Ecological Economics* 69:1657–1665.

Ambrosi, P., Hourcade, J.-C., Hallegatte, S., Lecocq, F., Dumas, P., and Duong, M. H. 2003. Optimal Control Models and Elicitation of Attitudes towards Climate Damages. *Environmental Modeling & Assessment* 8:133–147.

Asheim, G. and Zuber, S. 2014. Escaping the repugnant conclusion: Rank-discounted utilitarianism with variable population. *Theoretical Economics* 9:629–650.

Asheim, G. and Zuber, S. 2016. Evaluating intergenerational risks. *Journal of Mathematical Economics* 65:73–96.

Barro, R. J. 2006. Rare Disasters and Asset Markets in the Twentieth Century*. *Quarterly Journal of Economics* 121:823–866.

Barro, R. J. and Jin, T. 2011. On the Size Distribution of Macroeconomic Disasters. *Econometrica* 79:1567–1589.

Blackorby, C., Bossert, W., and Donaldson, D. 1996. Leximin population ethics. *Mathematical Social Sciences* 31:115–131.

Blackorby, C., Bossert, W., and Donaldson, D. 2005. Population Issues in Social Choice Theory, Welfare Economics, and Ethics. Cambridge University Press.

Bommier, A., Lanz, B., and Zuber, S. 2015. Models-as-usual for unusual risks? On the value of catastrophic climate change. *Journal of Environmental Economics and Management* 74:1–22.

Boucekkine, R., Fabbri, G., and Gozzi, F. 2014. Egalitarianism under population size: age-structure does matter. *Journal of Mathematical Economics* 55:86–100.

Broome, J. 2004. Weighing Lives. Oxford University Press.

Broome, J. 2012. Climate matters: ethics in a warming world. Amnesty International global ethics series. W.W. Norton, New York, 1st ed edition.

Clarke, H. R. and Reed, W. J. 1994. Consumption/pollution tradeoffs in an environment vulnerable to pollution-related catastrophic collapse. *Journal of Economic Dynamics and Control* 18:991–1010.

Cline, W. R. 1992. The economics of Global Warming. Institute for International Economics, Washington.

Cropper, M. 1976. Regulating activities with catastrophic environmental effects. *Journal of Environmental Economics and Management* 3:1–15.

Dennig, F., Budolfson, M. B., Fleurbaey, M., Siebert, A., and Socolow, R. H. 2015. Inequality, climate impacts on the future poor, and carbon prices. *Proceedings of the National Academy of Sciences* 112:15827–15832.

Dumas, P., Espagne, E., Perrissin-Fabert, B., and Pottier, A. 2012. Comprehensive Description of RESPONSE. CIRED Working Paper 41-2012, CIRED.

Fleurbaey, M. and Zuber, S. 2015. Discounting, beyond Utilitarianism. *Economics: The Open-Access, Open-Assessment E-Journal* 9:1–52.

Gjerde, J., Grepperud, S., and Kverndokk, S. 1999. Optimal climate policy under the possibility of a catastrophe. *Resource and Energy Economics* 21:289–317.

Golosov, M., Hassler, J., Krusell, P., and Tsyvinski, A. 2014. Optimal Taxes on Fossil Fuel in General Equilibrium. *Econometrica* 82:41–88.

Hope, C., Anderson, J., and Wenman, P. 1993. Policy analysis of the greenhouse effect - An application of the PAGE model. *Energy Policy* 21:327–338.

Hsiang, S. M., Burke, M., and Miguel, E. 2013. Quantifying the Influence of Climate on Human Conflict. *Science* 341:1235367–1235367.

IIASA 2014. AR5 Scenario Database.

Kolstad, C., Urama, K., Broome, J., Bruvoll, A., Cariño Olvera, M., Fullerton, D., Gollier, C., Hanemann, W., Hassan, R., Jotzo, F., Khan, M., Meyer, L., and Mundaca, L. 2014. Social, Economic, and Ethical Concepts and Methods. *In* O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. v. Stechow, T. Zwickel, and J. Minx (eds.), Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Lemoine, D. and Traeger, C. P. 2016. Economics of tipping the climate dominoes. *Nature Climate Change* 6:514–519.

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., and Schellnhuber, H. J. 2008. Tipping elements in the Earth's climate system. *Proceedings of the National Academy of Sciences* 105:1786–1793.

Lontzek, T. S., Cai, Y., Judd, K. L., and Lenton, T. M. 2015. Stochastic integrated assessment of climate tipping points indicates the need for strict climate policy. *Nature Climate Change* 5:441–444.

Martin, I. and Pindyck, R. 2017. Averting Catastrophes that Kill. Technical Report w23346, National Bureau of Economic Research, Cambridge, MA. DOI: 10.3386/w23346.

Martin, I. W. R. 2008. Disasters and the Welfare Cost of Uncertainty. *American Economic Review* 98:74–78.

Millner, A. 2013. On welfare frameworks and catastrophic climate risks. *Journal of Environmental Economics and Management* 65:310–325.

Ng, Y.-K. 1989. What should we do about future generations? impossibility of parfit's theory x. *Economics and Philosophy* 5:235–253.

Nordhaus, W. D. 1994. Managing the global commons: the economics of climate change. MIT Press, Cambridge, Mass.

PECK, S. C. AND TEISBERG, T. J. 1995. Optimal CO2 control policy with stochastic losses from temperature rise. *Climatic Change* 31:19–34.

PERRISSIN-FABERT, B., POTTIER, A., ESPAGNE, E., DUMAS, P., AND NADAUD, F. 2014. Why are climate policies of the present decade so crucial for keeping the 2 °C target credible? *Climatic Change* 126:337–349.

POSNER, R. A. 2004. Catastrophe: risk and response. Oxford Univ. Press, Oxford. OCLC: 611580585.

POTTIER, A., ESPAGNE, E., FABERT, B. P., AND DUMAS, P. 2015. The Comparative Impact of Integrated Assessment Models' Structures on Optimal Mitigation Policies. *Environmental Modeling & Assessment* 20:453–473.

RAMSEY, F. 1928. A mathematical theory of saving. *The Economic Journal* 38:543–559.

REUVENY, R. 2007. Climate change-induced migration and violent conflict. *Political Geography* 26:656–673.

SCHEFFER, M., CARPENTER, S., FOLEY, J. A., FOLKE, C., AND WALKER, B. 2001. Catastrophic shifts in ecosystems. *Nature* 413:591–596.

STERN, N. H. 2007. The economics of climate change: the Stern review. Cambridge University Press, Cambridge, UK ; New York.

THE WORLD BANK 2017. Gross savings (% of GDP). World Bank national accounts data, and OECD National Accounts data files., The World Bank.

TOL, R. S. J. 1997. On the optimal control of carbon dioxide emissions: an application of FUND. *Environmental Modeling & Assessment* 2:151–163.

TSUR, Y. AND ZEMEL, A. 1996. Accounting for global warming risks: Resource management under event uncertainty. *Journal of Economic Dynamics and Control* 20:1289–1305.

WEITZMAN, M. L. 2009. On Modeling and Interpreting the Economics of Catastrophic Climate Change. *Review of Economics and Statistics* 91:1–19.

WEITZMAN, M. L. 2012. GHG Targets as Insurance Against Catastrophic Climate Damages. *Journal of Public Economic Theory* 14:221–244.

Appendices

A Appendix - Proofs

A.1 The social discount rate

The social discount rate is

$$\rho_t = \left(\frac{c_t}{c_0}\right)^{\frac{\eta}{t}} \left(\frac{\sum_{T=0}^{\infty} P_T N_T^{\beta-1}}{\sum_{T=t}^{\infty} P_T N_T^{\beta-1}}\right)^{\frac{1}{t}} - 1.$$

Let us show that $h(\beta) = \frac{\sum_{T=0}^{\infty} P_T N_T^{\beta-1}}{\sum_{T=t}^{\infty} P_T N_T^{\beta-1}}$ is decreasing in β. Indeed,

$$\begin{aligned}
h'(\beta) &= \frac{\left(\sum_{T=0}^{\infty} P_T \ln(N_T) N_T^{\beta-1}\right)\left(\sum_{T=t}^{\infty} P_T N_T^{\beta-1}\right) - \left(\sum_{T=0}^{\infty} P_T N_T^{\beta-1}\right)\left(\sum_{T=t}^{\infty} P_T \ln(N_T) N_T^{\beta-1}\right)}{\left(\sum_{T=t}^{\infty} P_T N_T^{\beta-1}\right)^2} \\
&= \frac{\left(\sum_{T=0}^{t-1} P_T \ln(N_T) N_T^{\beta-1}\right)\left(\sum_{T=t}^{\infty} P_T N_T^{\beta-1}\right) - \left(\sum_{T=0}^{t-1} P_T N_T^{\beta-1}\right)\left(\sum_{T=t}^{\infty} P_T \ln(N_T) N_T^{\beta-1}\right)}{\left(\sum_{T=t}^{\infty} P_T N_T^{\beta-1}\right)^2}
\end{aligned}$$

But given that N_T is the total population of the T first generations (hence increasing in T), we have $N_T \leq N_{t-1}$ for all $T \leq t-1$, $N_T \geq N_t$ for all $T \geq t$ and $N_t > N_{t-1}$. Hence:

$$\begin{aligned}
h'(\beta) &= \frac{\left(\sum_{T=0}^{t-1} P_T \ln(N_T) N_T^{\beta-1}\right)\left(\sum_{T=t}^{\infty} P_T N_T^{\beta-1}\right) - \left(\sum_{T=0}^{t-1} P_T N_T^{\beta-1}\right)\left(\sum_{T=t}^{\infty} P_T \ln(N_T) N_T^{\beta-1}\right)}{\left(\sum_{T=t}^{\infty} P_T N_T^{\beta-1}\right)^2} \\
&\leq \frac{\ln(N_{t-1})\left(\sum_{T=0}^{t-1} P_T N_T^{\beta-1}\right)\left(\sum_{T=t}^{\infty} P_T N_T^{\beta-1}\right) - \ln(N_t)\left(\sum_{T=0}^{t-1} P_T N_T^{\beta-1}\right)\left(\sum_{T=t}^{\infty} P_T N_T^{\beta-1}\right)}{\left(\sum_{T=t}^{\infty} P_T N_T^{\beta-1}\right)^2} \\
&= \frac{\left(\ln(N_{t-1}) - \ln(N_t)\right)\left(\sum_{T=0}^{t-1} P_T N_T^{\beta-1}\right)\left(\sum_{T=t}^{\infty} P_T N_T^{\beta-1}\right)}{\left(\sum_{T=t}^{\infty} P_T N_T^{\beta-1}\right)^2} < 0.
\end{aligned}$$

Hence, $h'(\beta) < 0$ and the social discount rate is decreasing in β.

A.2 The social value of catastrophic risk reduction

The social value of catastrophic risk reduction is

$$\xi_t = -\frac{\sum_{T=0}^{\infty} \frac{\partial P_T}{\partial p_t} N_T^{\beta} AW_T(c)}{c_t^{-\eta} \sum_{T=t}^{\infty} P_T N_T^{\beta-1}},$$

with

$$AW_T(c) = \sum_{\tau=0}^{T} \frac{n_\tau}{N_T} \left[\frac{c_\tau^{1-\eta}}{1-\eta} - \frac{\bar{c}^{1-\eta}}{1-\eta} \right].$$

Here we explain how to derive Eq. (7). Note that $\frac{\partial P_T}{\partial p_t} = 0$ if $T < t$, $\frac{\partial P_t}{\partial p_t} = P_t/p_t$ and $\frac{\partial P_T}{\partial p_t} = -P_T/(1-p_t)$ if $T > t$. Let $P_{\geq t} = \prod_{\tau=0}^{t-1}(1-p_\tau)$ the probability of generation t existing and $P_T^{|t} = \frac{P_T}{P_{\geq t}}$ the probability of the world existing for exactly $T \geq t$ periods, conditional on generation t existing. Hence,

$$\begin{aligned}
\sum_{T=0}^{\infty} \frac{\partial P_T}{\partial p_t} N_T^{\beta} AW_T(c) &= P_{\geq t} N_t^{\beta} AW_t(c) - \sum_{T=t+1}^{\infty} \frac{P_T}{(1-p_t)} N_T^{\beta} AW_T(c) \\
&= \frac{P_{\geq t}}{(1-p_t)} \left((1-p_t) N_t^{\beta} AW_t(c) - \sum_{T=t+1}^{\infty} P_T^{|t} N_T^{\beta} AW_T(c) \right) \\
&= \frac{P_{\geq t}}{(1-p_t)} \left(N_t^{\beta} AW_t(c) - \sum_{T=t}^{\infty} P_T^{|t} N_T^{\beta} AW_T(c) \right),
\end{aligned}$$

and

$$\xi_t = \frac{\sum_{T=t}^{\infty} P_T^{|t} N_T^{\beta} AW_T(c) - N_t^{\beta} AW_t(c)}{(1-p_t)(c_t)^{-\eta} \sum_{T=t}^{\infty} P_T^{|t} N_T^{\beta-1}}. \quad \text{(A-1)}$$

A.3 Evolution of welfare differences

We show here in that $\Delta_p W$ is increasing with β and increasing in η. Recall that the planning horizon probability streams P_t^i and P_t^j are built from the hazard rate stream p_t according to the formulas: $P_t^i = p_t^i . \prod_{s<t}(1-p_s^i)$ and $P_t^j = p_t^j . \prod_{s<t}(1-p_s^j)$. Using previous notation, we have $P_{\geq t}^i = \prod_{s<t}(1-p_s^i) = \sum_{s\geq t} P_s^i$ and $P_{\geq t}^j = \prod_{s<t}(1-p_s^j) = \sum_{s\geq t} P_s^j$. Because $p_t^j \leq p_t^i$, we have obviously that $P_{\geq t}^j \geq P_{\geq t}^i$, which means that $\sum_{\tau=t}^{\infty}(P_\tau^j - P_\tau^i) \geq 0$. Note that $P_{\geq 0}^i = P_{\geq 0}^j = 1$. Then, we have the following lemma.

Lemma 1 *Let $(u_t)_{t\in\mathbb{N}}$ be a non null sequence such that $\sum_{s\geq t} u_s \geq 0$ for all t and $\sum_{s\geq 0} u_s = 0$. If a is an increasing (resp. decreasing) sequence then $\sum_{s\geq 0} a_s u_s$ is positive (resp. negative).*

Proof. The proof relies on the transformation of the sequence $(u_t)_{t\in\mathbb{N}}$. If we introduce $U_t = \sum_{s\geq t} u_s$, the conditions on $(u_t)_{t\in\mathbb{N}}$ become that U_t is non-negative. We have $u_t = U_t - U_{t+1}$. So $\sum_{s\geq 0} a_s u_s = \sum_{s\geq 0} a_s(U_s - U_{s+1}) = a_0.U_0 + \sum_{s\geq 1} U_s.(a_s - a_{s-1})$. □

Assuming that per capita consumption is increasing, $AW_t(c^j)$ is increasing. Furthermore, the derivative of N_t^β with respect to β is positive and increasing in t (because N_t is increasing). Thus, using Lemma 1 and the property of the sequence $P_t^j - P_t^i$, we obtain that $\Delta_p W$ is increasing in β.

On the other hand, $\Delta_p W$ can be written in the following way:

$$\Delta_p W = \sum_{t=0}^{\infty} N_t^{\beta-1} \cdot (P_t^j - P_t^i) \cdot TW_t(c^j),$$

where $TW_t(c^j) = \sum_{\tau=0}^{t} n_\tau \left[\frac{(c_\tau^j)^{1-\eta}}{1-\eta} - \frac{\bar{c}^{1-\eta}}{1-\eta}\right]$. Assuming that $c_\tau^j > \bar{c}$, we know by the same reasoning as above that each term $\frac{(c_\tau^j)^{1-\eta}}{1-\eta} - \frac{\bar{c}^{1-\eta}}{1-\eta}$ is decreasing in η. Hence, the derivative of $TW_t(c^j)$ with respect to η is decreasing in t (we keep adding negative derivatives), and so is $N_t^{\beta-1}$ because $\beta \leq 1$. Hence, by Lemma 1 and the property of the sequence $P_t^j - P_t^i$, we obtain that $\Delta_p W$ is decreasing in η.

B Appendix - The role of population ethics

B.1 Without climate damages

B.1.1 BAU vs. 3 °C

β											b (per °C)
0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[3.10^{-6}; 10^{-2}]$
BAU	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[10^{-6}; 2.10^{-6}]$
BAU	BAU	BAU	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-7}
BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	0

Δc_t causes ΔW, Δp_t counteracts (or plays no role)

Δp_t causes ΔW, Δc_t counteracts (or plays no role)

(a) $\eta = 2.0$

β											b (per °C)
0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[4.10^{-6}; 10^{-2}]$
BAU	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[10^{-6}; 3.10^{-6}]$
BAU	BAU	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-7}
BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	0

Δc_t causes ΔW, Δp_t counteracts (or plays no role)

Δp_t causes ΔW, Δc_t counteracts (or plays no role)

(b) $\eta = 2.5$

Table A: Preferred policy and contributions as a function of β and b without damages (BAU vs. 3 °C)

B.1.2 3 °C vs. 2 °C

β → 0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	b (per $^{\circ}C$)
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[10^{-4}; 10^{-2}]$
3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[5.10^{-6}; 10^{-5}]$
3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[3.10^{-6}; 4.10^{-6}]$
3°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2.10^{-6}
3°C	3°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-6}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-7}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

(light) Δc_t causes ΔW, Δp_t counteracts (or plays no role)

(dark) Δp_t causes ΔW, Δc_t counteracts (or plays no role)

(a) $\eta = 2.0$

β → 0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	b (per $^{\circ}C$)
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[10^{-4}; 10^{-2}]$
3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[3.10^{-6}; 10^{-5}]$
3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2.10^{-6}
3°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-6}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	2°C	10^{-7}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

(light) Δc_t causes ΔW, Δp_t counteracts (or plays no role)

(dark) Δp_t causes ΔW, Δc_t counteracts (or plays no role)

(b) $\eta = 2.5$

Table B: Preferred policy and contributions as a function of β and b without damages (3 °C vs. 2 °C)

B.2 With climate damages

B.2.1 BAU vs. 3 °C

β											b (per $°C$)
0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-2}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-3}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[10^{-7}; 10^{-4}]$
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

Δc_t causes ΔW, Δp_t counteracts (or plays no role)

Δp_t causes ΔW, Δc_t counteracts (or plays no role)

Δc_t and Δp_t cause ΔW

(a) $\eta = 2.5$

β											b (per °C)
0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[2.10^{-6}; 10^{-2}]$
BAU	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[10^{-7}; 10^{-6}]$
BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	0

Δc_t causes ΔW, Δp_t counteracts (or plays no role)

Δp_t causes ΔW, Δc_t counteracts (or plays no role)

(b) $\eta = 3.0$

Table C: Preferred policy and contributions as a function of β and b with damages (BAU vs. 3 °C)

B.2.2 3 °C vs. 2 °C

β											b (per $^\circ C$)
0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[10^{-4}; 10^{-2}]$
3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[3.10^{-6}; 10^{-5}]$
3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[10^{-6}; 2.10^{-6}]$
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	2°C	2°C	10^{-7}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

- Δc_t causes ΔW, Δp_t counteracts (or plays no role)
- Δp_t causes ΔW, Δc_t counteracts (or plays no role)
- Δc_t and Δp_t cause ΔW

(a) $\eta = 2.5$

β											b (per °C)
0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[10^{-4}; 10^{-2}]$
3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[10^{-6}; 10^{-5}]$
3°C	3°C	3°C	3°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	10^{-7}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

(b) $\eta = 3.0$

Table D: Preferred policy and contributions as a function of β and b with damages (3 °C vs. 2 °C)

C Appendix - The role of inequality aversion

C.1 Without climate damages

C.1.1 BAU vs. 3 °C

η										b (per $^\circ C$)
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[4.10^{-6}; 10^{-2}]$
3°C	3°C	3°C	3°C	BAU	BAU	BAU	BAU	BAU	BAU	3.10^{-6}
3°C	3°C	3°C	BAU	BAU	BAU	BAU	BAU	BAU	BAU	2.10^{-6}
3°C	3°C	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	$[10^{-7}; 10^{-6}]$
BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	0

Δc_t causes ΔW, Δp_t counteracts (or plays no role)

Δp_t causes ΔW, Δc_t counteracts (or plays no role)

(a) $\beta = 0.0$

η										b (% per $^\circ C$)
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[10^{-6}; 10^{-2}]$
3°C	3°C	BAU	BAU	BAU	3°C	3°C	3°C	3°C	3°C	10^{-7}
BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	0

Δc_t causes ΔW, Δp_t counteracts (or plays no role)

Δp_t causes ΔW, Δc_t counteracts (or plays no role)

(b) $\beta = 0.1$

η										b (% per $^\circ C$)
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[10^{-7}; 10^{-2}]$
BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	BAU	0

Δc_t causes ΔW, Δp_t counteracts (or plays no role)

Δp_t causes ΔW, Δc_t counteracts (or plays no role)

(c) $\beta = 1.0$

Table E: Preferred policy and contributions as a function of η and b without damages (BAU vs. 3 °C)

C.1.2 3 °C vs. 2 °C

η										
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	b (per $^\circ C$)
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[10^{-4}; 10^{-2}]$
2°C	2°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[10^{-6}; 10^{-5}]$
2°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-7}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

Δc_t causes ΔW, Δp_t counteracts (or plays no role)

Δp_t causes ΔW, Δc_t counteracts (or plays no role)

(a) $\beta = 0.0$

η										
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	b (% per $^\circ C$)
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[6.10^{-6}; 10^{-2}]$
2°C	2°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	5.10^{-6}
2°C	2°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	$[3.10^{-6}; 4.10^{-6}]$
2°C	2°C	3°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	2.10^{-6}
2°C	2°C	3°C	3°C	3°C	3°C	2°C	2°C	2°C	2°C	10^{-6}
2°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	2°C	2°C	10^{-7}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

Δc_t causes ΔW, Δp_t counteracts (or plays no role)

Δp_t causes ΔW, Δc_t counteracts (or plays no role)

(b) $\beta = 0.1$

η										
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	b (% per $^\circ C$)
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[10^{-7}; 10^{-2}]$
2°C	3°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-7}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

Δc_t causes ΔW, Δp_t counteracts (or plays no role)

Δp_t causes ΔW, Δc_t counteracts (or plays no role)

(c) $\beta = 1.0$

Table F: Preferred policy and contributions as a function of η and b without damages (3 °C vs. 2 °C)

C.2 With climate damages

C.2.1 BAU vs. 3 °C

η										b (per °C)
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-2}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-3}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[3.10^{-6}; 10^{-4}]$
3°C	3°C	3°C	3°C	3°C	3°C	3°C	BAU	BAU	BAU	2.10^{-6}
3°C	3°C	3°C	3°C	3°C	BAU	BAU	BAU	BAU	BAU	$[10^{-7}; 10^{-6}]$
3°C	3°C	3°C	3°C	3°C	BAU	BAU	BAU	BAU	BAU	0

(a) $\beta = 0$

η										b (per °C)
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-2}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-3}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[3.10^{-6}; 10^{-4}]$
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	2.10^{-6}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[10^{-7}; 10^{-6}]$
3°C	3°C	3°C	3°C	3°C	BAU	BAU	BAU	BAU	BAU	0

(b) $\beta = 0.1$

η										b (per °C)
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-2}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-3}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[3.10^{-6}; 10^{-4}]$
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	2.10^{-6}
3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[10^{-7}; 10^{-6}]$
3°C	3°C	3°C	3°C	3°C	BAU	BAU	BAU	BAU	BAU	0

- Δc_t causes ΔW, Δp_t counteracts (or plays no role)
- Δp_t causes ΔW, Δc_t counteracts (or plays no role)
- Δc_t and Δp_t cause ΔW

(c) $\beta = 1$

Table G: Preferred policy and contributions as a function of η and b with damages (BAU vs. 3 °C)

C.2.2 3 °C vs. 2 °C

η										b (per $°C$)
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-2}
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-3}
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-4}
2°C	2°C	2°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	10^{-5}
2°C	2°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	$[10^{-7}; 4.10^{-6}]$
2°C	2°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

(a) $\beta = 0$

η										b (per $°C$)
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-2}
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[9.10^{-6}; 10^{-3}]$
2°C	2°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	3.10^{-6}
2°C	2°C	3°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	$[10^{-6}; 2.10^{-6}]$
2°C	2°C	3°C	3°C	3°C	3°C	3°C	3°C	2°C	2°C	10^{-7}
2°C	2°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

(b) $\beta = 0.1$

η										b (per $°C$)
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-2}
2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	2°C	$[[10^{-6}; 10^{-3}]$
2°C	2°C	3°C	3°C	2°C	2°C	2°C	2°C	2°C	2°C	10^{-7}
2°C	2°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	3°C	0

- Δc_t causes ΔW, Δp_t counteracts (or plays no role)
- Δp_t causes ΔW, Δc_t counteracts (or plays no role)
- Δc_t and Δp_t cause ΔW

(c) $\beta = 1$

Table H: Preferred policy and contributions as a function of η and b with damages (3 °C vs. 2 °C)

C.3 Evolution of the welfare difference with inequality aversion

We examine the evolution of the terms of the welfare difference with η. Figure A shows the behaviour of $\Delta_c W$ and $\Delta_p W$ with η in the case ($\beta = 0.1$, $b = 10^{-6}$ per °C). This figure should be compared to the third row from the bottom of table Fb. The figure clearly shows that both terms decrease with η, but at a different pace.

The terms of the welfare difference cross twice. This coincides with the result that the preferred policy switches twice, first from the 2°C scenario to the 3°C scenario at low η, then from the 3°C scenario to the 2°C scenario at high η (for $b = 10^{-6}$ per °C). These results are consistent with the analysis presented in section 2.2.3 which showed that anything coud be expected regarding the evolution of the preferred policy with respect to ethical parameters, as the welfare difference ΔW is a difference between two positive quantities that behave in the same way with respect to each ethical parameter.

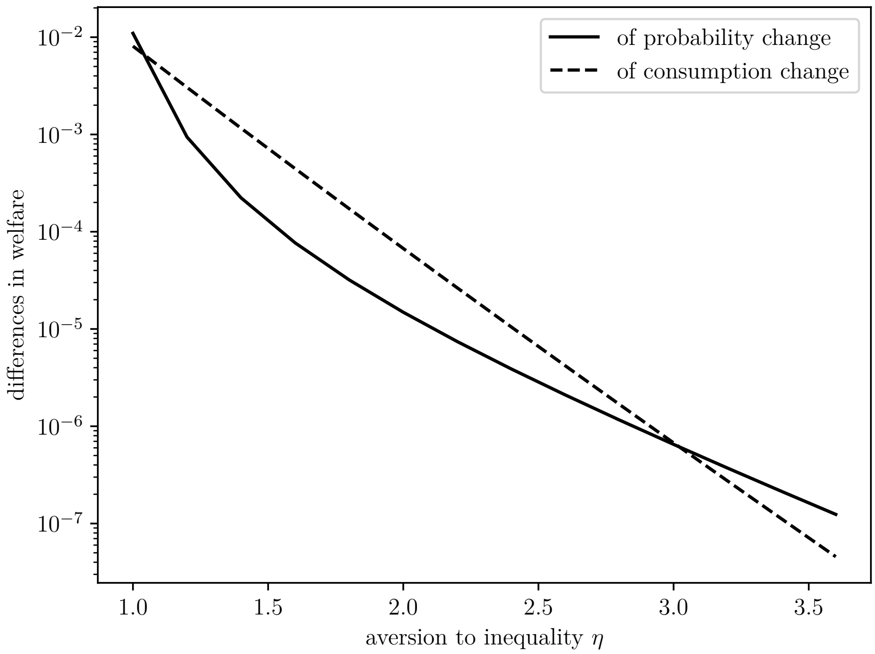


Figure A: Contributions to the welfare difference between *2°C* and *3°C* as functions of η ($\beta = 0.1$, $b = 10^{-6}$ per °C)