



## Catastrophic climate change, population ethics and intergenerational equity

Aurélie Méjean, Antonin Pottier, Marc Fleurbaey, Stéphane Zuber

### ► To cite this version:

Aurélie Méjean, Antonin Pottier, Marc Fleurbaey, Stéphane Zuber. Catastrophic climate change, population ethics and intergenerational equity. *Climatic Change*, 2020, 163, pp.873-890. 10.1007/s10584-020-02899-9 . halshs-01599453v2

**HAL Id: halshs-01599453**

**<https://shs.hal.science/halshs-01599453v2>**

Submitted on 16 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Catastrophic climate change, population ethics and intergenerational equity

Aurélie Méjean · Antonin Pottier ·  
Marc Fleurbaey · Stéphane Zuber

Received: date / Accepted: date

**Abstract** Climate change threatens irreversible and dangerous impacts, possibly leading to extinction. The most relevant trade-off then 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. Using the class of number-dampened utilitarian social welfare functions, we compare different climate policies that change the probability of catastrophic outcomes yielding an early extinction. We analyse the role of inequality aversion and population ethics. Low inequality aversion and a preference for large populations favour the most ambitious climate policy, although there are cases where the effect of inequality aversion on the preferred policy is reversed. This is due to the fact that a higher inequality aversion both decreases the welfare loss of reducing consumption of the current generation, and also decreases the welfare gain of reducing the future risk of extinction.

**Keywords** Climate change · Catastrophic risk · Equity · Population · Climate-economy model

---

A. Méjean  
CIRED, CNRS  
45 bis avenue de la Belle Gabrielle, 94736 Nogent-sur-Marne Cédex, France  
Tel.: +33 (0) 1 43 94 73 72  
E-mail: mejean@centre-cired.fr

A. Pottier  
CIRED, EHESS

M. Fleurbaey  
Paris School of Economics (CNRS)

S. Zuber  
Paris School of Economics (CNRS)

## 1 Introduction

The risk of abrupt and irreversible changes to the climate is one of the five reasons for concern identified by the IPCC (McCarthy et al., 2001; Field et al., 2014). Such abrupt changes are generally associated with tipping points (Lenton et al., 2008), like the shutoff of the Atlantic thermohaline circulation, the collapse of the West Antarctic ice sheet or the dieback of the Amazon rainforest. They may have indirect impacts, for instance through increased migration and conflicts (Reuveny, 2007; Hsiang et al., 2013). Catastrophic outcomes can also arise from discontinuities in the response of socio-economic systems to climate change.

What is the impact of climate change catastrophes on climate policy making? In this paper, we study a specific instance of such a climate change catastrophe, namely human extinction. There is evidence that climate change may bring about a global catastrophe. One reason is that the level of warming itself is very uncertain and may be much larger than expected. Wagner and Weitzman (2015) have produced estimates of possible extreme temperature rises in a medium-high emission pathway. They estimate the probability of eventual warming of 6 C to be around 10%, and the probability of 10 C to be around 3%. Such levels may threaten the Earth inhabitability for humans (Sherwood and Huber, 2010). The climate tipping points could induce a runaway greenhouse effect where feedback loops reinforce the warming trends, creating overheating spirals well past survivability (Steffen et al., 2018). Last, climate change may trigger general warfare, the use of dangerous technology or other events that may entail human extinction.

Climate change is very unlikely to cause human extinction. However, even if human extinction is a very unlikely event, it is an event with large welfare consequences, which may be sufficient to justify a substantial policy response. Due to the speculative nature of the risk studied, our analysis should be taken mainly as a thought experiment, pointing to orders of magnitude and qualitative judgements instead of sharp, quantitative statements.

When climate change entails a risk of extinction of the human population, population ethics matters. Population ethics assesses the value of population size for social welfare. For example, the utilitarian social welfare function may be extended to “average utilitarianism” (Hurka, 1982), by looking at the average of people’s well-being, or to “total utilitarianism” (Sidgwick, 1907), by looking at the sum of people’s well-being. But other approaches are possible, for instance “critical-level utilitarianism” (Blackorby et al., 1997; Broome, 2004) or “number-dampened utilitarianism” (Ng, 1989). Total utilitarianism is implicit in many Integrated Assessment Models (henceforth IAMs; see e.g., Nordhaus 1994). Although it might seem obvious that human extinction would be a very bad thing, the issue is actually unclear (Broome, 2010). There have been lots of debates in the population ethics literature as each of the existing ethical theories about the value of population has intuitively unattractive implications (Parfit, 1984). As climate change policies are expected to affect the size of the world’s population, and different ethical theories may imply very

different conclusions about the value of these policies, this raises a serious difficulty for evaluating policies aimed at mitigating climate change. This has largely been ignored in the literature (Broome, 2012; Millner, 2013; Kolstad et al., 2014).

Another important ethical consideration for climate change is that of the distribution of well-being across individuals belonging to different generations. The utilitarian approach, discussed above in the context of population ethics, is insensitive to this distribution because it only looks at total well-being. On the contrary, prioritarian approaches look at the sum of transformed well-being, where the transformation through a concave function implies that an increase in well-being at low levels enhances social welfare more than a similar increase at high levels (see Adler, 2012, for a discussion of these issues and a presentation of prioritarianism). The degree of priority for worse-off individuals is measured by inequality aversion. In IAMs, priority is typically expressed with respect to consumption, and inequality aversion plays an important role in shaping optimal climate policy. When introducing extinction risk, inequality aversion will play an additional role: it may change the value of the population in the case where extinction does not happen (and thus the value of avoiding extinction). It is thus important to jointly study the role of population ethics and inequality aversion.

Using a broad class of ethical value functions – number-dampened utilitarian social welfare functions – we show that accounting for a risk of human extinction due to climate change alters the standard results of cost-benefit analysis carried out using IAMs. This is true even if the risk we consider is very small (up to four orders of magnitude below other exogenous causes of extinction). When assuming enduring technical progress and under high inequality aversion, these models usually recommend that we postpone climate change mitigation and let future, richer generations pay. We show that when accounting for a risk of extinction due to climate change, there are cases where a high inequality aversion favours the most ambitious climate policy.

Most of the climate economics literature about catastrophes has modelled them as large drop in consumption: either as the certainty-equivalent loss of risky catastrophic damages (Nordhaus and Boyer, 2000), or 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 (Hope, 2006; Ackerman et al., 2010; Weitzman, 2012).<sup>1</sup> The case of human extinction is different, as it is an irreversible regime shift in the sense that post-catastrophe welfare is independent from pre-catastrophe actions. In this paper, we thus choose a simplified representation of human extinction so that (i) once started, the extinction process is irreversible; (ii) it happens within a time step in our analysis.<sup>2</sup>

<sup>1</sup> The stochastic nature of catastrophes is also sometimes explicitly accounted for (Peck and Teisberg, 1995; Lontzek et al., 2015; Lemoine and Traeger, 2016).

<sup>2</sup> This modelling strategy follows a strand of literature (Cropper, 1976; Clarke and Reed, 1994; Gjerde et al., 1999; Tsur and Zemel, 2009) that considers that abrupt climate change introduces an irreversible regime shift. Characteristic (ii) is not an accurate description of

In this paper, we explicitly account for the risk of extinction of future generations. We use the IAM RESPONSE (Ambrosi et al., 2003; Pottier et al., 2015), representing the interaction between climate and the economy, to assess the impacts of three scenarios (a business-as-usual case, a 2  C and a 3  C stabilisation policy) in terms of future consumption paths and extinction risk associated to climate change. We analyse the role of inequality aversion and population ethics for policy evaluation within the class of number-dampened utilitarian social welfare functions. This paper is thus at the crossroads of the literature on catastrophic climate change and the nascent literature on population ethics applied to climate change. Although the importance of population growth for future greenhouse gas emissions has long been recognised (Gaffin and O’Neill, 1997; Kelly and Kolstad, 2001), and thus population control has been mentioned as a possible way to mitigate climate change (see Cafaro, 2012, for an advocacy), it is only very recently that the role of population ethics for climate policy analysis has started to be discussed. Adler and Treich (2015) propose a prioritarian framework to evaluate scenarios with variable population in the case of climate change. Scovronick et al. (2017) show that putting a larger weight on population size warrants tighter climate policy whereas Lawson and Spears (2018) discuss how the preference for large populations can be overturned when environmental constraints are accounted for. This paper is the first to examine the role of population ethics in the context of catastrophic climate change.

The paper is structured as follows. Section 2 presents the framework used to evaluate climate policies. The numerical analysis presented in section 3 shows that the preferred climate policy depends on the relative impact of ethical parameters on the welfare gain due to a lower hazard rate (risk effect) and the welfare loss due to a lower consumption stream (consumption effect). In particular, we find cases where increasing inequality aversion favours the most ambitious climate policy. This new result is explained by the differentiated impact of inequality aversion on the risk and consumption effects. Section 4 concludes.

## 2 Evaluating the effects of climate policies

Evaluating climate policies involves two steps: first, each climate policy is translated into a social prospect using a climate-economy model; second, social prospects are evaluated using a social welfare function, policies are then ordered in terms of social welfare. We present the social welfare function to value social prospects in section 2.1 and the climate-economy model that computes social prospects from policies in section 2.2.

---

how human extinction might come about (it may take several years). But this simplification can be justified by saying that the extinction process is predetermined and rapidly yields a situation that is equivalent to extinction in terms of welfare (for instance because the extinction may involve a lot of suffering for many years where people live very bad lives).

## 2.1 Number-dampened utilitarianism with a risk of extinction

Policy choices may impact consumption, population size and the probability of stochastic events. This paper focuses on the risk of extinction and therefore the only uncertain variable is the date at which extinction occurs. For a given policy, consumption and population are deterministic but conditional on existence. Key data of social outcomes from a policy are the sequence  $n = (n_0, n_1, \dots, n_t, \dots)$  which gives each generation's population size *conditional on existence*, and the sequence  $c = (c_0, c_1, \dots, c_t, \dots)$  which gives each generation's consumption per capita *conditional on existence*. Indeed, we do not consider intra-generational inequalities, and we assume that consumption is equally distributed among individuals of the same generation, so that we can refer either to a generation or to a (representative) individual of that generation. Generation  $t$ , if it exists, is thus of size  $n_t$  and enjoys consumption per capita  $c_t$ . We now introduce the social welfare function used throughout the paper in two steps, starting with sure social outcomes.

In this paper, we use a social welfare approach to rank policies according to the prospects they induce. Relying on a value function (the social welfare function), our approach is concerned with the goodness of the outcomes or consequences of actions, and thus belongs to a broad class of consequentialist theories (see Adler, 2012, for a presentation of the social welfare approach and its ethical underpinnings). It however means that we ignore some important justice issues such as those related to the notions of harm, virtues or duties.<sup>3</sup> We think that even people who do not think that consequentialism is the only relevant ethical view may still think that consequences are part of the ethical considerations we should rely on. Using a social welfare approach is at least a way to discuss some relevant justice issues.

### 2.1.1 Sure social outcomes and main parameters

First, consider a social outcome for which the date of extinction is known for sure. If the time horizon is  $H$ , only the first  $H$  generations exist and social welfare is generated by the size of generations  $n = (n_0, n_1, \dots, n_H)$  and their consumption per capita  $c = (c_0, c_1, \dots, c_H)$ . We thus write the social outcome  $(n, c, H)$ , as this triplet fully describes it. To value this social outcome, we follow most of the existing literature and assume that aggregate welfare  $V$  from outcome  $(n, c, H)$  has the following generalised utilitarian form<sup>4</sup>.

<sup>3</sup> See for instance Jamieson (2007) or Shue (2014) for examples of such approaches to climate justice.

<sup>4</sup> The economic literature usually uses the term 'generalised utilitarian' rather than prioritarian as we did in the introduction. The formula here is prioritarian with respect to consumption; it may be viewed as utilitarian if function  $u(c) = \frac{c^{1-\eta}}{1-\eta}$  is the true utility function. There exists a limited literature in population ethics proposing alternatives to this generalised 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).

**Definition 1 (Number-dampened utilitarian social welfare functions)**

A social welfare function is a *number-dampened utilitarian social welfare function* (NDU SWF) if there exist real numbers  $\beta \in [0, 1]$ ,  $\underline{c} \in \mathbb{R}_{++}$  and  $\eta \in \mathbb{R}_+$  such that:

$$V(n, c, H) = N_H^\beta \sum_{t=0}^H \frac{n_t}{N_H} \left[ \frac{c_t^{1-\eta}}{1-\eta} - \frac{\underline{c}^{1-\eta}}{1-\eta} \right], \quad (1)$$

where  $N_H = \sum_{t=0}^H n_t$  is the total (cumulated) population that comes into existence in social outcome  $(n, c, H)$ . The utility enjoyed by an individual of generation  $t$  is  $\frac{c_t^{1-\eta}}{1-\eta} - \frac{\underline{c}^{1-\eta}}{1-\eta}$ . The utility averaged across individuals and generations is:

$$AU_H(n, c) = \sum_{t=0}^H \frac{n_t}{N_H} \left[ \frac{c_t^{1-\eta}}{1-\eta} - \frac{\underline{c}^{1-\eta}}{1-\eta} \right],$$

so that social welfare (1) is the product of average utility and a population weight  $N_H^\beta$ . Note that if we consider social outcomes with the same total population, the terms  $N_H^\beta$  and  $\frac{\underline{c}^{1-\eta}}{1-\eta}$  do not play any role: our social welfare function then gives the same ranking as the usual utilitarian social welfare function  $\sum_{t=0}^H n_t \frac{c_t^{1-\eta}}{1-\eta}$ . However, this is not true anymore when population varies and can endogenously change with the policy, as in the case of the extinction risk we consider afterwards.

The NDU SWF is a variant of number-dampened utilitarianism proposed by Ng (1986), where welfare is the product of average utility and a concave function of total population size. The slight difference is that Ng's concave function has a finite limit at infinity to avoid Parfit's Repugnant conclusion, whereas ours has not. However, Blackorby et al. (2005) use the term number-dampened (general) utilitarianism for a generalisation of our specification. Several other authors have used the same variant of NDU as we do, but averaging utility over periods, not over the entire history, and computing the weight accordingly, see for instance Boucekine et al. (2014) or Lawson and Spears (2018).

A key property of our NDU SWF is that it does not exhibit pure time discounting: contrary to most approaches in climate economics, it treats generations in a symmetric way.<sup>5</sup> This also implies that only the total number of individuals at a given consumption welfare matters, irrespective of their distribution across time. We describe the main axiomatic properties of our NDU SWF in the Electronic supplementary material (section S.A).

The NDU SWF embodies the preferences of an evaluator through three important ethical parameters.

<sup>5</sup> See Ramsey (1928) and Stern (2007) for arguments on why we should treat generations in a symmetric way. Observe that one reason for discounting, namely the possibility that future generations may not exist, is not yet accounted for, given that the existence (or non-existence) of generations is for the moment known for sure. The next section will show that the extinction risk creates endogenous discounting.

1. Parameter  $\eta$  is the (intergenerational) inequality aversion coefficient. It determines the marginal utility of individual consumption. A high value of  $\eta$  implies that the utility function of individuals is strongly concave in consumption, so that marginal utility at high consumption levels is much lower than at low consumption levels. A more inequality averse social welfare function means that we are willing to sacrifice more to equalise consumption levels across generations.
2. Parameter  $\beta$  is the population ethics coefficient and determines the value of large populations. Indeed, Eq. (1) embeds well-known views of utilitarianism when population size varies. *Total* or *classical utilitarianism* values the total sum of utilities, which corresponds to  $\beta = 1$ . *Average utilitarianism*, on the contrary, values the average utility, which corresponds to  $\beta = 0$ . Varying  $\beta$  between 0 and 1 spans cases between the total and the average utilitarian views.
3. Parameter  $\underline{c}$  is a consumption per capita threshold. The interpretation of this threshold is that it corresponds to the neutral level of utility, i.e. the level above which a life is worth living (life is assumed to be not worth living below that level of consumption).<sup>6</sup> In the case of total utilitarianism ( $\beta = 1$ ), it is also the *critical level of consumption*, i.e. the level of consumption such that, if enjoyed by an additional individual, total welfare is left unchanged when that individual is added to the population. This parameter plays a key role in critical-level generalised utilitarianism (Broome, 2004; Blackorby et al., 2005). When  $\beta \neq 1$  however, the critical level is not constant and depends on population size and average utility. Parameter  $\underline{c}$  still influences aggregate welfare and the value of changing population size.

### 2.1.2 Risky social prospects

Consider now a risky social prospect. As explained before, the only risk considered here is a risk on the existence of future generations. Hence, a social prospect is still partly described by a deterministic population stream  $n$  and a deterministic consumption stream  $c$ , i.e. population size is  $n_t$  and consumption per capita is  $c_t$ , conditional on existence at period  $t$ . Due to the risk of extinction, the time horizon is unknown. The risk of extinction is specified as an infinite stream of hazard rates  $p$ . The hazard rate  $p_t$  is the probability of extinction at the end of period  $t$ , conditional on existence at the beginning of period  $t$ . The survival probability at  $t$  is  $P_{\geq t} = \prod_{\tau=0}^{t-1} (1 - p_\tau)$  (probability that the time horizon is greater than or equal to  $t$ ). The probability that humanity becomes extinct right after generation  $t$  is  $P_t = p_t P_{\geq t}$ , which we name the horizon probability, i.e. the probability that the time horizon is  $t$ . The risky

<sup>6</sup> The notion of the neutral level of utility is important in the literature on population axiologies, as many paradoxes or principles (for instance “avoiding a repugnant conclusion”, or the “mere addition principle”), rely on the notion (see Arrhenius, forth.). The notion has been defended for instance by Holtug (2001). Still, there are discussions about the exact meaning of “a life is worth living” and the level of  $\underline{c}$ .



social prospect is fully represented by a triplet  $(n, c, p)$ . To extend the NDU SWF to social prospects with unknown time horizons, we simply assume that aggregate welfare  $W$  is the expected value of the NDU welfare of each possible outcome.

**Definition 2 (Expected number-dampened utilitarian social welfare functions)** A social welfare function is an *expected number-dampened utilitarian social welfare function* (ENDU SWF) if there exist real numbers  $\beta \in [0, 1]$ ,  $\underline{c} \in \mathbb{R}_{++}$  and  $\eta \in \mathbb{R}_+$  such that:

$$\begin{aligned} W(n, c, p) &= \sum_{H=0}^{\infty} P_H \left( N_H^{\beta} \sum_{t=0}^H \frac{n_t}{N_H} \left[ \frac{c_t^{1-\eta}}{1-\eta} - \frac{\underline{c}^{1-\eta}}{1-\eta} \right] \right) \\ &= \sum_{t=0}^{\infty} \underbrace{\left( \sum_{H=t}^{\infty} P_H N_H^{\beta-1} \right)}_{\theta_t} n_t \left[ \frac{c_t^{1-\eta}}{1-\eta} - \frac{\underline{c}^{1-\eta}}{1-\eta} \right]. \end{aligned} \quad (2)$$

By definition  $W(n, c, p) = \sum_{H=0}^{\infty} P_H V(n, c, H)$ , so that  $V(n, c, H) = W(n, c, 1_H)$  (where  $1_H$  is the prospect assigning probability 1 to extinction at date  $H$  and 0 otherwise). So there are two ways by which population size may vary in our model. First, the deterministic path  $n$  of population size conditional on existence may vary. Second, the probability  $p$  of the extinction risk may vary, making expected population size and the distribution of population sizes endogenous. For instance, comparing prospects  $(n, c, 1_H)$  and  $(n, c, 1_{H+1})$  requires to be able to compare populations with different sizes. In the rest of the paper, we will assume that policy does not affect  $n$ , i.e. that it does not affect a generation's population size, conditional on its existence.<sup>7</sup> However, given that policy will affect the hazard rates  $p$ , the existence of future generations and the distribution of population size will be endogenous. Population ethics considerations will therefore matter for policy assessment.

In Eq.(2), the scalar  $\theta_t$  is tantamount to a time discount factor on the utility of generation  $t$ . This discount factor is not related to pure time discounting but stems from the uncertainty about the horizon, a rationale for discounting already discussed in the Stern Review (2007). It is actually an endogenous discount factor that depends on the stream of hazard rate  $p$ . It also depends on the attitudes towards population size as embodied in the population ethics coefficient  $\beta$ . This role of views about population ethics on discounting has been discussed before by Fleurbaey and Zuber (2015) and Scovronick et al. (2017). When  $\beta = 1$ , the discount factor  $\theta_t$  on the utility of generation  $t$  is exactly its probability of existence. When  $\beta < 1$ , this discount factor is the probability of existence multiplied by a factor which is decreasing with time, so that future generations' utility is discounted beyond what follows from their probability of existence.<sup>8</sup>

<sup>7</sup> Section S.A. in the Electronic supplementary material considers the possibility that policy choices affect population size from a theoretical point of view.

<sup>8</sup> See also section S.A.2. of the Electronic Supplementary Material for a clarification of the discount rate and the underlying Ramsey formula in our framework.

## 2.2 Modelling the social outcomes of climate policies

We present the climate-economy model (section 2.2.1) and the way the risk of extinction is modelled (section 2.2.2).

### 2.2.1 The climate-economy model

To translate climate policies into social prospects to be evaluated by the ENDU SWF, we use the IAM RESPONSE (Ambrosi et al., 2003), which belongs to the tradition of compact IAMs (Nordhaus, 1994; Hope et al., 1993; Tol, 1997). It combines a simple representation of the economy and a climate module. The economic module is a Solow-like growth model with capital accumulation and exogenous population. It includes climate mitigation costs and a climate damage function, that is chosen here to be equivalent to that of DICE. A key parameter is exogenous technical progress (Total Factor Productivity growth), which decreases over time.<sup>9</sup> A thorough description of the model and its equations can be found in (Dumas et al., 2012) and (Pottier et al., 2015).

A policy (or scenario)  $i$  is specified as the decision variables of the model, which consists of a saving rate stream and an abatement stream. The model computes outputs from these decision variables. Model outputs include the stream of consumption per capita  $c^i$ . Because population (conditional on existence) is exogenous, the generation size stream  $n$  does not depend on policy  $i$  (it will thus be common to all social prospects that will be evaluated). To complete the description of the social prospect  $(n, c^i, p^i)$  of policy  $i$ , we only need to specify how the hazard rate  $p$  is computed.

### 2.2.2 The risk of extinction

Before describing how the hazard rate is modelled, let us motivate the introduction of a risk of extinction due to climate change. In our model (which follows standard specifications), climate damages on consumption are not sufficient *per se* to trigger the extinction of the human population. We consider that the risk of extinction does not stem from dwindling consumption, but from a societal regime shift induced by climate change. Such a regime shift can be initiated by extreme climate events, climate tipping points, or for instance an epidemic outbreak due to global warming. Such initiating events would provoke a chain reaction, e.g. through inappropriate policy responses, large migration waves, general warfare, which would challenge the physical, political and social infrastructures of global society, leading to its collapse. This is the type of event described under the label “risk of extinction”: a catastrophe that destroys society and its institutions.

<sup>9</sup> In this paper, we assume a decreasing Total Factor Productivity (TFP) growth due to the very long time horizon considered (of the order of 10,000 years). Indeed, assuming constant TFP 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).

We model extinction as an abrupt event that occurs over the course of one period (a decade). This is of course a harsh simplification of actual processes, which would be gradual and involve feedbacks between institutions, consumption, mortality and fertility. In reality, the ultimate consequences of such a catastrophe would probably take decades to unfold. Following the work by Cropper (1976) and Clarke and Reed (1994) in modelling environmental catastrophes, we consider that what is crucial for our analysis is that post-catastrophe variables do not depend on pre-catastrophe variables, so that a transfer of wealth (or population, which is more important for population ethics) cannot occur across the time of the catastrophe, i.e., the course of post-catastrophe events is inevitably set and cannot be acted upon. Since post-catastrophe welfare cannot be changed, we abstract from the details of the extinction process and adopt the simplifying assumption that what triggers the catastrophe also represents the end of time.

Whether such a catastrophic chain of events would actually lead to the extinction of all human beings is controversial and may be considered as highly unlikely. An important point of this paper is that even a remote possibility of human extinction may be sufficient to justify a policy response (see Section 3.2). The probability of the extinction risk is however difficult to determine as its calibration cannot rely on the frequentist approach (no similar event has ever happened). Given the large uncertainty surrounding the probability of such events, we decided to remain neutral on the issue and to explore many possibilities. We thus attempt to cover a wide range of possible values.

To further simplify this first inquiry into the role of the risk of extinction, we assume that the hazard rate  $p$  depends on the temperature increase  $T$  only, and not, for example, on wealth or on a proxy of adaptive capacity. We therefore disregard any factor, social or natural, that may mitigate the effect of climate change on the risk of extinction. More precisely, we model the hazard rate  $p$  as a linear function of temperature increase  $T$  (in  $^{\circ}\text{C}$  above the pre-industrial level), above a temperature threshold  $T_0$  (set at  $1^{\circ}\text{C}$ ). The hazard rate at date  $t$ ,  $p_t$ , is thus  $p(T_t)$ , a function of temperature increase  $T_t$  at date  $t$ . We assume that this function is valid in a range of temperature increase spanning from  $1^{\circ}\text{C}$  to  $10^{\circ}\text{C}$ , which includes the values of temperature increase presented in the numerical analysis (cf. section 3).

$$p(T) = \begin{cases} p_0, & \text{if } T \leq T_0 \\ p_0 + b \cdot (T - T_0), & \text{if } T_0 \leq T \leq T_0 + \frac{1-p_0}{b} \\ 1, & \text{if } T_0 + \frac{1-p_0}{b} \leq T, \end{cases} \quad (3)$$

where  $p_0$  is the exogenous minimum hazard rate (set at  $10^{-3}$  per annum), and  $b$  is the marginal hazard rate per  $^{\circ}\text{C}$  above  $T_0$ .

As argued above, there is no available data about its realisation from which the parameters of the hazard rate function could be inferred. The calibration can thus only be illustrative and tentative. We retain a range of  $b$  between  $10^{-1}$  and  $10^{-7}$  per  $^{\circ}\text{C}$  per annum (see the Electronic supplementary material, section S.B.1, for a detailed description).

### 2.3 Decomposition of the welfare impacts of policy change

Climate policies affect social prospects, as they alter both the consumption path and the probability of catastrophic outcomes yielding an early extinction. For example, a stringent mitigation policy that reduces greenhouse gas emissions compared to a business-as-usual case will reduce consumption now and in the short run, but increase consumption in the long run by avoiding climate damages; the hazard rate in all future periods will be also reduced, as temperature increase is lower.

We want to better understand how the preferred policy varies with the ethical parameters of the social welfare function, namely the inequality aversion parameter  $\eta$  and the population ethics parameter  $\beta$ . We may think that these parameters actually have different effects depending on whether we look at the impact of policy on consumption or at its impact on the extinction risk. We use a welfare decomposition method (described in Electronic supplementary material, section S.C.1 in particular Table S.3) in the numerical analysis below to determine which effect (the risk effect or the consumption effect) is larger depending on the value of ethical parameters  $\eta$  and  $\beta$ .

## 3 Numerical evaluation of three climate policies

A numerical analysis is necessary to understand the role of ethical parameters on the preferred policy. We set out three climate policies (section 3.1) and rank them using the ENDU SWF. We examine specifically the role of the risk of extinction (section 3.2) and the role of ethical parameters (section 3.3) in determining the preferred climate policy among the set of policies considered.

### 3.1 The numerical experiment: climate policies and their social outcomes

We introduce the climate policies and parameter ranges used in the numerical experiment. We consider three climate policies: 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-industrial level to 3 °C and 2 °C. We provide a detailed description of the policies considered in the Electronic supplementary material, section S.B.2.

In terms of ethical parameters used for the evaluation, 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. We test a wide range of values for the inequality aversion coefficient  $\eta$ , spanning from 0.5 to 5. The threshold per capita consumption parameter is calibrated on an extreme poverty consumption level, as in (Bommier et al., 2015). The parameter value ranges are summarised in Table S.2 in the Electronic supplementary material.

We do not formally consider other sources of uncertainty than the risk of extinction in this paper. Our aim is to illustrate and give broad orders of magnitude about the trade-off between present consumption and the existence of future generations. We keep the rest of the analysis as simple as possible, and use central values for uncertain parameters such as the climate sensitivity, the convexity of the damage function or the rate of technical progress in the numerical assessment.

### 3.2 Results: The role of the risk of extinction and climate damages

This section examines the effect of the risk of extinction on the preferred policy between a business-as-usual scenario (BAU) and a 3  C scenario<sup>10</sup>, with and without accounting for climate damages. When not accounting for climate damages on consumption, we 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  $c$  and the stream of hazard rate  $p$  over the whole period<sup>11</sup>. The most ambitious climate policy would therefore be favoured because of a reduction in the risk of extinction, while the least ambitious policy would be favoured because of an increase in consumption in early periods.

We examine the preferred policy option between BAU and 3  C as a function of the marginal hazard rate  $b$  for a given degree of inequality aversion ( $\eta = 2.0$ ) and for a given value of the population ethics coefficient ( $\beta = 1.0$ , i.e. the case of total utilitarianism).

		$b$ (per $^{\circ}C$ )							
		0	$10^{-7}$	$10^{-6}$	$10^{-5}$	$10^{-4}$	$10^{-3}$	$10^{-2}$	$10^{-1}$
w/o damages	BAU	3°C	3°C	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	3°C	3°C
		<div><div></div>social outcome preferred due to higher consumption</div> <div><div></div>social outcome preferred due to lower risk</div> <div><div></div>social outcome preferred due to higher consumption and lower risk</div>							

**Table 1:** Preferred policy (BAU vs. 3  C) with and without damages ( $\beta = 1.0$ ,  $\eta = 2.0$ )

<sup>10</sup> We do not include the 2  C scenario in this section for the sake of clarity.

<sup>11</sup> 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 the case if climate damages are accounted for, cf. Figure S.3 in the Electronic supplementary material.

The first row of table 1 shows the case without climate change damages. The results show that in that case, 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). With a marginal hazard rate equal to zero, mitigation has no impact on extinction and only reduces consumption. So, climate policy necessarily reduces welfare if we exclude climate damages. With a marginal hazard rate  $b$  greater than zero, there is a chance that climate action may avoid extinction, and the 3°C scenario is favoured over the BAU. The first row of table 1 shows that the difference in the hazard rate always explains the preference for the 3°C scenario, while the difference in consumption counteracts (dark grey cells). Conversely, the difference in consumption explains the preference for the BAU scenario when the hazard rate is purely exogenous (i.e., for a marginal hazard rate  $b = 0$  per °C per annum), as hazard rate streams are then 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.

The second row of table 1 shows the case with climate damages. 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 in table 1 show that the preferred policies are unchanged whether or not climate damages are accounted for, with the exception of  $b = 0$  per °C per annum, i.e. for a purely exogenous hazard rate. The second row of table 1 shows that when damages are accounted for, the 3°C scenario is preferred due to both the risk and the consumption effects (black cells, again with the exception of  $b = 0$  per °C per annum). This contrasts with the case without climate damages, where the 3°C scenario was preferred due to the risk effect alone (dark grey cells), while the consumption effect played in favour of the BAU (light grey cells). 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 per annum), 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  $10^{-3}$  per annum, 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 climate damages, the 3°C scenario is no longer preferred due to higher consumption when  $b$  increases from  $10^{-3}$  to  $10^{-2}$  per °C per annum. 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  $b$ , 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.

### 3.3 Results: 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 population ethics coefficient ( $\beta$ ), and the role of inequality aversion ( $\eta$ ).

#### 3.3.1 The role of population ethics

In table 2, we examine preferred policy options (between the BAU, 3  C and 2  C scenarios)<sup>12</sup> as a function of the marginal hazard rate  $b$  for various population ethics coefficients ( $\beta$ ) and for given values of the inequality aversion coefficient ( $\eta = 1.5$  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 2.3).

$\beta$											$b$ (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	$[10^{-5}; 10^{-1}]$
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)  $\eta = 1.5$**

$\beta$											$b$ (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	$[10^{-4}; 10^{-1}]$
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)  $\eta = 2.5$**

social outcome preferred due to higher consumption

social outcome preferred due to lower risk

**Table 2:** Preferred policy (BAU vs. 3  C vs. 2  C) as a function of  $\beta$  and  $b$  (with climate damages)

<sup>12</sup> See section S.E.2. of the Electronic supplementary material for the comparison between the 3  C and BAU scenarios, and between the 3  C and 2  C scenarios.

For a given marginal hazard rate  $b$ , a higher population ethics coefficient 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  $b$ , the value of  $\beta$  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 (dark grey cells). Conversely, the difference in consumption always explains the preference for the least ambitious climate scenario (light grey cells), except when the hazard rate is purely exogenous (i.e. for a marginal hazard rate  $b = 0$  per °C per annum). Note that the results in table 2 show that for the values of  $\eta$  considered, the BAU scenario is never preferred when climate damages are accounted for.

### 3.3.2 The role of inequality aversion

We here examine preferred policy options as a function of the risk of extinction (i.e. as a function of the marginal hazard rate  $b$ ) for various degrees of inequality aversion ( $\eta$ ) and for given values of the population ethics coefficient ( $\beta = 0, 0.1, 1$ ). We examine the preferred policy between the BAU, 3 °C and 2 °C scenarios<sup>13</sup> in table 3.

In the case of average utilitarianism ( $\beta = 0$ , table 3a), at a given level of risk (i.e., at a given  $b$ ), a higher inequality aversion favours the least ambitious climate policy. As the risk of extinction is reduced (i.e., as the marginal hazard rate  $b$  decreases), the minimum level of inequality aversion that justifies the least ambitious scenario is reduced: with a lower risk of extinction the catastrophe is postponed and richer generations are added, which enhances inequalities between generations because the present generation is still assumed to be relatively poorer. For a high extinction risk (i.e., for large 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 lower risk and higher consumption (black cells) for relatively low values of inequality aversion ( $\eta \leq 1.0$ ). 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 plays in favour of the least ambitious policy (see section S.E.1. of the Electronic supplementary material). The difference in consumption is indeed expected to favour the most ambitious climate policy for low inequality aversion only, 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 BAU scenario due to technical change, this effect only occurs for low inequality aversion.

Note that, as before, the contributions of the difference in extinction risk and the contribution of the difference in consumption in determining the preferred policy are calculated when comparing the preferred policy and the next

<sup>13</sup> See section S.E.2 of the Electronic supplementary material for the comparison between the 3 °C and BAU scenarios, and between the 3 °C and 2 °C scenarios.



$\eta$										$b$ (per $^{\circ}C$ )
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	$[10^{-2}; 10^{-1}]$
2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	$10^{-3}$
2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	$10^{-4}$
2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	$10^{-5}$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	$[8.10^{-6}; 9.10^{-6}]$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	$[4.10^{-6}; 7.10^{-6}]$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	$3.10^{-6}$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	BAU	BAU	BAU	$2.10^{-6}$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	BAU	BAU	BAU	BAU	BAU	$10^{-6}$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	BAU	BAU	BAU	BAU	BAU	$10^{-7}$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	BAU	BAU	BAU	BAU	BAU	0

**(a)  $\beta = 0$**

$\eta$										$b$ (per $^{\circ}C$ )
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	$[10^{-2}; 10^{-1}]$
2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	$[4.10^{-6}; 10^{-3}]$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	$3.10^{-6}$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	$2.10^{-6}$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	$10^{-6}$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	$10^{-7}$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	BAU	BAU	BAU	BAU	BAU	0

**(b)  $\beta = 0.1$**

$\eta$										$b$ (per $^{\circ}C$ )
0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	$[10^{-2}; 10^{-1}]$
2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	$[10^{-6}; 10^{-3}]$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	2 $^{\circ}C$	$10^{-7}$
2 $^{\circ}C$	2 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	3 $^{\circ}C$	BAU	BAU	BAU	BAU	BAU	0

**(c)  $\beta = 1.0$**

social outcome preferred due to higher consumption

social outcome preferred due to lower risk

social outcome preferred due to higher consumption and lower risk

**Table 3:** Preferred policy (BAU vs. 3 $^{\circ}C$  vs. 2 $^{\circ}C$ ) as a function of  $\eta$  and  $b$  (with climate damages)

best policy. When the 3 $^{\circ}C$  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 3a), and when the next best policy is the 2 °C, the corresponding cell is light grey (i.e., the 3 °C scenario is preferred to the 2 °C due to higher consumption); when the next best policy is the BAU, the corresponding cell is dark grey (i.e., the 3 °C scenario is preferred to the BAU due to the difference in hazard rate). More results comparing policies in pairs, with and without accounting for climate damages, are presented in the Electronic supplementary material (section S.E.).

The pattern is different for  $\beta = 0.1$  (table 3b). In this case, increasing the inequality aversion  $\eta$  still favours the least ambitious climate policy for low values of  $\eta$ , but the effect is reversed for  $\eta > 2.0$ . There are two opposite effects of increasing the inequality aversion coefficient: it decreases the value of reducing the risk of extinction but it also decreases the cost of implementing the policy by making the first (relatively poorer) generations pay. These intuitions are formalised in Propositions S.1 and S.2 in the Electronic supplementary material (section S.C.2). Both effects are also at play in the case  $\beta = 0$ , as shown in the analytical results presented in the section S.C.2 Electronic supplementary material. However, in the numerical results, we find instances where the decrease in the costs of implementing the policy dominates only for  $\beta > 0$ .

The choice of  $\beta$  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 ( $\beta = 0$ , table 3a), inequality aversion plays a role on the preferred policy option for a wider range of  $b$  values than in the case of total utilitarianism ( $\beta = 1$ , table 3c). The choice of  $\beta$  does not influence the preferred policy when the extinction risk is exogenous (i.e for a marginal hazard rate  $b$  equal to zero).

## 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 which preserves the consumption of the present, poorer generation. On the other hand, abatement can prevent extinction, which can favour ambitious climate policies. Our main results are summarised below.

The preferred policy depends on the relative effects of inequality aversion and of the preference for large populations on the welfare gain due to a lower risk of extinction, and on the welfare loss due to lower consumption. A high inequality aversion reduces the welfare gain due to higher consumption. It also reduces the value of postponing extinction. A preference for large populations increases both the welfare loss due to lower consumption, and the welfare gain as the size of the cumulative population increases due to a lower risk of extinction.

A preference for large populations always favours the most ambitious climate scenario. This result is consistent with our first intuition: a preference for large populations gives a larger weight to the welfare of future generations. Low inequality aversion 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 a higher inequality aversion favours the most ambitious climate policy. This new result is explained by the relative effect of inequality aversion on the risk and consumption components of the difference in welfare between policies. 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.

The effect of ethical choices on the preferred policy should be further explored. One may for instance want 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. One could also include the possibility of partial extinction, i.e. 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.

As mentioned before, our work makes several simplifying assumptions and future work could explore whether those can be relaxed to provide more realistic assessments. First, the probability of extinction depends only on global temperature. One could think that the adaptive capacities of human may depend to some extent on income or wealth. If costly adaptation measures can be taken to survive (for instance we can get protection from heat through air conditioning), humanity would not face extinction even if temperature increase is very large. Also, we could study the consequence of a global catastrophe that is not extinction but a large civilisation collapse. The problem here would be to choose a precise scenario for such a collapse (followed perhaps by a recovery).

**Acknowledgements** The authors would like to thank those who provided valuable feedbacks, in particular three anonymous reviewers.

## Funding

This research has been supported by the Agence nationale de la recherche through the Fair-ClimPop project (ANR-16-CE03-0001-01) and Investissements d’Avenir program (ANR-10-LABX-93). Financial support from the Swedish foundation for humanities and social sciences is also gratefully acknowledged (Anslag har erhållits från Stiftelsen Riksbankens Jubileumsfond).

---

**Conflict of interest**

The authors declare that they have no conflict of interest.

**Author attribution**

A.M., A.P., M.F., and S.Z. jointly wrote the paper, designed the research, and analysed the results. A.M. and A.P. conducted model simulations. A.M. generated figures and tables. A.P. developed the original model.

## References

- Ackerman F, Stanton EA, Bueno R (2010) Fat tails, exponents, extreme uncertainty: Simulating catastrophe in DICE. *Ecological Economics* 69(8):1657–1665, DOI 10.1016/j.ecolecon.2010.03.013
- Adler MD (2012) *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. Oxford University Press
- Adler MD, Treich N (2015) Prioritarianism and Climate Change. *Environmental and Resource Economics* 62(2):279–308
- Ambrosi P, Hourcade JC, Hallegatte S, Lecocq F, Dumas P, Duong MH (2003) Optimal Control Models and Elicitation of Attitudes towards Climate Damages. *Environmental Modeling & Assessment* 8(3):133–147, DOI 10.1023/A:1025586922143
- Arrhenius G (forth.) *Population Ethics*. Oxford University Press, Oxford
- Asheim G, Zuber S (2014) Escaping the repugnant conclusion: Rank-discounted utilitarianism with variable population. *Theoretical Economics* 9:629–650
- Asheim G, Zuber S (2016) Evaluating intergenerational risks. *Journal of Mathematical Economics* 65:73–96
- Blackorby C, Bossert W, Donaldson D (1996) Leximin population ethics. *Mathematical Social Sciences* 31:115–131
- Blackorby C, Bossert W, Donaldson D (1997) Critical-Level Utilitarianism and the Population-Ethics Dilemma. *Economics and Philosophy* 13(02):197, DOI 10.1017/S026626710000448X
- Blackorby C, Bossert W, Donaldson D (2005) *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*. Cambridge University Press
- Bommier A, Lanz B, 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, DOI 10.1016/j.jeem.2015.07.003
- Boucekkine R, Fabbri G, 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 (2010) The most important thing about climate change. In: Boston J, Bradstock A, Eng D (eds) *Public Policy: Why Ethics Matters*, ANU Press, pp 101–116
- Broome J (2012) *Climate matters: ethics in a warming world*, 1st edn. Amnesty International global ethics series, W.W. Norton, New York
- Cafaro P (2012) Climate ethics and population policy. *Wiley Interdisciplinary Reviews: Climate Change* 3(1):45–61
- Clarke HR, Reed WJ (1994) Consumption/pollution tradeoffs in an environment vulnerable to pollution-related catastrophic collapse. *Journal of Economic Dynamics and Control* 18(5):991–1010, DOI 10.1016/0165-1889(94)90042-6
- Cropper M (1976) Regulating activities with catastrophic environmental effects. *Journal of Environmental Economics and Management* 3(1):1–15, DOI 10.1016/0095-0696(76)90009-7

- Dumas P, Espagne E, Perrissin-Fabert B, Pottier A (2012) Comprehensive Description of RESPONSE. CIRED Working Paper 41-2012, CIRED
- Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD (eds) (2014) Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of the Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press
- Fleurbaey M, Zuber S (2015) Discounting, beyond Utilitarianism. *Economics: The Open-Access, Open-Assessment E-Journal* 9(2015-12):1–52
- Gaffin SR, O'Neill BC (1997) Population and global warming with and without co2 targets. *Population and Environment* 18(4):389–413
- Gjerde J, Grepperud S, Kverndokk S (1999) Optimal climate policy under the possibility of a catastrophe. *Resource and Energy Economics* 21(3-4):289–317, DOI 10.1016/S0928-7655(99)00006-8
- Holtug N (2001) On the value of coming into existence. *The Journal of Ethics* 5:361–384
- Hope C (2006) The marginal impact of CO2 from PAGE2002: An integrated assessment model incorporating the IPCC's five reasons for concern. *Integrated assessment* 6(1):19–56
- Hope C, Anderson J, Wenman P (1993) Policy analysis of the greenhouse effect - An application of the PAGE model. *Energy Policy* 21(3):327–338
- Hsiang SM, Burke M, Miguel E (2013) Quantifying the Influence of Climate on Human Conflict. *Science* 341(6151):1235367–1235367, DOI 10.1126/science.1235367
- Hurka T (1982) Average utilitarianisms. *Analysis* 42(2):65–69
- IIASA (2014) AR5 Scenario Database
- Jamieson D (2007) When utilitarians should be virtue theorists. *Utilitas* 19:160–183
- Kelly DL, Kolstad CD (2001) Malthus and Climate Change: Betting on a Stable Population. *Journal of Environmental Economics and Management* 41(2):135–161
- 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, Mundaca L (2014) Social, Economic, and Ethical Concepts and Methods. In: Edenhofer O, Pichs-Madruga R, Sokona Y, Farahani E, Kadner S, Seyboth K, Adler A, Baum I, Brunner S, Eickemeier P, Kriemann B, Savolainen J, Schlömer S, Stechow Cv, Zwickel T, Minx J (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
- Lawson N, Spears D (2018) Optimal population and exhaustible resource constraints. *Journal of Population Economics* 31(1):295–335
- Lemoine D, Traeger CP (2016) Economics of tipping the climate dominoes. *Nature Climate Change* 6(5):514–519, DOI 10.1038/nclimate2902
- Lenton TM, Held H, Kriegler E, Hall JW, Lucht W, Rahmstorf S, Schellnhuber HJ (2008) Tipping elements in the Earth's climate system. *Proceedings of*

- the National Academy of Sciences 105(6):1786–1793, DOI 10.1073/pnas.0705414105
- Lontzek TS, Cai Y, Judd KL, Lenton TM (2015) Stochastic integrated assessment of climate tipping points indicates the need for strict climate policy. *Nature Climate Change* 5(5):441–444, DOI 10.1038/nclimate2570
- McCarthy JJ, Canziani OF, Leary NA, Dokken DJ, White KS (eds) (2001) *Climate Change 2001: Impacts, Adaptation, and Vulnerability. Contribution of the Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press
- Millner A (2013) On welfare frameworks and catastrophic climate risks. *Journal of Environmental Economics and Management* 65(2):310–325, DOI 10.1016/j.jeem.2012.09.006
- Ng YK (1986) Social criteria for evaluating population change: An alternative to the Blackorby-Donaldson criterion. *Journal of Public Economics* 29(3):375–381
- Ng YK (1989) What should we do about future generations? impossibility of parfit’s theory x. *Economics and Philosophy* 5:235–253
- Nordhaus WD (1994) *Managing the global commons: the economics of climate change*. MIT Press, Cambridge, Mass.
- Nordhaus WD, Boyer J (2000) *Warming the world: economic models of global warming*. MIT Press, Cambridge, Mass
- Parfit D (1984) *Reasons and Persons*. Oxford University Press
- Peck SC, Teisberg TJ (1995) Optimal CO<sub>2</sub> control policy with stochastic losses from temperature rise. *Climatic Change* 31(1):19–34, DOI 10.1007/BF01092979
- Pottier A, Espagne E, Fabert BP, Dumas P (2015) The Comparative Impact of Integrated Assessment Models’ Structures on Optimal Mitigation Policies. *Environmental Modeling & Assessment* 20(5):453–473, DOI 10.1007/s10666-015-9443-9
- 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(6):656–673, DOI 10.1016/j.polgeo.2007.05.001
- Scovronick N, Budolfson MB, Dennig F, Fleurbaey M, Siebert A, Socolow RH, Spears D, Wagner F (2017) Impact of population growth and population ethics on climate change mitigation policy. *Proceedings of the National Academy of Sciences* 114(46):12338–12343
- Sherwood SC, Huber M (2010) An adaptability limit to climate change due to heat stress. *Proceedings of the National Academy of Sciences* 107(21):9552–9555, DOI 10.1073/pnas.0913352107
- Shue H (2014) *Climate Justice. Vulnerability and Protection*. Oxford University Press
- Sidgwick H (1907) *The Methods of Ethics*. McMillan
- Steffen W, Rockström J, Richardson K, Lenton TM, Folke C, Liverman D, Summerhayes CP, Barnosky AD, Cornell SE, Crucifix M, Donges JF, Fetzer I, Lade SJ, Scheffer M, Winkelmann R, Schellnhuber HJ (2018) Trajecto-

- ries of the earth system in the anthropocene. *Proceedings of the National Academy of Sciences* 115(33):8252–8259
- Stern NH (2007) *The economics of climate change: the Stern review*. Cambridge University Press, Cambridge, UK ; New York
- Tol RSJ (1997) On the optimal control of carbon dioxide emissions: an application of FUND. *Environmental Modeling & Assessment* 2(3):151–163, DOI 10.1023/A:1019017529030
- Tsur Y, Zemel A (2009) Endogenous Discounting and Climate Policy. *Environmental and Resource Economics* 44(4):507
- Wagner G, Weitzman ML (2015) *Climate Shock – The Economic Consequences of a Hotter Planet*. Princeton University Press
- Weitzman ML (2012) GHG Targets as Insurance Against Catastrophic Climate Damages. *Journal of Public Economic Theory* 14(2):221–244, DOI 10.1111/j.1467-9779.2011.01539.x