Optimal climate policy under tipping risk and temporal risk aversion

Romain Fillon[1a,b], Céline Guivarch[b], Nicolas Taconet[b,c,d]

[a] *Université Paris-Saclay, INRAE, AgroParisTech, Paris-Saclay Applied Economics, F-91120, Palaiseau, France*
[b] *CIRED, AgroParisTech, CIRAD, CNRS, EHESS, Ecole des Ponts ParisTech, Univ Paris-Saclay, UMR 8568. CIRED, 94736 Nogent-sur-Marne, France*
[c] *Potsdam Institute for Climate Impact research - Member of the Leibniz Association, Potsdam, Germany*
[d] *Technical University Berlin, Germany*

Abstract

We investigate the implications of absolute risk aversion with respect to intertemporal utility, i.e. *temporal* risk aversion, in the presence of a stylized climate tipping risk affecting productivity irreversibly. Optimal climate policy is more stringent under temporal risk aversion, in order to reduce all present and future probabilities of crossing the tipping point and avoid a situation where all generations are badly off. Temporal risk aversion implies a 30% increase in the social cost of carbon (SCC) under our benchmark calibration and for a 10% irreversible increase in the level of economic damage from climate change. The optimal SCC under temporal risk aversion increases sharply with the level of damage brought by a potential tipping point.

Keywords : stochastic climate-economy modelling, risk-sensitive recursive preferences, environmental policy, risk aversion. **JEL classification** : D61, D63, D71, D81, Q54, Q58.

[1] Corresponding author: rfillon@protonmail.com

1. Introduction

When it comes to decision-making, risk is all around. But the concept is equivocal. First, it can refer to a univariate risk bearing on a single prospect. The seminal work from Pratt [42] and Arrow [6] introduced this risk into the analysis of decision-making through univariate measures of absolute and relative risk aversion within expected utility theory. A substantial body of literature has developed to generalise these measures of risk aversion to mutivariate risks [31]. A risk-averse portfolio manager does not sum the risk of each asset, but considers the aggregate risk bearing on the portfolio. Indeed, a positive correlation between these asset risks increases the aggregate risk. In intertemporal settings, the absolute risk aversion with respect to aggregate intertemporal risk is called the *temporal risk aversion* [14]. The standard discounted expected utility model assumes temporal risk-neutrality [2]. This assumption has large implications as it implies that the decision-maker has no preference on the correlation between individual risks. Introducing absolute risk aversion with respect to intertemporal utility, i.e. temporal risk aversion, on the other hand, allows to consider risk bearing on aggregate intertemporal utility. It can explain agent's intertemporal decisions [11, 15]. It is also of interest from a normative point of view, to define optimal policies in risky social situations that involve several successive generations whose welfare is correlated.

A prominent example of intertemporal social risk management is climate policy-making. A major concern of climate policy-making is the possibility of non-linearities such as tipping points in the climate system. Once some thresholds for greenhouse gas concentrations in the atmosphere are exceeded, the state of the climate system could be radically and irreversibly altered. Tipping elements with significant economic implications have been identified, including the slowdown of the Atlantic Meridional Overturning Circulation, the West Antarctic ice sheet disintegration, the Amazon rainforest dieback, or the Greenland ice sheet disintegration [5]. In the states of the world where the tipping point occurs, the welfare of all subsequent generations is affected by this qualitative regime change. Consequently, considering absolute risk aversion with respect to intertemporal utility becomes imperative due to the substantial impact on intertemporal welfare.

Temporal risk aversion can be interpreted as positive intertemporal correlation aversion [44], as positive intertemporal correlation implies a larger aggregate risk over intertemporal utility. A temporally risk-averse social planner prefers the welfare of different generations to be negatively or not correlated rather than positively correlated, in order to lower the risk on the aggregate outcome. In other words, the temporally risk averse social planner would be ready to give up some social welfare to prevent a situation where the tipping point is crossed and all subsequent generations are badly off. Thus, this social diversification strategy is appealing from a normative point of view when facing irreversible catastrophic tipping risks. Also, from a positive point of view, empirical elicitations of individual preferences suggest that individual agents might exhibit positive correlation aversion [24, 3, 27, 45, 34].

In this article, we investigate how temporal risk aversion may affect optimal climate policy. We analyze both analytically and numerically why, how and by how much two social planners, i.e. a temporally risk-neutral and a temporally risk-averse planner, differ in their optimal policy under risk. We focus on a specific type of risk: a climate tipping risk. We use a dynamic stochastic climate-economy model [28, 48] and extend it to an alternative social welfare function which allows the analysis of temporal risk aversion: the risk-sensitive preferences axiomatized in Hansen and Sargent [29]. By comparing optimal climate policies under risk-sensitive preferences with those under the standard additive form of expected discounted utility, which assumes temporal risk-neutrality, we shed light on the implications of temporal risk aversion for policy design.

We find that, in the presence of a tipping risk, climate policy is more stringent under risk-sensitive preferences. The social planner under risk-sensitive preferences is willing to sacrifice more today to reduce all present and future probabilities of crossing the tipping point to avoid a situation of low overall intertemporal utility level. The difference in optimal climate policy between the two planners increases more than proportionally to the increase in the possible shock or in the temporal risk aversion. Under our benchmark calibration, a change from additive to risk-sensitive preferences implies a 30% increase in the social cost of carbon (SCC) for a 10% irreversible increase in the damage factor. Switching from additive to risk-sensitive preferences under a 10% possible shock is equivalent to a 5 percentage points increase in the shock if we keep additive preferences. The difference between the two social choice criteria increases steeply with risk. Furthermore, other things being equal, a 50% decrease in pure time preference (from 1.5% to 1% yearly) is needed to obtain the same optimal policy under additive preferences as under risk-sensitive preferences for a 10% tipping risk and under our benchmark calibration. Thus, a change in the structure of the social welfare function can be directly compared to a change in the value of some parameters that have been highly debated. Finally, we use an analytical decomposition of our optimal policy program to derive the key channels through which a tipping risk affects optimal policy under both social welfare functions.

Our work contributes to the literature aiming to enhance the integration of different types of risk, particularly the risk of climate tipping points [36, 52, 19], into stochastic integrated assessment models (IAM). The first integrated climate-economy models were deterministic, e.g. Nordhaus [40]. These models did not allow for a proper consideration of risk and uncertainty in planner's decisions, even when Monte Carlo analyses were conducted [21]. In parallel, contributions to modeling endogenous catastrophic environmental risk were mostly stylized [20, 50, 14]. In particular, these models are based on the assumption that welfare after the catastrophic event is exogenous and independent of the planner's actions. Tipping points are less extreme than catastrophes after which production and consumption would be exogenous and independent of the planner's decisions. Indeed, these are ecological regime shifts with large economic consequences rather than complete economic or institutional collapses. These events are also different from reversible extreme events that occur as one-off catastrophes along a smoothly evolving climate regime with fluctuations, traditionally modelled with Poisson and Wiener processes in the macroeconomics literature on disasters, e.g. in Bretschger and Vinogradova [17]. Departing from the assumption of a geometric Brownian motion with rare and reversible catastrophic events, we study irreversible regime changes. This modelling approach has counterparts in the real business cycles literature studying markov switching rational expectations models with Bayesian learning, e.g. in Bullard and Singh [18].

Our contribution confronts the standard discounted expected utility model with an alternative criterion: a risk-sensitive criterion steming from social choice theory and axiomatized in Bommier et al. [13]. Exploration of alternative social choice criteria under endogenous climate change was undertaken to introduce relative risk aversion under Epstein-Zin-Weil preferences [8, 51], a robust control penalty [46] and ambiguity aversion under isoelastic preferences in a setting with uncertainty [37]. In comparison with EZW preferences, risk-sensitive preferences are the only recursive preferences axiomatized by Kreps and Porteus [33] that admit a separation of risk and intertemporal attitudes, while being monotone [13]. This desirable normative property ensures that a more risk-averse planner consistently prioritizes risk reduction. Those preferences can be defined through the following recursion [29, 13]:

$$V_t = \begin{cases} (1-\beta)\, u_t + \beta \mathbb{E}[V_{t+1}] & \text{if} \quad \epsilon = 0 \\ u_t - \frac{\beta}{\epsilon} \ln[\mathbb{E}(\exp[-\epsilon V_{t+1}])] & \text{if} \quad \epsilon \neq 0 \end{cases} \tag{1}$$

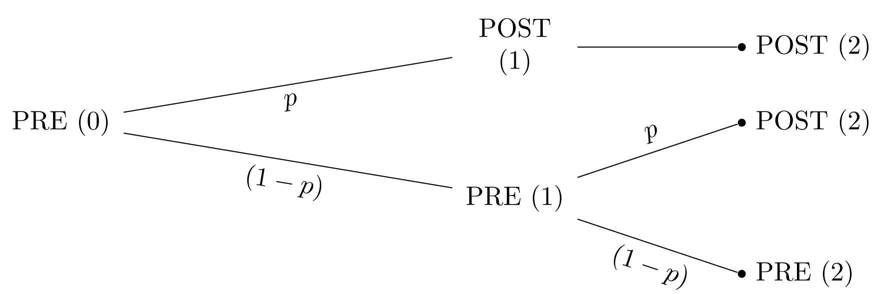
with u_t the instantaneous utility at time t, β a discount factor derived from pure time preference and ϵ the temporal risk aversion. We hereafter use the denomination of risk-sensitive preferences only for those stationary preferences for which the social planner is at least as risk averse ($\epsilon > 0$) as a standard planner with additive preferences. Cases where the social planner is temporally risk-seeking ($\epsilon < 0$) are not discussed because of potential nonconvexity issues [15]. A temporally risk-seeking planner would choose a max-max strategy and positive correlation between the social gambles. If $\epsilon = 0$, then the social planner is temporally risk-neutral, which comes down to the additive form.

Firstly, we present our modelling approach (section 2): a dynamic stochastic climate-economy integrated model with a stylized tipping risk, in which we compare two alternative social welfare functions. Then, we discuss analytically how temporal risk aversion affects optimal policy under a tipping risk (section 3). Finally, we quantify numerically the differences between the two social welfare functions under a tipping risk (section 4).

2. A dynamic climate-economy stochastic model

2.1. A simple illustration

Firstly, we illustrate the significance of temporal risk aversion in the analysis of climate tipping risks using a simplified scenario. Consider three consecutive time periods, representing distinct generations. Two climate regimes exist: pre-tipping (PRE) and post-tipping (POST), each associated with different levels of economic damage. Each generation t can either be in a high (PRE) or a low (POST) welfare regime, described by the variable u_i^t, $i \in (pre, post)$, $t \in 1, 2$. We assume that instantaneous welfare function in each potential situation is the same for both generations, i.e. $u_i^1 = u_i^2$.



We assume away time discounting and assume that the social planner has no preference on the order of the attributes and no preference for early resolution of uncertainty. Under the conditions listed above, a temporally risk-neutral social planner would be indifferent between the two following lotteries [10]:

$$\begin{cases} (u_1^{post}, u_2^{post}) \text{ with probability } 1/3 \\ (u_1^{pre}, u_2^{post}) \text{ with probability } 1/3 \\ (u_1^{pre}, u_2^{pre}) \text{ with probability } 1/3 \end{cases} \sim \begin{cases} (u_1^{post}, u_2^{post}) \text{ with probability } 1/2 \\ (u_1^{pre}, u_2^{pre}) \text{ with probability } 1/2 \end{cases} \quad (2)$$

A social planner under additive preferences would be indifferent between the two social lotteries A and B as the additive form assumes temporal risk-neutrality, while a temporally risk-averse social planner has a preference for lottery A. In other words, a temporally risk-averse social planner is willing to pay a temporal risk premium to hedge risks across generations and reduce the probability of complete failure across all generations.

In addition to positive intertemporal correlation aversion, temporal risk aversion bears preference for catastrophe avoidance[2] [14], i.e. preference for a mean-preserving contraction in the distribution of catastrophic risks. The preference for catastrophe avoidance is highly debated in the literature for two main reasons. First, it is not clear that individual agents are catastrophe-averse [43]. Furthermore, preference for catastrophe avoidance may be seen as unethical as a catastrophe-averse planner prefers to concentrate risk on a single generation rather than spreading it evenly [25]. Consequently, Fleurbaey [26] highlights that catastrophe aversion might be appealing only if the catastrophe has a multiplier effect through externalities in society. The possible nonconvexities in the human-environment system, enhanced by ecological thresholds like climate tipping points, do have this multiplier property. Indeed, in the states of the world where the tipping point occurs, the regime change is *irreversible* and has an impact on all future generations.

We have described in a simple illustration the importance of temporal risk aversion in risky intertemporal settings. We now present a full-fledged stochastic climate-economy model to analyse and quantify the importance of temporal risk aversion for the definition of optimal climate policy under a tipping risk.

2.2. The model

A climate-economy integrated assessment model aims to study the interactions between the economy and the climate system. We introduce a simple growth model *à la Ramsey*, add a stylized representation of the climate dynamics and an endogenous stochastic tipping point in the climate system. We build on Guivarch and Pottier [28] and Taconet et al. [48], update the economic dynamics to match DICE-2016 [39] and use an alternative social welfare function.

[2] If the social planner is temporally risk-seeking ($\epsilon < 0$), she favors risk equity, i.e. equalizing and spreading the risk among generations.

Economy In our global model, a single good is produced at each period t using two production factors, endogenous capital K_t and exogenous labour L_t, through a Cobb-Douglas production function $F(K_t, L_t) = A_t K_t^{\alpha} L_t^{1-\alpha}$ with exogenous Hicks-neutral technological change. The gross output $F(K_t, L_t)$ is affected by a damage factor $\Omega_t(T_t)$ that increases with global average temperature T_t. Net output Y_t is derived from the gross output net of damage: $Y_t = \Omega_t(T_t)F(K_t, L_t)$. Capital dynamics is determined by δ, the per-period capital depreciation, and s_t, the savings rate. It writes: $K_{t+1} - K_t = -\delta K_t + s_t Y_t$. Gross output induces emissions, which can be mitigated at a certain cost. The social planner trades off consumption C_t, mitigation costs (which represent a share Λ_t of Y_t), and investment: $C_t = Y_t(1 - \Lambda_t - s_t)$. The mitigation cost Λ_t depends on the abatement rate μ_t and on the cost of the abatement technology that decreases due to exogenous technical progress. The cost of the abatement technology is calibrated on Nordhaus [39] as other parameters of the economic module.

Climate We use a simple representation for the climate system with a linear formula linking temperature change to the stock of carbon emissions [23]. This approach avoids overestimating the delay between emissions and temperature rise. Indeed, the link between cumulative emissions and temperature has been shown to be almost independent of time and emissions pathways except for very high emission pathways [35] such as the RCP 8.5: it should thus hold for any reasonable optimal policy scenario. Emissions are derived from output: $E_t = \sigma_t Y_t(1 - \mu_t)$, where σ_t is the carbon content of production that decreases exogenously over time. Emissions increase carbon concentration in the atmosphere and there is no decay. Equation for temperature change is:

$T_t = \psi\left(CE_0 + \sum_{s=0}^{t} E_s\right) = \psi S_t$ where T_t is the global temperature increase (in comparison with the pre-industrial era) at time t, CE_0 is cumulated emissions up to the first period of the model, E_s the emissions at time s, S_t the carbon stock in the atmosphere at time t and ψ the transient climate response to cumulative carbon emissions (TCRE, $\psi = 1.65$°C per TtC, according to Masson-Delmotte et al. [38]).

Tipping risk We model one stylized endogenous tipping point that may decrease the output via an increase in the damage factor affecting the productivity. The tipping point is endogenous as its probability of occurence is a function of global average surface temperature. If the tipping point is crossed, the damage factor Ω faces an irreversible J% increase. The pre-tipping damage function writes: $\Omega_1(T) = 1 - \pi T^2$. Once the tipping point is crossed, the damage increase by J% and the new damage function writes: $\Omega_2(T) = (1 - J)(1 - \pi T^2)$. The damage occurs with no delay. The probability of tipping is modeled with a uniform distribution between initial temperature increase with respect to pre-industrial era and an upper temperature threshold[3] to make as few assumptions as possible about the precise temperature at which a tipping event may occur. Along the path, this specification allows learnings from the bayesian policy-maker as she updates her beliefs on the location of the threshold in the state space and on the probability of tipping at each period. The key assumption from this specification of the potential tipping event is that there is no tipping risk if the temperature is stabilized [36]. At each period t, the tipping point is

[3]The lower bound is the 2015 current excess temperature in comparison with the preindustrial era (0.87°C in 2015). The upper bound is set to 5.7°C according to the upper bound of the temperature increase reached in 2100 in RCP 8.5 [5]. See Appendix E.1 for a sensitivity analysis.

crossed with probability h_t:

$$h_t(T_t, T_{t-1}) = \begin{cases} \frac{T_t - T_{t-1}}{T_{max} - T_{t-1}} & \text{if} \quad T_t < T_{max} \\ 1 & \text{if} \quad T_t \geq T_{max} \end{cases} \tag{3}$$

We have presented above a stochastic model with a stylized tipping point. A second step is to use a social welfare function that allows the study of temporal risk aversion. We present this function in more depth below. To allow comparison with previous literature, we compare how two forms of social preferences behave in a risky intertemporal social setting. The first form is the additive one. The second form is the one of risk-sensitive preferences with positive temporal risk aversion.

2.3. Social preferences

In our model, we write two Bellman equations for the two possible situations, pre- and post-tipping, under the additive and the risk-sensitive social welfare functions, as welfare is affected by a J% increase in the damage factor once the tipping point is crossed. If the tipping point is crossed, the Bellman equation writes the same way for the two programs. The two social welfare functions yield the same policy in the risk-free post-tipping situation: temporal risk aversion plays no role in these risk-free situations, whatever its level. Once the tipping is crossed, all risk is solved: the tipping risk is the sole risk we study here. The state variables of our optimization program are $x_t = (S_t, K_t)$ respectively the cumulative emissions stock and the capital stock at time t. The control variables are $y_t = (\mu_t, s_t)$, respectively the abatement rate and the savings rate at time t. The instantaneous utility function writes: $u_t(x_t, y_t) = C_t^{1-\eta}/(1-\eta)$ with η the elasticity of marginal utility.

Additive preferences Under additive preferences, once the tipping point is crossed, we have: $U_t^{post}(x_t, y_t) = \underset{y_t}{max} \; \left[u_t(x_t, y_t) + \beta U_{t+1}^{post}(x_{t+1})\right]$ under the constraints: $x_{t+1} = G(x_t, y_{t+1})$ and $y_t \in \Gamma(x_t)$, with Γ the space of possible (positive) values for the control variables and G a transfer function. If the tipping point has not been crossed yet at time t, then it may be crossed at time t+1 with probability h_{t+1} or the world can stay in a pre-tipping situation with a probability $(1 - h_{t+1})$. The pre-tipping Bellman equation under additive preferences and under the same constraints as above writes:

$$U_t^{pre}(x_t, y_t) = \underset{y_t}{max} \ \left[u_t(x_t, y_t) + \beta[(1 - h_{t+1})U_{t+1}^{pre}(x_{t+1}) + h_{t+1}U_{t+1}^{post}(x_{t+1})]\right] \quad (4)$$

Risk-sensitive preferences Once the tipping point is crossed, the program under risk-sensitive preferences reduces to the additive one. If $\epsilon = 0$, the program under risk-sensitive preferences reduces to the additive one. Finally, it should be noted that $V^{post} = U^{post}$. The Bellman equation under the same constraints in the pre-tipping situation writes:

$$V_t^{pre}(x_t, y_t) = \underset{y_t}{max} \ \left(u_t(x_t, y_t) - \frac{\beta}{\epsilon} \ln\left[(1 - h_{t+1}) \exp(-\epsilon[V_{t+1}^{pre}(x_{t+1})]) + h_{t+1} \exp(-\epsilon[V_{t+1}^{post}(x_{t+1})])\right]\right) \quad (5)$$

2.4. Comparison with alternative social preferences

We compare the additive expected utility model to risk-sensitive preferences in order to study temporal risk aversion. Two main other frameworks have been used to study risk aversion under endogenous catastrophic climate change: the Epstein-Zin-Weil framework (hereafter, EZW) and the multiplicative preferences.

2.4.1. Epstein-Zin-Weil preferences

EZW preferences have been widely used in risky intertemporal settings to discuss optimal policy, e.g. in Cai and Lontzek [19], because of their flexibility, which allows to disentangle preference over time and preference over states of the world. We depart from it for two main reasons.

The first reason is that these preferences are monotone with respect to first-order stochastic dominance[4] [13] only in the limit cases where relative risk aversion equals the inverse of the elasticity of intertemporal substitution (they reduce to the standard additive model) or when the elasticity of intertemporal substitution equals one (EZW preferences are then risk-sensitive). If EZW preferences are well ordered in terms of risk aversion 'in the large' (willingness to pay to eliminate all risks), those preferences are not well ordered in terms of risk aversion 'in the small' (willingness to pay for marginal risk reductions). Thus, a social planner under EZW preferences might choose dominated strategies in social settings where it is not possible or optimal to eliminate all risk which may precisely be the case with climate change. In particular, it has been shown in the theoretical and applied literature that this non-monotonicity can lead to two types of counter-intuitive behaviours. On the one hand, the EZW agent can make more precautionary choices than necessary, choosing to build up more precautionary savings in a risky situation than the savings chosen in the worst state of the world that could occur under this risk if it happened deterministically [13]. This leads to a more extreme behavior than a max-min approach. On the other hand, the role of risk aversion could be non-monotone, meaning that for a higher relative risk aversion and the same risk, the planner can be less precautionous [32, 15]. The fact that such dominated strategies can be chosen, even if not always, makes this criterion less appealing for the definition of the optimal policy. Unlike the EZW framework, risk-sensitive preferences are monotone with respect to first-order stochastic dominance, which means that dominated strategies are never chosen. In particular, in our setting, we show in annex Appendix C that the risk premium is always positive and increasing

[4] A social planner has preferences that respect first-order dominance if, for two lotteries A and B with A dominating B, she prefers A to B regardless of her utility function, as long as it is weakly increasing. The lottery A dominates B if it gives more wealth than B realization by realization.

in the temporal risk aversion ϵ. When relative risk aversion is lower than the inverse of the elasticity of intertemporal substitution, EZW preferences show preference for late resolution of uncertainty and a negative risk premium, while risk-sensitive preferences exhibit preference for early resolution of uncertainty whenever $\epsilon > 0$. Risk-sensitive preferences thus allow a more rational social choice while preserving the flexibility and recursivity properties of the Kreps and Porteus [33] framework.

The second reason why we use risk-sensitive preferences rather than EZW preferences is that the coefficient of relative risk aversion studied in EZW preferences does not directly compare with the absolute risk aversion with respect to intertemporal utility studied under risk-sensitive preferences[5], as a reduction in relative risk does not always come with a reduction in aggregate risk [12]. A relative risk averse agent prefers to have non-extreme payoffs across states of the world within periods, while a temporally risk-averse planner prefers to have non-extreme payoffs across states of the world over the whole time horizon considered.

2.4.2. Multiplicative preferences

The second form are the multiplicative preferences [14] that rule out pure time preference so that different generations are not given different utility weights because they were born at different dates. Instead, we use an intermediate form

[5] Risk-sensitive preferences use a constant *absolute* risk aversion certainty equivalent, whereas EZW preferences use a constant *relative* risk aversion certainty equivalent [14]. When comparing temporal lotteries of consumption, constant absolute risk aversion has been seen as irrealistic because risk aversion is the same for all levels of wealth under this assumption. Here, the constant absolute risk aversion certainty equivalent is applied to distributions of utility levels rather than consumption levels. This assumption is made under risk-sensitive preferences as monotonicity implies that risk aversion is considered with respect to aggregate utility. Thus, in order to preserve history independence, constant asbolute risk aversion with respect to aggregate intertemporal risk ensures that the utility of the first periods does not impact social choice afterwards [13].

of risk-sensitive preferences that does not assume away time discounting for three reasons. Firstly, we do not include an extinction risk, so that without pure time preference, our undiscounted dynamic program would be too sensitive to the arbitrary terminal value and limit the comparability between the two programs. The second reason is that we want to analyze the sole role of temporal risk aversion on social choice rather than intertwining this questioning with the debate between discounted and undiscounted utilitarianism [47, 40]. The third reason is the comparability between additive and risk-sensitive preferences. Indeed, additive and risk-sensitive social planners have the same rankings over deterministic consumption paths regardless of the value of the temporal risk aversion ϵ. We can therefore simply vary ϵ within a reasonable value range and make comparisons between the two social choice criteria under risk for different values of ϵ.

We have characterized the additive and the risk-sensitive social welfare functions and explained how temporal risk aversion can be an important determinant of climate policy. We now assess analytically the impact of temporal risk aversion on optimal climate policy under a tipping risk.

3. How does temporal risk aversion affect optimal policy under a tipping risk?

Firstly, we derive analytically the impact of temporal risk aversion on the optimal policy under a tipping risk. We decompose the pre-tipping value functions (4) and (5) which incorporate the risk of tipping and analyze the case where a single state variable determines the chance of crossing the threshold. We focus solely on S_t, the cumulated stock of emissions at time t. As we are considering optimal climate policy, we focus on the abatement rate μ_t and derive the first-order condition of our policy programs. Our analytical decomposition is a two-step procedure. First, we decompose the immediate short-term effect on next-period welfare of a marginal variation in abatement rate departing from the optimum, following Lemoine and Traeger [36]. Then, we derive the complete long-term effect of a marginal variation in the cumulative emissions stock on all future probabilities of tipping. The decomposition is done for the additive and risk-sensitive preferences: thus, we can derive how the channels through which a tipping risk affects optimal policy under additive preferences adjust to temporal risk aversion, in both the short and long term.

From the first-order condition of our policy programs, we show that the tipping risk affects optimal policy through three short-term channels. The first channel, the marginal hazard effect *mhe*, measures the impact of the control variable on the immediate probability of tipping. The second channel, the differential welfare impact *dwi*, measures the differential impact of the control variable on welfare depending on the situation, i.e. pre- or post-tipping, and if the tipping point is crossed. The last channel, the marginal impact pre-tipping *mpre*, defines the decrease in next-period's welfare resulting from an increase in the abatement policy if the tipping point has not been crossed yet: possible future tipping points are included in this last channel. Removing all arguments that are independent of μ_t in equation (3), the value of the optimal policy program in the pre-tipping situation under additive preferences writes:

$$u_t[\mu_t^*] + \beta \underbrace{\left[h_{t+1}(\mu_t^*)U_{t+1}^{post}(\mu_t^*) + (1 - h_{t+1}(\mu_t^*))U_{t+1}^{pre}(\mu_t^*)\right]}_{U_{t+1}^{eff}} \tag{6}$$

The first term of equation (8) corresponds to the level of instantaneous utility at time t for an optimal choice of the control variable μ_t^*. The second term gives the expected welfare at time t+1 when there is a probability of tipping point under temporal risk neutrality and for an optimal choice of the control variable, scaled by the discount factor β. Varying μ_t gives us the immediate decomposition under additive preferences characterizing optimal policy: $u'_t = \beta(dwi^{add}_{t+1} + mhe^{add}_{t+1} + mpre^{add}_{t+1})$, with the following channels:

$$\begin{cases} mhe^{add}_{t+1} = \frac{\partial h_{t+1}}{\partial S_{t+1}}\frac{\partial S_{t+1}}{\partial \mu_t}(U^{pre}_{t+1} - U^{post}_{t+1}) \\ dwi^{add}_{t+1} = h_{t+1}\frac{\partial S_{t+1}}{\partial \mu_t}(\frac{\partial U^{pre}_{t+1}}{\partial S_{t+1}} - \frac{\partial U^{post}_{t+1}}{\partial S_{t+1}}) \\ mpre^{add}_{t+1} = -\frac{\partial U^{pre}_{t+1}}{\partial S_{t+1}}\frac{\partial S_{t+1}}{\partial \mu_t} \end{cases} \tag{7}$$

The risk-sensitive social planner maximizes at time t a utility function V_t which is linked to the random continuation utility V_{t+1} through the following recursion: $V_t = u_t + \beta\phi^{-1}(\mathbb{E}[\phi(V_{t+1})])$. The function ϕ writes $\phi(V) = (1 - \exp(-\epsilon V))/\epsilon$. It is increasing and strictly concave for any $\epsilon > 0$. The value of the optimal policy program in the pre-tipping situation under risk-sensitive preferences is:

$$u_t[\mu^*_t] + \beta\,\underbrace{\phi^{-1}\left[h_{t+1}(\mu^*_t)\phi(V^{post}_{t+1}(\mu^*_t)) + (1 - h_{t+1}(\mu^*_t))\phi(V^{pre}_{t+1}(\mu^*_t))\right]}_{V^{eff}_{t+1}} \tag{8}$$

The immediate decomposition under risk-sensitive preferences writes: $u'_t = \beta(dwi^{rs}_{t+1} + mhe^{rs}_{t+1} + mpre^{rs}_{t+1})$, with the following channels:

$$\begin{cases} mhe^{rs}_{t+1} = \frac{B_{t+1}}{\epsilon}\left(\frac{\partial h_{t+1}}{\partial S_{t+1}}\frac{\partial S_{t+1}}{\partial \mu_t}[exp(-\epsilon V^{post}_{t+1}) - exp(-\epsilon V^{pre}_{t+1})]\right) \\ dwi^{rs}_{t+1} = B_{t+1}\left(h_{t+1}\frac{\partial S_{t+1}}{\partial \mu_t}\left[\frac{\partial V^{pre}_{t+1}}{\partial S_{t+1}}exp(-\epsilon V^{pre}_{t+1}) - \frac{\partial V^{post}_{t+1}}{\partial S_{t+1}}exp(-\epsilon V^{post}_{t+1})\right]\right) \\ mpre^{rs}_{t+1} = -B_{t+1}\left(\frac{\partial V^{pre}_{t+1}}{\partial S_{t+1}}\frac{\partial S_{t+1}}{\partial \mu_t}exp(-\epsilon V^{pre}_{t+1})\right) \\ \text{with } B_{t+1} = \left((1 - h_{t+1})exp(-\epsilon V^{pre}_{t+1}) + h_{t+1}exp(-\epsilon V^{post}_{t+1})\right)^{-1} \end{cases} \tag{9}$$

We highlight how temporal risk aversion implies an adjustment on these channels in comparison with additive temporally risk-neutral preferences. We extend the reasoning of Lemoine and Traeger [37] under uncertainty and ambiguity aversion to a related setting with risk and risk-sensitive preferences[6] and

[6]They use an isoelastic function for the transformation with uncertainty aversion in a setting with an ambiguous tipping point. The equivalent of risk-sensitive preferences in an

use their general approximations for the adjustments on the channels implied by a concave transformation of the additive social welfare function under a tipping risk. The complete procedure is depicted in Appendix A. The measure of absolute temporal risk aversion $\frac{-\phi''}{\phi'}\Big|_{V^{eff}}$ is equal to ϵ. We adjust the temporally risk-neutral marginal hazard effect channel mhe^{add} obtained from additive preferences to find the risk-sensitive marginal hazard effect mhe^{rs}:

$$mhe^{rs} \approx mhe^{add}\left[1+\epsilon(V^{eff}-\frac{V^{pre}+V^{post}}{2})\right] \tag{10a}$$

where V^{post} is the continuation value if the tipping point has already been crossed, V^{pre} the continuation value if the tipping point has not been crossed yet and V^{eff} the random continuation value for an optimal choice of the policy variable. The amplitude and the sign of the adjustment can not be derived analytically. Indeed, an increase in temporal risk aversion ϵ is counter-balanced by its negative impact on V^{eff} as V^{eff} is decreasing in ϵ. In comparison with the arithmetic mean $(V^{pre}+V^{post})/2$, the two possible regimes in V^{eff} are weighted by the probability of (not) tipping, lower (higher) than one half in any optimal policy paths considered here. We thus expect the marginal hazard effect to be increasing with ϵ in our setting. The marginal hazard effet, depicting the marginal impact of a marginal increase in abatement on the immediate probability of tipping, relates to the social value of catastrophic risk reduction [14] and the VSL-like parameter of Weitzman [53]. This channel is associated with self-protection in Lemoine and Traeger [36].

We then adjust[7] the temporally risk-neutral differential welfare impact dwi^{add} to obtain the risk-sensitive differential welfare impact dwi^{rs}. This channel is de-

uncertain setting would be the multiplier criterion [13].

[7]Taken from [37], the approximation holds for a low shock.

picted as self-insurance in Lemoine and Traeger [36]. The adjustment writes:

$$dwi^{rs} \approx dwi^{add} + \epsilon h \left[(V^{eff} - V^{pre})(\frac{\partial V^{pre}}{\partial \mu}) - (V^{eff} - V^{post})(\frac{\partial V^{post}}{\partial \mu}) \right] \tag{10b}$$

Similarly, the sign of the adjustment of temporal risk aversion on the risk-neutral DWI cannot be determined analytically. An increase in the temporal risk aversion ϵ decreases V^{eff} and both terms in the bracket, so that the overall sign depends on the relative level of the marginal welfare impact of the change in policy variable in the pre-threshold and the post-threshold worlds as in the temporally risk-neutral case. The adjustment decreases with the probability of tipping. We expect this channel and the adjustment to be negligible. Indeed, they depend on the value and the trajectory of the tipping probability with respect to ϵ. But the larger ϵ is, the lower the probability of tipping, because optimal policy under large temporal risk aversion is expected to be stricter. In our specification as in Lemoine and Traeger [36, 37], the *dwi* might be completely overwhelmed by the *mhe*.

One can finally adjust the last channel: the direct impact of the change in policy variable on the welfare if one stays in a pre-tipping situation in the next period:

$$mpre^{rs} = mpre^{add} \frac{\phi'(V^{pre})}{\phi'(V^{eff})} \tag{10c}$$

The adjustment implied by temporal risk aversion is the relative slope of the transformed continuation value if we stay in a pre-tipping situation on the slope of the transformed random continuation value. The size of the adjustment depends on the concavity of ϕ, i.e., the strength of temporal risk aversion ϵ. This term is equal to one when there is no tipping risk, i.e. if the temperature is stabilized, and goes to 0 if the probability of tipping h increases. The adjustment implied by temporal risk aversion decreases *mpre* unambiguously as $V_{pre} > V_{eff}$.

We have focused on the immediate impact of a marginal variation of the policy variable around the optimum and identified the channels through which the tipping risk affect next-period welfare under additive and risk-sensitive preferences. So far, we have only analyzed the immediate channels (*mhe* and *dwi*) and left all future impacts of a marginal change in the policy variable in the pre-tipping continuation value included in *mpre* as in Lemoine and Traeger [37]. Indeed, today's emissions also affect all future probabilities of triggering the tipping point. In order to recover the full impact of temporal risk aversion on the optimal policy under a tipping risk, we need to decompose further this *mpre* channel. We do not focus on the marginal impact of an increase in a control variable (i.e. the abatement rate), but on the marginal impact on the pre-tipping value function of a marginal increase in a state variable (the concentration stock S). As we assume that there is no decay, a marginal increase in the concentration stock can be analyzed as a marginal increase in carbon emissions. As in Jensen and Traeger [30], we assume that the dynamic system is well-defined so that the shadow value of the carbon concentration increase $\partial V^{pre}/\partial S$ grows sufficiently slowly along the optimal path to make the limit approach zero over our large time horizon. We can advance the derivative of our pre-tipping value function with respect to emissions by one period and reinsert it in itself:

$$\frac{\partial V_t^{pre}}{\partial S_t} = u'_t - \beta \left(mhe^{rs}_{t+1} + dwi^{rs}_{t+1} - B_{t+1} exp(-\epsilon V^{pre}_{t+1}) \left[u'_{t+1} - \beta(mhe^{rs}_{t+2} + dwi^{rs}_{t+2} - B_{t+2} exp[-\epsilon V^{pre}_{t+2}] \frac{\partial V^{pre}_{t+2}}{\partial S_{t+2}}) \right] \right) \tag{11}$$

Iterating the procedure eventually yields a general expression of the marginal impact of a marginal increase in carbon emissions on all present and future periods. The complete decomposition under risk-sensitive preferences writes:

$$\frac{\partial V_t^{pre}}{\partial S_t} = u'_t - \beta[mhe^{rs}_{t+1} + dwi^{rs}_{t+1}] + \sum_{i=t+1}^{\infty} \beta^{i-t} \left(\prod_{k=t+1}^{i} \underbrace{\frac{\phi'(V_k^{pre})}{\phi'(V_k^{eff})}}_{adjustment(mpre)} \right) (u'_i - \beta[mhe^{rs}_{i+1} + dwi^{rs}_{i+1}]) \tag{12}$$

The *mpre* channel of the immediate decomposition disappears. To differentiate them from the immediate decomposition terms, the full decomposition terms are in capital letters. The complete decomposition $\partial V_t^{pre}/\partial S_t = U_t' - MHE_t^{rs} - DWI_t^{rs}$ now includes all present and future effects:

$$\begin{cases} U_t' = u_t' + \sum_{i=t+1}^{\infty} \beta^{i-t} \left(\prod_{k=t+1}^{i} \frac{\phi'(V_k^{pre})}{\phi'(V_k^{eff})} \right) u_i' \\ MHE_t^{rs} = \beta mhe_{t+1}^{rs} + \sum_{i=t+1}^{\infty} \beta^{i-t+1} \left(\prod_{k=t+1}^{i} \frac{\phi'(V_k^{pre})}{\phi'(V_k^{eff})} \right) mhe_{i+1}^{rs} \\ DWI_t^{rs} = \beta dwi_{t+1}^{rs} + \sum_{i=t+1}^{\infty} \beta^{i-t+1} \left(\prod_{k=t+1}^{i} \frac{\phi'(V_k^{pre})}{\phi'(V_k^{eff})} \right) dwi_{i+1}^{rs} \end{cases} \tag{13}$$

The complete MHE^{rs} and DWI^{rs} depend on the sign and amplitude of all the present and future immediate mhe^{rs} and dwi^{rs}, and all future effects are scaled by the discount factor and the positive adjustment implied by temporal risk aversion. We have described analytically how temporal risk aversion changes the various channels through which a tipping risk affects a decision-maker, both short and long term. We assess numerically the impact of temporal risk aversion in a dynamic climate-economy stochastic model under a tipping risk and quantify the different channels depicted.

4. A numerical investigation

4.1. Calibration

We use the same specifications for the macroeconomic model as Nordhaus [39]. We use typical ranges of possible values for the key parameters. The pure rate of time preference ρ is 1.5% [39]. The marginal utility parameter η is set to 1.5 with a sensitivity analysis from 0.5 to 2.5. We explore a large range for the shock J, ranging from 0 to 10% as explored in van der Ploeg and de Zeeuw [51], Cai and Lontzek [19] and van der Ploeg and de Zeeuw [52].

Social planners under additive and risk-sensitive preferences have the same ordering over deterministic consumption paths[8]. Thus, we can make compar-

[8] On the contrary, this is not the case for all values of ϵ under multiplicative preferences that

isons between the two social choice criteria under risk for different values of ϵ. We look for a range of plausible values for this parameter and a benchmark value within it to set a default value and perform a sensitivity analysis. The range of values used in the literature is large. Anderson [4] uses 0.1, 1 and 2 to study the dynamics of optimal Pareto allocations of risk-sensitive agents. When studying precautionary savings, Bommier et al. [13] explore large values ranging from 0 to 4, and Bommier and Le Grand [15] explore very large values, up to 100. In order to reduce the plausible range, we use the fact that, when the elasticity of intertemporal substitution is set to one, the EZW preferences are risk-sensitive preferences [49]. Indeed, risk-sensitive and EZW preferences are special cases of the more general family of recursive Kreps and Porteus [33] preferences. An analytical relation between the temporal risk aversion on the one hand and pure time preference ρ and relative risk aversion χ of EZW preferences on the other hand can thus be formulated in this precise case: $\epsilon = -(1-\beta)(1-\chi)$ with χ the coefficient of relative risk aversion with respect to atemporal wealth gambles, and β the discount rate. Following the IAM literature calibration for χ [1, 19], we use $\chi = 10$ as a benchmark and run a sensitivity analysis around this value. In our benchmark case, with $\chi = 10$ and $\rho = 1.5\%$ yearly, we have $\epsilon = 0.133$. A low $\chi = 1.1$ would yield $\epsilon = 0.0015$ while a large $\chi = 20$ would yield $\epsilon = 0.3$. The lower the pure time preference, the lower the difference between additive and risk-sensitive preferences [14]. Our benchmark measure may not be adapted to social settings : a welfare-maximizing social planner might be more temporally risk averse than individuals when a catastrophic and irreversible risk bears on all future generations. In an empirical elicitation of the aversion towards

are undiscounted ($\rho = 0$). Thus, Bommier et al. [14] have to rely on a specific calibration of ϵ so that additive and multiplicative preferences yield the same discount rate and are comparable. The calibration of ϵ under multiplicative preferences depends on the form of the instantaneous utility, the level of pure time preference and the post-tipping exogenous consumption.

correlated risks in the context of donations to risky aid projects, Gangadharan et al. [27] find that individuals are more averse to correlated risks when they donate other people's money. This is an interesting line of thought for climate change, where the contemporary social planner has to choose an appropriate level of temporal risk aversion for other generations than the one he belongs to. Thus, our benchmark value for the temporal risk aversion is conservative and in the lower bound of those estimates.

4.2. A comparison of the two social welfare functions under risk

We derive the optimal climate policy under the two social welfare functions in a risky intertemporal social setting using dynamic programming. Details of the resolution are in Appendix B. A key instrument to compare optimal policy along the trajectory is the social cost of carbon (SCC) at initial time. For both specifications, it writes: $-\beta(\partial_S \mathbb{E}[W_1]|_{y_1} \,/\, \partial_C W_0|_{x_0,y_0*})$ with y_0* the optimal abatement and investment of the program at initial time given x_0 and β the discount rate derived from pure time preference. W, the value function, can be U (additive) or V (risk-sensitive). Figure 1 gives the absolute value of the SCC ($/tC) under additive and risk-sensitive preferences at initial time for a range of irreversible increase in the damage factor J and the ratio of the SCC under risk-sensitive preferences to the SCC under additive preferences for various ϵ and J.

We can draw three conclusions from the graphs above. First, optimal climate policy under risk-sensitive preferences is more stringent for any value of ϵ and J than under additive preferences. An increase in temporal risk aversion unambiguously leads to an increase in the social cost of carbon due to the monotonicity of risk-sensitive preferences. The second conclusion is that switching from additive to risk-sensitive preferences under a tipping risk induces a large change in optimal policy: the form of the social welfare function matters, as

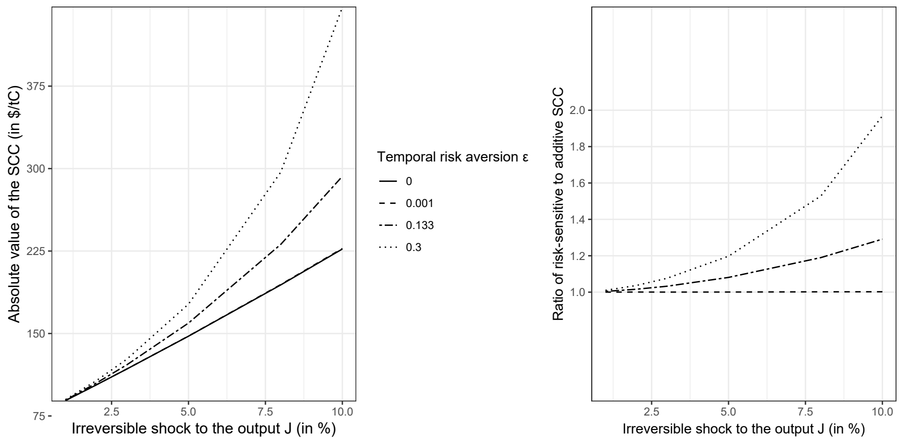


Figure 1: Absolute values of the additive and risk-sensitive SCC (in \$ per tC) at initial time (left) and ratio of the risk-sensitive SCC to the additive SCC (right) for various J and ϵ under our benchmark calibration ($\eta = 1.5$, $\rho = 1.5\%$). The curves overlap for the two smallest values of ϵ in the left-hand graph. Same graphs with a wider range for J are given in Appendix E.1.

already highlighted in Bommier et al. [14] for more catastrophic collapses. For a tipping point inducing a 10% irreversible increase in the damage factor, the SCC under risk-sensitive preferences is 30% higher than under additive preferences under our benchmark $\epsilon = 0.133$: it goes from 227\$ per tC to 293\$ per tC. This difference is increasing with the size of the possible shock J: the larger the tipping risk, the larger the difference between the optimal policies. Finally, temporal risk aversion plays a key role: under risk-sensitive preferences, for the largest $\epsilon = 0.3$, the SCC at initial time is 2-times higher than under additive preferences for a 10% shock. The slope of the ratio of the risk-sensitive SCC to the additive SCC is also increasing with ϵ. Increasing temporal risk aversion increases unambiguously the weights attributed to the catastrophic states of the world where numerous generations are badly off with a low intertemporal utility level. We run a sensitivity analysis in Appendix E.3 to check if our result is not affected by the calibration of the inequality aversion parameter η. The ratio of the risk-sensitive to the additive SCC is increasing in the value of the inequality aversion η. The SCC under risk-sensitive preferences is larger than under additive preferences for any value of η explored here.

To illustrate the magnitude of the change in optimal climate policy arising from temporal risk aversion, we show how switching from additive to risk-sensitive preferences compares with changes in the value of some parameters under additive preferences. We focus on two parameters that have been subject to debates in the literature. On the one hand, we consider the rate of pure time preference ρ [47, 40]. On the other hand, we focus on the value of the economic damage generated by climate change [41], and more specifically by a climate tipping point. First, Figure 2 (left) shows how a change from additive to risk-sensitive preferences compares to a change in ρ under additive preferences. Switching from additive to risk-sensitive preferences under a 10% tipping risk and for our benchmark calibration of the temporal risk aversion ($\epsilon = 0.133$) is equivalent to a 50% decrease in the value of ρ under additive preferences. In other words, the optimal policy derived from risk-sensitive preferences for our benchmark calibration ($\epsilon = 0.133$, $\eta = 1.5$, $\rho = 1.5\%$) and under a 10% tipping risk is obtained under additive preferences when $\rho = 1\%$ other things being equal. Figure 2 (right) shows that it takes a 14% shock for the additive preferences to give the same SCC as for a 10% irreversible increase in the damage factor under risk-sensitive preferences. The difference between the two approaches becomes more pronounced as the level of risk intensifies.

The numerical estimation of the channels analytically depicted in section (3) provides an understanding of the channels through which a tipping risk affects a temporally risk-averse planner. In our analytical decomposition, we firstly derived the channels through which a marginal increase in the policy variable departing from the optimum affects welfare in the next period under a tipping

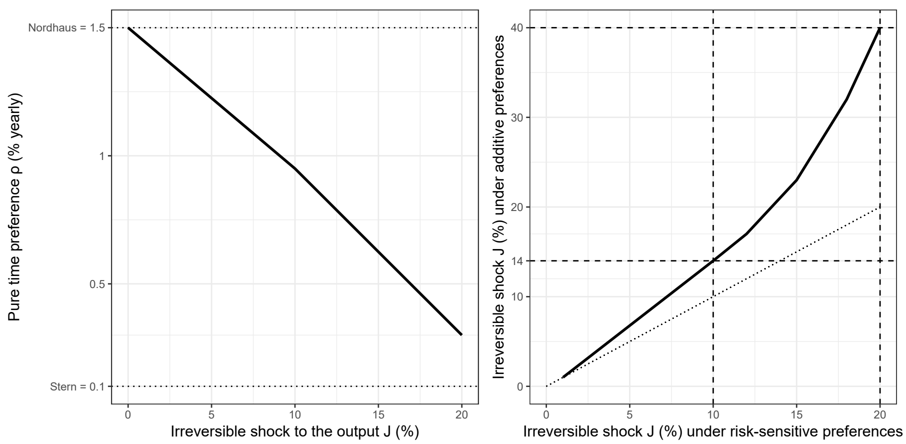


Figure 2: Equivalence in ρ (left) and J (right) needed to obtain the same SCC at initial time under additive and risk-sensitive preferences under our benchmark calibration. On the left, we represent the pure time preference (ρ) that is needed under additive preference (all else being equal) to match the risk-sensitive SCC for various J. On the right, we represent the irreversible shock J that is needed under additive preference (all else being equal) to match the risk-sensitive SCC for various J. The dotted line from the right-hand graph is the identity function.

risk and for a risk-sensitive planner. Thus, we have left aside all the future impacts on subsequent periods in this immediate decomposition, in particular the impact of the change in the policy variable on all future probabilities of crossing the threshold. Then, we have performed a full decomposition to take into account the impact of this change in policy on welfare in all future periods: this is the complete decomposition. We have shown that there are two channels through which tipping risk can influence optimal policy: the marginal hazard effect (immediate and complete) and the differential welfare impact (immediate and complete). We now run a numerical estimation of these channels to understand how temporal risk aversion may affect the channels through which the tipping risk affects the planner.

We can draw two conclusions from Figure 3. First, we see from our numerical estimation that the main channel is the marginal hazard effect. Indeed, the planner is ready to give up welfare in order to reduce all present (immediate)

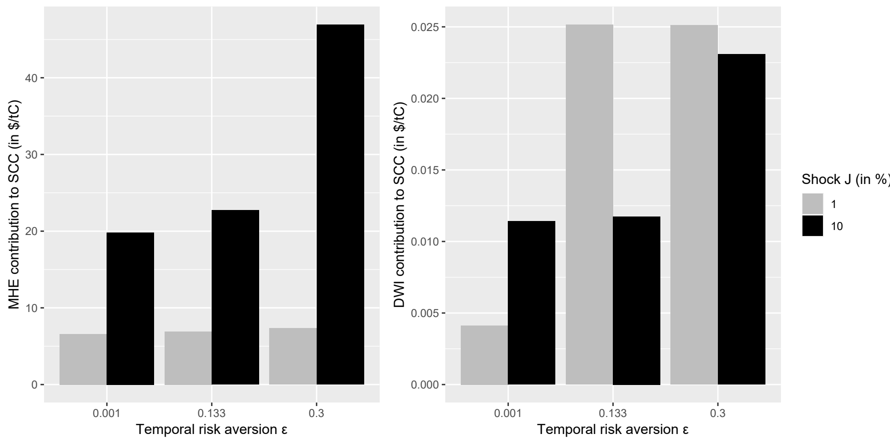


Figure 3: The graphs give the marginal contributions to the SCC at initial time (in \$) of the complete *MHE* (left) and the complete *DWI* (right) under risk-sensitive preferences for various J and ϵ. The scales for the two graphs are different because the two channels are several orders of magnitude apart. We give the same graphs for the immediate decomposition in Appendix E.4.

and future (complete) probabilities of crossing the threshold. This is partly due to the setting generally chosen in the literature to represent tipping points, as we do not model adaptation as an endogenous choice, which could for example decrease the level of damage J. Whether some form of adaptation can decrease the damage of such regime shifts remains uncertain, thus justifying its exclusion from our framework. The second conclusion from the numerical estimation is that the marginal hazard effect channel is increasing in the possible shock and in temporal risk aversion. A higher temporal risk aversion increases the stringency of the optimal policy, as highlighted above, and decreases even further the relative weight of the differential welfare impact in comparison with the marginal hazard effect.

5. Discussion

We study analytically and numerically in an integrated model with a stochastic tipping risk the role of absolute risk aversion with respect to intertemporal utility, i.e. temporal risk aversion. We compare the optimal climate policy arising from the expected discounted utility model to a risk-sensitive social welfare function exhibiting temporal risk aversion. A temporally risk-averse social planner maximising the welfare of successive generations prefers to lower the possibility of an irreversible damage bearing on all subsequent generations. In this sense, she adopts a social risk diversification strategy to hedge against potential environmental regime shifts.

First, while the two social welfare functions yield the same optimal climate policy in a risk-free setting, they differ once a tipping risk is introduced. Assumptions regarding the structure of the social welfare function appear as least as important as the debated value of some parameters in the expected utility model, such as the damage from a tipping point or the value of pure time preference. It should be emphasized that the assumption of temporal risk neutrality embedded in expected utility, while justifiable in risk-free models with smooth climate change, may not adequately capture possible non-linearities and abrupt regime changes in the climate system, which have been extensively documented in climate science [5]. Ignoring temporal risk aversion may lead to underestimating the severity of climate risks and result in more lenient climate policies. Therefore, considering temporal risk aversion becomes crucial when studying correlated intertemporal social risks.

Second, optimal policy under temporal risk aversion is more stringent than under temporal risk neutrality. The difference between the two social welfare functions increases more than proportionally to the increase in the shock J or the temporal risk aversion ϵ. For a 10% irreversible increase in the damage factor, the SCC under temporal risk aversion is 30% higher than the SCC under risk neutrality under our benchmark calibration. Our key take-away is that if one believes that major catastrophes bearing large multiplier effects such as irreversible regime shifts are possible, the social planner's aversion towards those risks bearing on intertemporal utility should be accounted for. On the other hand, if there is no such risk or if the possible damage is low, then we should stick to the additive model as it does not come with the ethical drawbacks catastrophe aversion bears.

The last conclusion is that optimal climate policy in our setting is mainly driven by the marginal hazard effect. The tipping risk affects optimal policy as the social planner wants to reduce all present and future probabilities of crossing the tipping point. This channel is increasing in the possible shock J and increasing in the temporal risk aversion. The risk-sensitive planner is willing to give up more wealth to avoid the catastrophic event.

Our analysis suffers three main limitations. Firstly, our model, although including a stochastic risk, suffers from the limitations often pointed out in integrated climate-economy models: the specification of the damage function, the exogenous technological change dynamics and the assumptions regarding future growth are for example uncertain. Secondly, our representation of tipping points is limited, as we focus on a single tipping point and do not consider various characteristics, such as their probability of occurrence, reversibility, abruptness, and time horizons. Additionally, our tipping probability is solely a function of global temperature, while other drivers, such as deforestation, can also contribute to tipping points. These limitations leave room for further research to provide a more comprehensive and precise representation of climate tipping points and damages. Lastly, our model assumes known probabilities for the tipping risk. Under ambiguity about the tipping points, a temporally risk averse planner might not prefer higher diversification [9].

Finally, we do not take any stance on what the *right* social welfare function is. This question remains open to scientific and public debates. In particular, the risk-sensitive social planner is not an expected utility maximizer. This may be defensible as one may 'accept the sure-thing principle for individual choice but not for social choice, since it seems reasonable for the individual to be concerned solely with final states while society is also interested in the process of choice' [22]. Temporal risk aversion helps us understand the specificity of the social choice issue climate change raises when it is considered not as a linear and smooth phenomenon, but as a phenomenon that can give rise to non-linearities and abrupt regime changes. A future research avenue could be to elicit the value that individuals would give to this parameter in the context of normative intergenerational social choice.

If our analysis is applied to a stylized climatic tipping risk, we believe that risk-sensitive preferences and temporal risk aversion might be used for the study of more standard smooth risks, as long as they are endogenous and correlated. Indeed, as risk-sensitive preferences exhibit preference for catastrophe avoidance when the social planner has temporal risk aversion, they comply with a weaker pareto axiom in comparison with additive preferences [16]: this axiom states that there is no difference between the social planner's and the individuals' preferences as long as uncorrelated risks are considered, but that some divergence may occur when correlated risks are at play. This intertemporal social choice criterion might thus bear critical implications for the management of correlated risks, for instance the large aggregate social risks due to potential ecological thresholds (e.g. biodiversity collapse).

Appendix A. Analytic decomposition details

We follow Lemoine and Traeger [37] to find an analytic approximation of how the risk-neutral channels adjust under temporal risk aversion. In addition, we disentangle mpre from dwi. Starting from expression of mhe^{add} and mhe^{rs} in equation (7) and (9), we write:

$$mhe^{rs}_{t+1} = \frac{\partial h_{t+1}}{\partial S_{t+1}} \frac{\partial S_{t+1}}{\partial \mu_t} \left(\frac{\phi(V^{pre}_{t+1}) - \phi(V^{post}_{t+1})}{\phi'(V^{eff}_{t+1})} \right) \tag{A.1}$$

Thus:

$$mhe_{rs} = mhe_{add} \underbrace{\frac{\phi(V^{pre}) - \phi(V^{post})}{\phi'(V^{eff})(V^{pre} - V^{post})}}_{adj_{mhe}} \tag{A.2}$$

and recall that $\phi(V) = (1 - \exp(-\epsilon V))/\epsilon$. A second order Taylor expansion for $\phi(V^i)$ around $\phi(V^{eff})$ gives: $\phi(V^i) \approx \phi(V^{eff}) + \phi'(V^{eff})[V^i - V^{eff}] + \frac{1}{2}\phi''(V^{eff})[V^i - V^{eff}]^2 + O([V^i - V^{eff}]^3)$. We have:

$$\phi(V^{pre}) - \phi(V^{post}) \approx \phi'(V^{eff})[\phi(V^{pre}) - \phi(V^{post})] + \frac{1}{2}\phi''(V^{eff})[(V^{pre})^2 - (V^{post})^2 + 2V^{eff}(V^{post} - V^{pre})] \tag{A.3}$$

And :

$$adj_{mhe} \approx 1 + \left. \frac{-\phi''}{\phi'} \right|_{V^{eff}} \left[V^{eff} - \frac{V^{pre} + V^{post}}{2} \right] \tag{A.4}$$

This yields our final expression for the adjustment implied by temporal risk aversion on mhe^{rs}. Expression for dwi^{add} is in equation (7). For dwi^{rs} in equation (9), we use a more restricted expression than Lemoine and Traeger [37]. Indeed, we exclude mpre from dwi and consider only the differential impact of a marginal increase in the pre and post tipping if the tipping point is actually crossed (with probability h). The expression writes:

$$dwi^{rs}_{t+1} = h_{t+1} \frac{\partial S_{t+1}}{\partial \mu_t} \left(\frac{\phi'(V^{pre})}{\phi'(V^{eff})} \frac{\partial V^{pre}_{t+1}}{\partial S_{t+1}} - \frac{\phi'(V^{post})}{\phi'(V^{eff})} \frac{\partial V^{post}_{t+1}}{\partial S_{t+1}} \right) \tag{A.5}$$

Then, we can write:

$$dwi_{t+1}^{rs} = dwi_{t+1}^{add} + h_{t+1}\frac{\partial S_{t+1}}{\partial \mu_t}\left(\left[\frac{\phi'(V^{pre})}{\phi'(V^{eff})} - 1\right]\frac{\partial V_{t+1}^{pre}}{\partial S_{t+1}} - \left[\frac{\phi'(V^{post})}{\phi'(V^{eff})} - 1\right]\frac{\partial V_{t+1}^{post}}{\partial S_{t+1}}\right) \quad \text{(A.6)}$$

We do a first-order approximation of $\phi'(V^i)$ for i $\in \{pre, post\}$ as Lemoine and Traeger [37], assuming that the tipping point does not cause too large a welfare loss, to obtain $\frac{\phi'(V^i)}{\phi'(V^{eff})} - 1 \approx \frac{\phi'(V^{eff})+\phi''(V^{eff})[V^i - V^{eff}]}{\phi'(V^{eff})} - 1 \approx -\epsilon(V^i - V^{eff})$. This approximation, together with equation (A.6), yields equation (10b). Finally, equation (10c) is derived from equations (7) and (9).

Appendix B. Resolution

We solve our recursive programs using dynamic programming. For each social welfare function, we approximate the value function in the post-tipping world and then in the pre-tipping world using the solution from the post-threshold problem. We interpolate recursively starting from the last period and approximate the unknown value functions with Chebyshev polynomials. We choose a 10^{-3} tolerance for the solver: our result is not affected by stricter tolerance. In each regime (pre- and post-tipping), the value functions are expected to be smooth as the tipping risk is the only risk we consider. We use a four degree complete Chebyshev approximation in the two-dimensional state space. Additional degrees do not affect the results. The state variables are the carbon stock in the atmosphere S_t and the stock of capital K_t at time t. The time-dependent approximation space is defined around a deterministic growth path derived from Ramsey formula. Once we have interpolated recursively at each time step, we simulate the optimal path for each control and state variables starting from the first period. In the stochastic case with a tipping point, we run 1.000 simulations. An increase in the number of simulations does not affect significantly the median path. A key element is the definition of a terminal value in the program. The calculation is done on a finite horizon (T =

600 years) as an approximation of the infinite program. The terminal value is defined as the sum of all the period utilities from time T to infinity. The assumption made is that the consumption will grow for a constant capital per efficient capita and total abatement, with a deterministic path for the capital derived from Ramsey. The terminal constraint uses a modified discount factor [7]. The choice of the terminal value does not affect the program : a 10% increase in the terminal value does not significantly affect the optimal path. It writes: $TVF = u(\overline{c})/(1-\beta(1+GA))^{\delta(\frac{1-\eta}{1-\alpha})}$ with $\overline{c}$ the consumption for constant capital per efficient capita and total abatement, β the discount rate, δ the time step, η the marginal utility parameter, α the capital elasticity in the production function, and GA the annual growth rate of productivity from the last period.

Appendix C. Risk-sensitive preferences and the risk premium

We show that the risk premium is positive for all ϵ under risk-sensitive preferences. We give the share of the risk-sensitive SCC under expected damages in the risk-sensitive stochastic SCC for various values of ϵ, J and for our benchmark $\eta = 1.5$, following Taconet et al. [48]. In particular, we show numerically on the graph on the left below that the risk premium is positive for all values of ϵ in $\mathbb{R}^+$ under risk-sensitive preferences, unlike for EZW preferences. As some pure risk is already priced under additive preferences with η, we also want to highlight how much the risk premium is increased by temporal risk aversion under our benchmark calibration: we plot on the graph on the right, for different values of ϵ and under a benchmark $J = 10\%$ and $\eta = 1.5$, the share of the additive risk premium in the risk-sensitive risk-premium: $100 * (SCC^{add}_{stoch} - SCC^{add}_{ed})/(SCC^{rs}_{stoch} - SCC^{rs}_{ed})$. The additive risk premium is always lower than the risk-sensitive risk premium for all $\epsilon \in \mathbb{R}^+$, i.e. when the social planner has temporal risk aversion.

On the graph on the left, we see that the share of expected damages in the

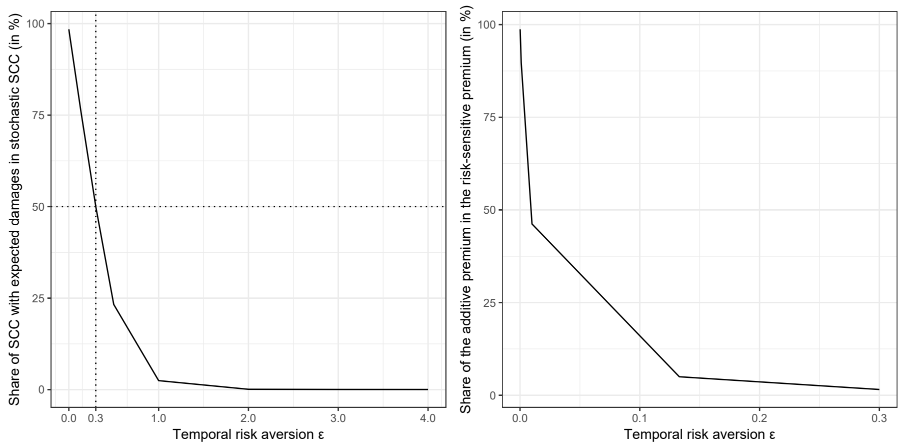


Figure C.4: The graph on the left gives the share of the stochatic SCC that is explained by expected damages (in %). The lowest value explored for ϵ is 0.0001 and the share would converge to 100 for $\epsilon \to 0$. The graph on the right gives the share of the risk-sensitive risk premium that is already priced under additive preferences. The lowest value explored for ϵ is 0.0001 and the share would converge to 100 for $\epsilon \to 0$. Both graphs are given for various ϵ and a benchmark J=10% and $\eta = 1.5$. The two graphs do not have the same scale for ϵ as the share goes quickly to 0 for values above $\epsilon > 0.3$ for the graph on the right

.

stochastic SCC is 50% for $\epsilon = 0.3$. In the remaining 50% of the stochastic SCC that are due to pure risk for $\epsilon = 0.3$, we see on the graph on the right that the pure risk already priced under additive preferences represents around 2% of the risk-sensitive risk premium. Most of the risk premium under risk-sensitive preferences stems from temporal risk aversion.

Appendix D. Time paths

We provide some time paths for our key variables.

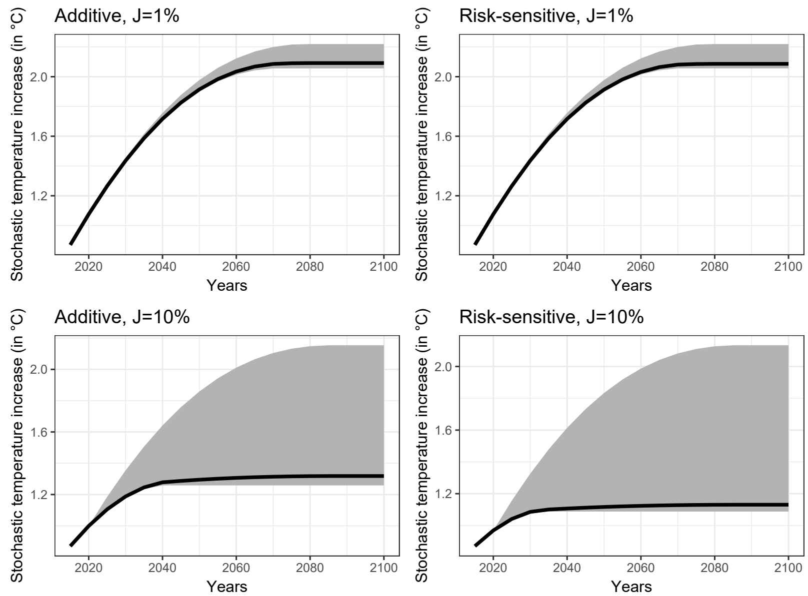


Figure D.5: The graphs give the time paths of the mean temperature increase until 2100 under additive (left) and risk-sensitive (right) preferences, for J=1% (up) and a J=10% (down), for $\epsilon = 0.133$. We give the mean (solid line) and [5% : 95%] confidence interval (shaded area) over 1.000 stochastic runs.

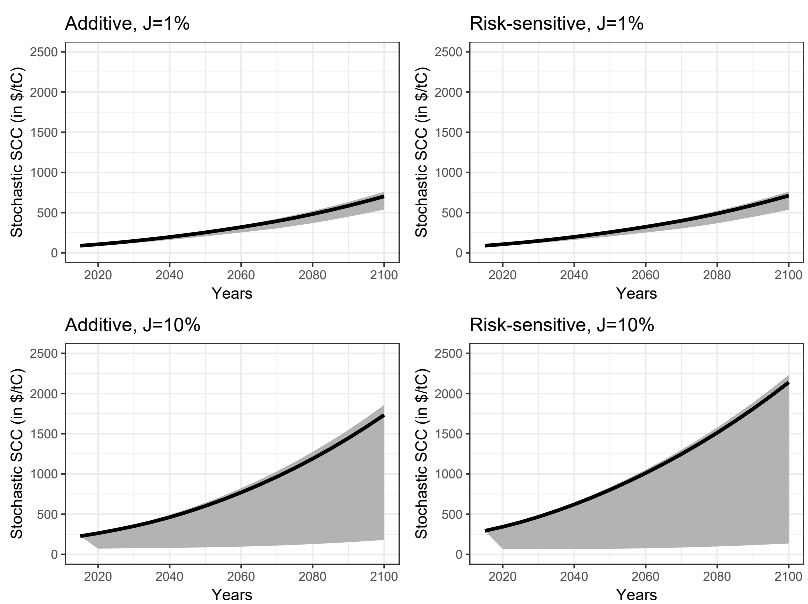


Figure D.6: Same graph as above but with the social cost of carbon.

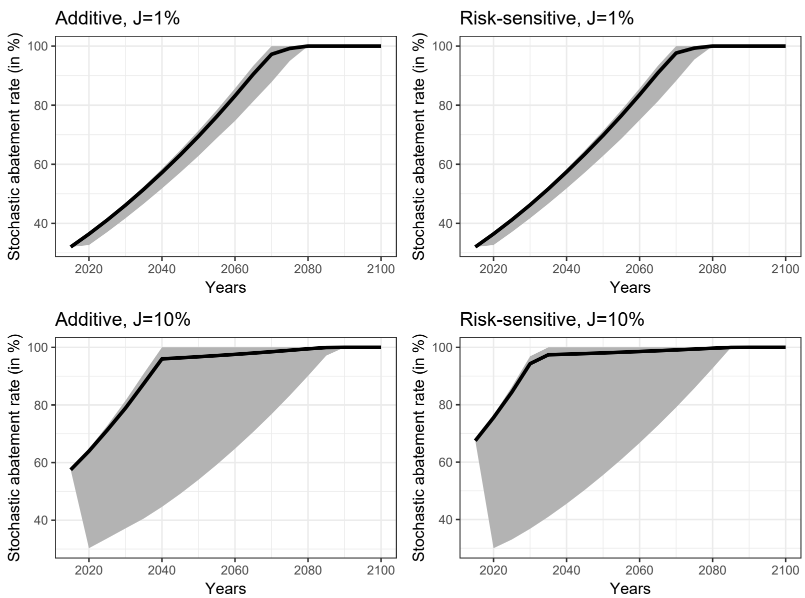


Figure D.7: Same graph as above but with the abatement rate.

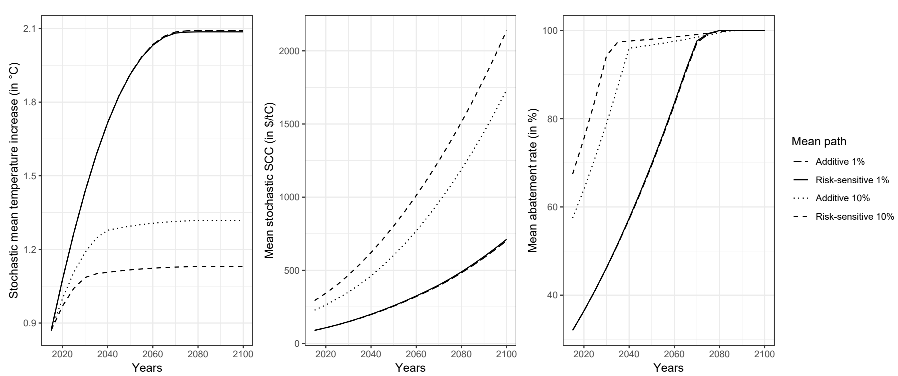


Figure D.8: The graphs give the mean time paths of the temperature increase in °C with respect to preindustrial era (left), the SCC (middle) and the abatement rate (in %) until 2100. In the risk-sensitive case and for J=10% (J=1%), the tipping point is crossed 4.4% (22.8%) of the 1000 runs over the whole time horizon considered. In the additive case and for J=10% (J=1%), it is 7.4% (26.9%).

Appendix E. Sensitivities

Appendix E.1. Upper temperature threshold

The ratio of the SCC under risk-sensitive preferences to the SCC under additive preferences decreases with the upper threshold, i.e. a lower probability of tipping decreases the difference between the two criteria.

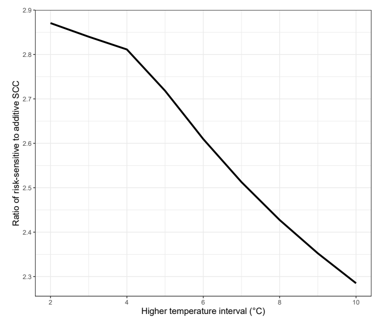


Figure E.9: Ratio of risk-sensitive to additive SCC for a benchmark $\epsilon = 0.133$, $J = 20\%$ and various upper temperature threshold.

Appendix E.2. Higher tipping damage J

We give the same graph as in the main text but for a larger range of J.

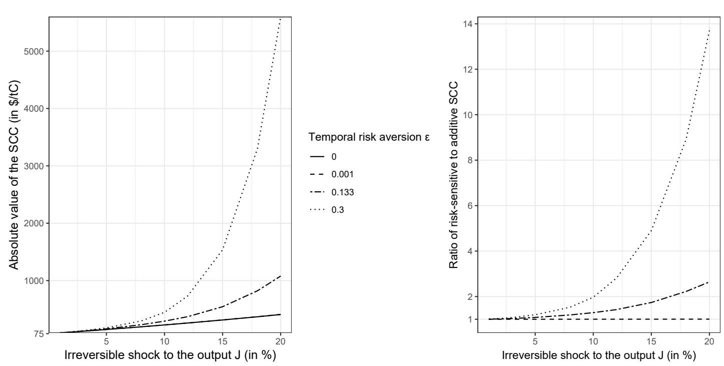


Figure E.10: Ratio of risk-sensitive to additive SCC at initial time for a benchmark $\epsilon = 0.133$ and various J.

Appendix E.3. Inequality aversion

We plot the log ratio of the SCC under risk-sensitive preferences to the SCC under additive preferences for different J and η, with $\epsilon = 0.133$. The ratio increases with the elasticity of marginal utility.

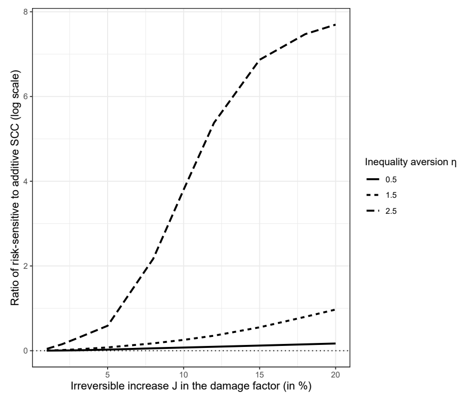


Figure E.11: Ratio of the SCC at initial time under risk-sensitive preferences on the SCC under additive preferences (benchmark calibration) for different J and η.

Appendix E.4. Sensitivity - immediate decomposition

We give the numerical decomposition for the immediate channels under risk-sensitive preferences for various J and ϵ.

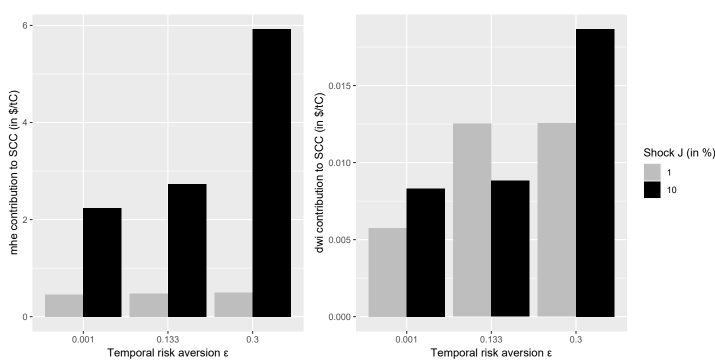


Figure E.12: Marginal contribution to the SCC at initial time (in \$) of the immediate *mhe* (left) and the immediate *dwi* (right) under risk-sensitive preferences for various J and ϵ. Results based on 1000 stochastic runs.

Disclosure statement

No competing interest to declare.

Acknowledgments

We thank two anonymous referees for their invaluable help in improving the manuscript. We also thank Thomas Douenne, Johannes Emmerling, Simon Jean, Vincent Martinet, Aurelie Méjean, Eddy H.F. Tam and Stephane Zuber for comments on earlier versions of this work. Any remaining errors are ours. Céline Guivarch received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101081604 (PRISMA).

References

[1] Ackerman, F., E. A. Stanton, and R. Bueno (2013). Epstein–zin utility in dice: Is risk aversion irrelevant to climate policy? *Environmental and Resource Economics 56*(1), 73–84.

[2] Ahn, C. M. (1989). The effect of temporal risk aversion on optimal consumption, the equity premium, and the equilibrium interest rate. *The Journal of Finance 44*(5), 1411–1420.

[3] Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2018). Multiattribute utility theory, intertemporal utility, and correlation aversion. *International Economic Review 59*(2), 537–555.

[4] Anderson, E. W. (2005). The dynamics of risk-sensitive allocations. *Journal of Economic theory 125*(2), 93–150.

[5] Arias, P., N. Bellouin, E. Coppola, R. Jones, G. Krinner, J. Marotzke, V. Naik, M. Palmer, G.-K. Plattner, J. Rogelj, et al. (2021). Climate change

2021: The physical science basis. contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change; technical summary.

[6] Arrow, K. J. (1971). The theory of risk aversion. chapter 3 in essays in the theory of risk bearing. *New York: American Elsevier*.

[7] Barr, J. L. and A. S. Manne (1967). Numerical experiments with a finite horizon planning model. *Indian Economic Review 2*(1), 1–31.

[8] Belaia, M., M. Funke, and N. Glanemann (2017). Global warming and a potential tipping point in the atlantic thermohaline circulation: the role of risk aversion. *Environmental and Resource Economics 67*(1), 93–125.

[9] Berger, L. and L. Eeckhoudt (2021). Risk, ambiguity, and the value of diversification. *Management Science 67*(3), 1639–1647.

[10] Bommier, A. (2007). Risk aversion, intertemporal elasticity of substitution and correlation aversion. *Economics Bulletin 4*(29), 1–8.

[11] Bommier, A. (2013). Life-cycle preferences revisited. *Journal of the European Economic Association 11*(6), 1290–1319.

[12] Bommier, A., A. Chassagnon, and F. Le Grand (2012). Comparative risk aversion: A formal approach with applications to saving behavior. *Journal of Economic Theory 147*(4), 1614–1641.

[13] Bommier, A., A. Kochov, and F. Le Grand (2017). On monotone recursive preferences. *Econometrica 85*(5), 1433–1466.

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

[15] Bommier, A. and F. Le Grand (2019). Risk aversion and precautionary savings in dynamic settings. *Management Science 65*(3), 1386–1397.

[16] Bommier, A. and S. Zuber (2008). Can preferences for catastrophe avoidance reconcile social discounting with intergenerational equity? *Social Choice and Welfare 31*(3), 415–434.

[17] Bretschger, L. and A. Vinogradova (2019). Best policy response to environmental shocks: Applying a stochastic framework. *Journal of Environmental Economics and Management 97*, 23–41.

[18] Bullard, J. and A. Singh (2012). Learning and the great moderation. *International Economic Review 53*(2), 375–397.

[19] Cai, Y. and T. S. Lontzek (2019). The social cost of carbon with economic and climate risks. *Journal of Political Economy 127*(6), 2684–2734.

[20] Clarke, H. R. and W. J. Reed (1994). Consumption/pollution tradeoffs in an environment vulnerable to pollution-related catastrophic collapse. *Journal of Economic Dynamics and Control 18*(5), 991–1010.

[21] Crost, B. and C. P. Traeger (2013). Optimal climate policy: uncertainty versus monte carlo. *Economics Letters 120*(3), 552–558.

[22] Diamond, P. A. (1967). Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: Comment. *Journal of political economy 75*(5), 765–766.

[23] Dietz, S. and F. Venmans (2019). Cumulative carbon emissions and economic policy: in search of general principles. *Journal of Environmental Economics and Management 96*, 108–129.

[24] Ebert, S. and G. van de Kuilen (2015). Measuring multivariate risk preferences. *Available at SSRN 2637964*.

[25] Fleurbaey, M. (2010). Assessing risky social situations. *Journal of Political Economy 118*(4), 649–680.

[26] Fleurbaey, M. (2018). Welfare economics, risk and uncertainty. *Canadian Journal of Economics/Revue canadienne d'économique 51*(1), 5–40.

[27] Gangadharan, L., G. W. Harrison, and A. D. Leroux (2019). Are risks over multiple attributes traded off? a case study of aid. *Journal of Economic Behavior & Organization 164*, 166–198.

[28] Guivarch, C. and A. Pottier (2018). Climate damage on production or on growth: what impact on the social cost of carbon? *Environmental Modeling & Assessment 23*(2), 117–130.

[29] Hansen, L. P. and T. J. Sargent (1995). Discounted linear exponential quadratic gaussian control. *IEEE Transactions on Automatic control 40*(5), 968–971.

[30] Jensen, S. and C. P. Traeger (2014). Optimal climate change mitigation under long-term growth uncertainty: Stochastic integrated assessment and analytic findings. *European Economic Review 69*, 104–125.

[31] Kihlstrom, R. E. and L. J. Mirman (1974). Risk aversion with many commodities. *Journal of Economic Theory 8*(3), 361–388.

[32] Kimball, M. and P. Weil (2009). Precautionary saving and consumption smoothing across time and possibilities. *Journal of Money, Credit and Banking 41*(2-3), 245–284.

[33] Kreps, D. M. and E. L. Porteus (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 185–200.

[34] Lanier, J., B. Miao, J. K.-H. Quah, and S. Zhong (2022). Intertemporal consumption with risk: A revealed preference analysis. *Review of Economics and Statistics*.

[35] Leduc, M., H. D. Matthews, and R. de Elía (2015). Quantifying the limits of a linear temperature response to cumulative co2 emissions. *Journal of Climate 28*(24), 9955–9968.

[36] Lemoine, D. and C. Traeger (2014). Watch your step: optimal policy in a tipping climate. *American Economic Journal: Economic Policy 6*(1), 137–66.

[37] Lemoine, D. and C. P. Traeger (2016). Ambiguous tipping points. *Journal of Economic Behavior & Organization 132*, 5–18.

[38] Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, et al. (2021). Climate change 2021: the physical science basis. *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change 2*.

[39] Nordhaus, W. (2018). Evolution of modeling of the economics of global warming: changes in the dice model, 1992–2017. *Climatic change 148*(4), 623–640.

[40] Nordhaus, W. D. (2008). *A question of balance*. Yale University Press.

[41] Piontek, F., L. Drouet, J. Emmerling, T. Kompas, A. Méjean, C. Otto, J. Rising, B. Soergel, N. Taconet, and M. Tavoni (2021). Integrated perspective on translating biophysical to economic impacts of climate change. *Nature Climate Change*, 1–10.

[42] Pratt, J. W. (1964). Risk aversion in the large and in the small. *Econometrica 32*(1-2), 122–136.

[43] Rheinberger, C. M. and N. Treich (2017). Attitudes toward catastrophe. *Environmental and Resource Economics 67*(3), 609–636.

[44] Richard, S. F. (1975). Multivariate risk aversion, utility independence and separable utility functions. *Management Science 22*(1), 12–21.

[45] Rohde, K. I. and X. Yu (2022). Intertemporal correlation aversion-a model-free measurement. *Available at SSRN 4079812*.

[46] Rudik, I. (2020). Optimal climate policy when damages are unknown. *American Economic Journal: Economic Policy 12*(2), 340–73.

[47] Stern, N. (2006). Stern review: The economics of climate change.

[48] Taconet, N., C. Guivarch, and A. Pottier (2021). Social cost of carbon under stochastic tipping points. *Environmental and Resource Economics 78*(4), 709–737.

[49] Tallarini Jr, T. D. (2000). Risk-sensitive real business cycles. *Journal of monetary Economics 45*(3), 507–532.

[50] Tsur, Y. and A. Zemel (1996). Accounting for global warming risks: Resource management under event uncertainty. *Journal of Economic Dynamics and Control 20*(6-7), 1289–1305.

[51] van der Ploeg, F. and A. de Zeeuw (2018). Climate tipping and economic growth: precautionary capital and the price of carbon. *Journal of the European Economic Association 16*(5), 1577–1617.

[52] van der Ploeg, F. and A. de Zeeuw (2019). Pricing carbon and adjusting capital to fend off climate catastrophes. *Environmental and Resource Economics 72*(1), 29–50.

[53] Weitzman, M. L. (2009). On modeling and interpreting the economics of catastrophic climate change. *The review of economics and statistics 91*(1), 1–19.