

Belaia, Mariia; Funke, Michael; Glanemann, Nicole

**Working Paper**

## Global Warming and a Potential Tipping Point in the Atlantic Thermohaline Circulation: The Role of Risk Aversion

CESifo Working Paper, No. 4930

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Belaia, Mariia; Funke, Michael; Glanemann, Nicole (2014) : Global Warming and a Potential Tipping Point in the Atlantic Thermohaline Circulation: The Role of Risk Aversion, CESifo Working Paper, No. 4930, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/102251>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



# Working Papers

[www.cesifo.org/wp](http://www.cesifo.org/wp)

## Global Warming and a Potential Tipping Point in the Atlantic Thermohaline Circulation: The Role of Risk Aversion

Mariia Belaia  
Michael Funke  
Nicole Glanemann

CESIFO WORKING PAPER NO. 4930  
CATEGORY 10: ENERGY AND CLIMATE ECONOMICS  
AUGUST 2014

*An electronic version of the paper may be downloaded*

- from the SSRN website: [www.SSRN.com](http://www.SSRN.com)
- from the RePEc website: [www.RePEc.org](http://www.RePEc.org)
- from the CESifo website: [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

# Global Warming and a Potential Tipping Point in the Atlantic Thermohaline Circulation: The Role of Risk Aversion

## Abstract

Common integrated assessment models produce the counterintuitive result that higher risk aversion does not lead to stronger near-term abatement. This paper re-examines this result with a DICE model that is fully coupled with a thermohaline circulation model. It also features Epstein-Zin utility and uncertainty about climate sensitivity that resolves after some time. The simulations show that aversion to this tipping point risk has little effect. For climate sensitivity of realistic magnitude, a collapse of the circulation occurs in the distant future, which allows acting after learning. Furthermore, the anticipated damage costs are not sufficiently great to justify precautionary measures.

JEL-Code: Q540, Q560, C610, C630.

Keywords: integrated assessment modeling, risk aversion, Epstein-Zin utility, DICE, thermohaline circulation, climate sensitivity, uncertainty.

*Mariia Belaia*  
*Department of Economics*  
*Hamburg University / Germany*  
*mariia.belaia@wiso.uni-hamburg.de*

*Michael Funke*  
*Department of Economics*  
*Hamburg University / Germany*  
*michael.funke@uni-hamburg.de*

*Nicole Glanemann*  
*Department of Economics*  
*Hamburg University / Germany*  
*nicole.glanemann@wiso.uni-hamburg.de*

July 2014

We would like to thank Frank Ackerman, Ramon Bueno and Elisabeth A. Stanton for sharing the GAMS code of their EZ-DICE model. We would also like to acknowledge Kirsten Zickfeld for providing background information on the four-box THC model and for offering the MATLAB code. Furthermore, we are grateful to Chao Li for his expert advice on Atlantic thermohaline circulation. The research was financially supported by the Cluster of Excellence “Integrated Climate System Analysis and Prediction” (CliSAP) of the University of Hamburg, which is funded in the framework of the Excellence Initiative of the German federal and state Governments.

# 1 Introduction

A stable climate system in the future is a valuable international public good. Like a stable world financial system, its benefits are largely taken for granted until they are lost. National authorities and international institutions are guardians of this public good. Their responsibility is to implement climate policies that preserve it. The possibility of large-scale climate risk in the distant future poses a particular challenge to climate policy assessment. In this respect, special attention is paid to tipping points, which define abrupt and in some cases even irreversible transitions in the state of a major component of the Earth (Lenton et al., 2008). An example of a tipping element is given by the Atlantic thermohaline circulation (THC), which transports immense amounts of heat northward and is thus an important guarantor of the heat supply in the North Atlantic region (Ganachaud and Wunsch, 2000). According to IPCC (2007), the ongoing warming of the Atlantic is very likely to lead to a slowdown in the THC during the 21st century, which raises concerns about a shutdown of the heat supply in the next century.<sup>1</sup> The likelihood of such a tipping point happening is thus affected by actions over a long period of time. Consequently, climate policy assessment must be equipped with methods that treat long periods of time and the involved risks adequately. Lately, most of the integrated assessment models (IAMs) used for climate policy recommendations have faced increasing criticism stating that they appear greatly to underestimate these risks (e.g. Stern (2013)). The main arguments put forward are, inter alia, that they do not encompass sufficient modelling of the ubiquitous uncertainties involved (e.g. Pindyck (2013)), that the risks of climate catastrophes are not satisfactorily accounted for (e.g. Lenton and Ciscar (2013)) and that the climate damage functions are not reconcilable with current research on the impacts of climate change (e.g. Ackerman et al. (2009) and Kopp et al. (2013)). Another main point of criticism is that the assumptions implicitly made by the commonly used concept of intertemporally additive expected utility, such as the constant relative risk aversion utility function (CRRA), provide a contorted representation of a decision maker's preferences. More precisely, the coefficient of relative risk aversion and the inverse of the elasticity of intertemporal substitution are merged as one parameter. An increase in this parameter could be attributed to a change in the preference of the time and interpreted as greater aversion to a transfer of wealth from the present to the future. As a result, the social discount rate also increases and policy is less concerned about climate change. The bias of the concept, however, is caused when the decision maker becomes more risk averse. Then the coefficient also increases, which, as already argued, leads to less abatement. Accordingly, the

---

<sup>1</sup>Long-run predictions on the THC, however, involve tremendous uncertainties and the estimates of future circulation strength are therefore somewhat speculative (Matei et al. (2012)). As reviewed by Lenton and Ciscar (2013), the expert elicitation study by Kriegler et al. (2009) and the IPCC (2007) suggest that rather high temperatures ( $>4^{\circ}\text{C}$ ) are needed to push the THC towards its collapse. However, these models are criticized for being biased towards the stability of the THC dynamics (Drijfhout et al. (2011)) and observations indicate a higher vulnerability of the THC (Drijfhout et al. (2011) and Hawkins et al. (2011)).

application of such a utility concept leads to counterintuitive results concerning the role of risk and risk aversion. By utilizing a more realistic preference specification, this paper aims to reexamine the optimal policy response to risk, in particular to the risk of a THC collapse, and investigates the role of risk aversion in the stringency of the optimal policy response. The limitations of the parsimonious and easily tractable intertemporally additive expected utility functions were first identified in finance by demonstrating the “equity premium puzzle” (Mehra and Prescott 1985). This puzzle illustrates the difficulties in explaining why the rates on return are rather high in terms of equity and rather low in relation to risk-free assets. Although it has not been possible to resolve this problem so far, possible explanations for this puzzle could be provided and it has resulted in some fundamental concepts being questioned.<sup>2</sup> An intuitively appealing approach is pursued by the “generalized expected utility” concept proposed by Epstein and Zin (1989, 1991) and Kreps and Porteus (1978), which allows a separate preference specification for risk and time in a recursive utility model. Since then, the Epstein-Zin preferences have gained in importance and found many applications. It is not the goal of this paper to give a literature review of these applications, but some striking examples of the contribution of this concept shall be mentioned. For instance, the Epstein-Zin preferences are found to illuminate many patterns of asset pricing, especially in the case of long-term risk (e.g. Bansal and Yaron (2004)). An important contribution to macroeconomics is exemplified by Tallarini (2000), showing that the Epstein-Zin preferences significantly increase the welfare costs of the business cycle by providing refined model predictions for the asset market.

In climate economics, most of the IAMs, such as the famous DICE (Nordhaus, 2008), FUND (Tol, 2009) and PAGE (Hope, 2006), rely on the CRRA utility specification. The equivalent to the equity premium puzzle given by a “negative risk premium on high climate change outcomes” (Nordhaus, 2008) in DICE is attributed to the preference specification and to the absence of tail risk in the model (Kaufman (2012)). Recent advances in accounting for Epstein-Zin preferences and for fat-tailed risk indicate a significant understatement of the optimal emission reduction rates in DICE (e.g. Ackerman et al. (2013), Cai et al. (2013), Crost and Traeger (2010), Jensen and Traeger (2014), Kaufman (2012), Lemoine and Traeger (2014)). Of particular importance to this paper is the study by Ackerman et al. (2013). Ackerman et al. (2013) report the results generated by their “EZ-DICE” model to be rather insensitive to risk aversion and only to show meaningful effects owing to the intertemporal elasticity of substitution. Ackerman et al. (2013) capture climate risk by accounting for a fat-tailed climate sensitivity distribution, where the decision maker knows that the true climate sensitivity value will be disclosed in the middle of the second half of the century. At this point in time, however, DICE allows for jumping to 100% abatement at relatively low

---

<sup>2</sup>For a recent survey see DeLong and Magin (2009).

cost, which can completely avoid the damage induced by higher climate sensitivity values. Consequently, irrespective of the level of risk aversion, it is always beneficial to respond to the risk only after the true state of nature is revealed and then to act accordingly. A more realistic modelling of the abatement options and of the climate risks is surmised to provide a stronger risk aversion effect (Ackerman et al. (2013)). A strong effect is also supported by the simple, albeit reduced-form, framework in Barro (2013). Our research tackles the question “Is risk aversion irrelevant to climate policy” raised by Ackerman et al. (2013) and develops an enhanced version of the EZ-DICE model that satisfies these features towards building a more realistic modelling framework.

In our model, uncertainty about climate sensitivity translates into uncertainty about a slow-down of the THC. We consider modelling climate risk in terms of the potential for a THC collapse as particularly interesting because the dynamics involved lag behind the emission of greenhouse gases by several decades and also, as mentioned above, it exhibits an abrupt transition to a complete shut-down. Accordingly, once the true climate sensitivity value is identified, it may be too late to prevent with certainty the THC from collapsing. Risk aversion may then play a more meaningful role for climate policy. The literature incorporating a more or less refined representation on the THC dynamics into an IAM is rich and may be categorized according to the representation of oceanographic dynamics. The more refined representations are given by the two-box Stommel model (Stommel, 1961) or extensions of the same (e.g. the Simple Climate Demonstrator by Schneider and Thompson (2000) and a four-box model by Zickfeld et al. (2004)), in which the heat transport is explicitly given. To our knowledge, none of the IAMs that are coupled with such a box model account for the more realistic Epstein-Zin preferences. However, this strand of research acknowledges the problem of accounting for a risky event in the far future. Mastrandrea and Schneider (2001) demonstrate the importance of choosing a more realistic discounting by opting for hyperbolic discounting as otherwise the damages from a THC collapse vanish from today’s welfare calculus. To avoid discounting the impacts altogether, other studies (e.g. Zickfeld and Bruckner (2003) and Zickfeld and Bruckner (2008)) employ cost effectiveness analysis in which only policies that (almost certainly) prevent the THC from collapsing are permitted. At the other end of the spectrum, an example of a stylized representation of the threat of a THC collapse is given by hazard rates (e.g. Lontzek et al. (2012) and Nævdal and Oppenheimer (2007)). Lontzek et al. (2012) calibrate this hazard rate on the basis of elicitation studies by Kriegler et al. (2009) and Zickfeld et al. (2007). Following this idea, Cai et al. (2013) adopt this simplistic concept to model (unspecified) tipping points and can also account for Epstein-Zin preferences. Again, Epstein-Zin preferences are shown to increase the stringency of optimal climate policy. We believe that our research question on the role of risk aversion benefits from the more precise modelling of the THC given by the box models. Thus, we build a climate model with geophysical microfoundations that then, as an emer-

gent property, exhibits multiple steady states and transition dynamics. As argued above, the explicitly modelled transition dynamics and in particular the characteristic of the THC significantly lagging behind temperature evolution may play an important role for the effects of risk aversion and are thus better modelled by a fully-fledged box model.

In a nutshell, our model is a fully coupled continuous-time EZ-THC-DICE model. More precisely, it combines the continuous-time DICE version by Cai et al. (2012a) with the continuous-time four-box THC model by Zickfeld et al. (2004). As in Ackerman et al. (2013), the “Epstein-Zin policy maker” needs to decide on an optimal policy path while facing a fat-tailed probability distribution for the true climate sensitivity parameter. In the year 2075, uncertainty is resolved and the policy path can be adjusted. For this specific set-up, we aim to clarify whether the climate policy path until 2075 is significantly affected by the fat-tailed risk of a THC collapse and, more importantly, by the policy maker’s attitude towards this risk. We also investigate how restrictions on the policy adjustment affect our results. Despite ever advancing technology, ceasing to emit instantaneously might be still impossible in the year 2075. Apart from technological limitations, it is also conceivable that policy inertia or inertia in the economic system will slow emission reduction efforts down.

The paper is structured as follows. Section 2 elucidates how the two models are coupled and goes into detail about new and less well known parts. Section 3 describes the calibration of the model and presents the resulting simulations together with an analysis. Section 4 draws conclusions based on the findings.

## 2 The Model

To represent the interaction between the economy, the climate and the THC, we choose to couple fully a version of the original DICE-2007 model (Nordhaus, 2008) and the continuous-time four-box THC model by Zickfeld et al. (2004). To guarantee the consistency of the time steps in the two sub-models, we choose to keep the continuous-time formulation of the THC model and to adopt the continuous-time version of DICE-2007 (DICE-CJL) by Cai et al. (2012a).<sup>3</sup> A rough illustration of our model is provided by Figure 1, which highlights the changes to DICE-CJL and how the two models are coupled. DICE-CJL in itself links a Ramsey-type optimal growth model of the global economy, which produces emissions, with a simplified climate model, which describes the effect of the emissions on the increase in global mean temperature  $\Delta T$ . The THC model then takes the global mean temperature increase  $\Delta T$  as input to generate an assessment concerning the strength of the THC overturning  $m$ . This value discloses whether the THC is slowing down or whether it has already collapsed

---

<sup>3</sup>The most important modifications to DICE-2007 implemented by Cai et al. (2012a) involve recalibrations of the parameters owing to the different time units. In addition, the atmospheric temperature response function is adjusted to rule out warming being affected by future atmospheric carbon concentrations. Note that “continuous” refers to a time-discretization of one-year time steps.

and thus incurs additional climate damage costs. To account for these costs, we develop an alternative formulation for the damage cost function that increases with global warming  $\Delta T$  and with a lower overturning strength  $m$ .

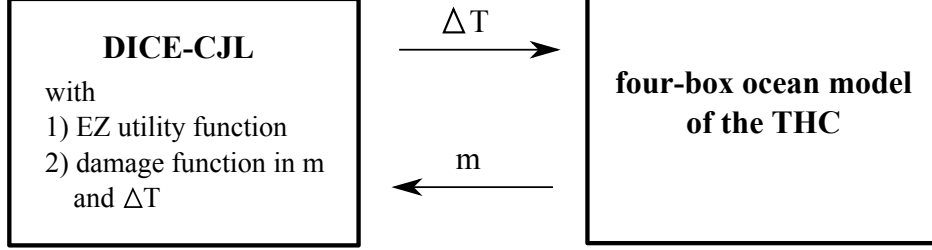


Figure 1: Illustration of the model coupling: the global mean temperature increase  $\Delta T$  is used as input for the THC model, while the THC model computes the overturning strength  $m$ . The new damage cost function comprises the impacts of both  $\Delta T$  and  $m$ .

The remainder of this section explains the parts of the model that are new or less well known in detail. As DICE is presumably the most frequently applied IAM in climate economics, we avoid giving a detailed description of DICE and instead proceed directly to explaining the changes to this model. To ensure clarity, we also point out the most important equations in the THC model.

The first alteration to DICE replaces the commonly used CRRA utility function

$$U(c_t, l_t) = \frac{\left(\frac{c_t}{l_t}\right)^{1-\gamma}}{1-\gamma} l_t, \quad (1)$$

where  $c_t$  is the optimized consumption path over time and  $l_t$  specifies the exogenously evolving labour supply over time. The motivation for dismissing the CRRA utility function can be explained by directing attention towards  $\gamma$ , which is referred to as the coefficient of relative risk aversion. Through the calibration of  $\gamma$  the curvature of the utility is configured and thereby the decision maker's attitude towards risk. The larger this parameter, the more risk averse the decision maker, who will refrain from gambling on uncertain outcomes. Assuming a different attitude towards risk also affects the “Ramsey equation” (Cass (1965), Koopmans (1963) and Ramsey (1928)) in which  $\gamma$  also appears. This states that for optimal intertemporal allocation the social discount rate  $r$  equals

$$r = \delta + \gamma g, \quad (2)$$

where  $\delta$  is the rate of pure time preference and  $g$  is the rate of growth per capita consumption. In this context,  $\gamma$  is referred to as the elasticity of the marginal utility of consumption or the inverse of the intertemporal elasticity of substitution. When assuming a more risk-averse climate policy maker,  $\gamma$  increases, which leads to a more concave utility function in equation (1) and a higher social discount rate  $r$  in equation (2). Consequently, future damage costs receive less weight in the welfare assessment, leading to less concern about future risks.

The CRRA utility function thus cannot mirror concerns about risk, without implying time preferences. In doing so, only two preference specifications are allowed: less concern for the future and high concern for current risks or high concern for the future and less concern for current risks. Disentangling time preferences and risk aversion is achieved by the recursive utility function provided by Epstein and Zin (1989, 1991), which derives utility at time  $t$ ,  $U_t$ , by drawing on current consumption and the certainty equivalent of the utility in the future,

$$U_t = \left( (1 - \beta) \frac{\left(\frac{c_t}{l_t}\right)^{1-\psi}}{1-\psi} l_t + \beta \left(\mathbb{E}_t [U_{t+1}^{1-\alpha}]\right)^{\frac{1-\psi}{1-\alpha}} \right)^{\frac{1}{1-\psi}}, \quad (3)$$

where  $\beta$  is the discount factor for utility and relates to the rate of pure time preferences by  $\delta = \frac{1-\beta}{\beta}$ . The expectation operator at time  $t$  is described by  $\mathbb{E}_t$ . The inverse of the intertemporal elasticity of substitution is now given by  $\psi$  and  $\alpha$  is the risk aversion parameter. For  $\psi = \alpha$ , equation (3) would collapse into a time separable specification as in equation (1). However, Bansal and Yaron (2004) point out that this case would be rather unrealistic.

The recursive structure of the Epstein-Zin utility specification causes great analytical and computational costs. To compute current utility, not only current consumption but also the expected value of the next period's utility must be known. The next period's utility, in turn, depends on the expected utility of the following period and so on. Therefore, current utility requires computation of the full branching tree of all possible futures. The implementation of equation (3) in IAMs has been accomplished by reducing complexity in one way or another: (i) Crost and Traeger (2010) opt to give a more stylized version of the climate dynamics to transform the problem into a Bellman equation; (ii) Ha-Duong and Treich (2004) and Kaufman (2012) reduce the number of branches by accounting only for a few periods; (iii) Ackerman et al. (2013) thin out the tree by limiting the number of possible values for the uncertain parameter. We choose to follow closely the method used by Ackerman et al. (2013) for two reasons. First, as our research addresses the research question posed by Ackerman et al. (2013), we shall ensure comparability of the results. Second, we consider this method rather advantageous compared to the other approaches as it retains the long time horizon structure in climate change decision problems. In addition, it allows us to implement additional dynamics given by the THC and thus even to increase model complexity. Following Ackerman et al. (2013), we assume that five climate sensitivity values are the possible candidates for being the true state of nature. Each climate sensitivity value is assigned a different probability.<sup>4</sup> In the first 70 years, the decision maker is required to decide on one policy path factoring in the uncertainty concerning all five climate sensitivity values. In 2075, uncertainty is fully resolved and the policy maker is then allowed to adjust policy accordingly. Basically, this design makes it possible to consider a tree that exhibits

---

<sup>4</sup>The next section describes in detail how these values and their probabilities are derived.

one branch until 2075, which goes out only once into five branches. The Epstein-Zin utility function for the first branch draws on current consumption and on the expected value of the utility after 2075. The expected value comprises the five different utilities that relate to the realization of one of the possible “worlds” or branches. These five utility functions are deterministic and reflect the present value over the remaining time. The recursive structure of the problem is therefore only given for the time until 2075. This simplified representation of uncertainty allows us to expand the IAM used by Ackerman et al. (2013) and also to account for detailed representation of the THC.

The conceptual four-box model we implement is designed to reproduce key processes such as the possibility of a sudden collapse and the driving mechanisms of the overturning. A schematic representation of the model is provided in Figure 2. The Atlantic is partitioned into the South Atlantic (box 1), the North Atlantic (box 2) and the Tropics (box 3 stands for the ocean surface and box 4 is associated with the deep ocean). The bold arrows indicate the direction of water circulation, which is determined by global density gradients generated by surface heat and freshwater fluxes. In the South Atlantic, the water is warmed and flows northward, where it cools en route, sinks down and then flows back towards the South, where it warms again. The double arrows depict the heat exchange with the atmosphere in the different regions. In particular, the double arrow above box 2 describes the warmth provided by the THC to the North Atlantic region. The temperatures  $T_i^*$ ,  $i \in \{1, 2, 3\}$  are referred to as the “restoring temperatures”. The curved arrows hint at the atmospheric water vapour transport  $F_1$  and  $F_2$  determining the degree of salinity in the surface boxes. The third curved arrow, which is not connected to any of the other boxes, indicates an additional freshwater flux  $F_2^{MW}$  provided by the meltwater runoff from the Greenland ice sheet and the meltwater from the Arctic sea ice.

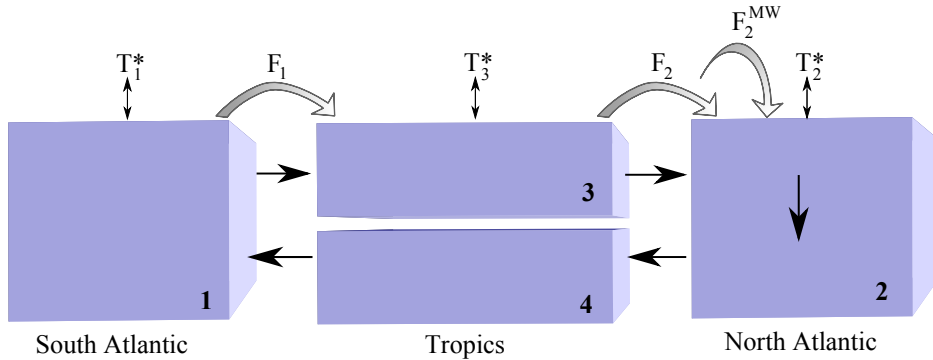


Figure 2: Schematic representation of the THC box model

The global mean temperature  $\Delta T$  provided by DICE affects the THC model through two channels. The global temperature increase can be scaled to derive the regional restoring

temperatures:

$$\Delta T_i^*(t) = p_i \Delta T(t), \quad i \in \{1, 2, 3\}, \quad (4)$$

where  $p_i$  are regional temperature constants. In addition, the freshwater forcing evolves according to  $\Delta T$ :

$$\Delta F_1(t) = h_1 p_{SH} \Delta T(t), \quad (5)$$

$$\Delta F_2^{tot} = \Delta F_2(t) + \Delta F_2^{MW}(t) = (h_2 + h_2^{MW}) p_{NH} \Delta T(t) \equiv h_2^{tot} p_{NH} \Delta T(t), \quad (6)$$

where  $p_{SH}$  and  $p_{NH}$  are temperature constants for the Southern and Northern Hemispheres, respectively, and  $h_1$ ,  $h_2$  and  $h_2^{MW}$  are hydrological sensitivity parameters. We define  $h_2^{tot}$  as the sum of the two hydrological sensitivity parameters  $h_2$  and  $h_2^{MW}$ .

Water temperature  $T_i$ ,  $i \in \{1, \dots, 4\}$ , in the boxes is influenced by the restoring temperatures  $T_i^*$  and the temperature of the water that flows into the box  $i$ , as indicated by the bold arrows in Figure 2. More specifically,

$$\dot{T}_1 = \frac{m}{V_1} (T_4 - T_1) + \lambda_1 (T_1^* - T_1), \quad (7)$$

$$\dot{T}_2 = \frac{m}{V_2} (T_3 - T_2) + \lambda_2 (T_2^* - T_2), \quad (8)$$

$$\dot{T}_3 = \frac{m}{V_3} (T_1 - T_3) + \lambda_3 (T_3^* - T_3), \quad (9)$$

$$\dot{T}_4 = \frac{m}{V_4} (T_2 - T_4), \quad (10)$$

where  $V_i$  are the box volumes,  $\lambda_i$  thermal coupling constants and  $m$  the above-mentioned meridional volume transport. The dot over the variable is used to denote the time derivative. The salinity levels in the boxes  $S_i$ ,  $i \in \{1, \dots, 4\}$ , are determined by the freshwater fluxes  $F_1$  and  $F_2^{tot}$  and the salinity level of the water flowing into the box  $i$ .

$$\dot{S}_1 = \frac{m}{V_1} (S_4 - S_1) + \frac{S_0 F_1}{V_1}, \quad (11)$$

$$\dot{S}_2 = \frac{m}{V_2} (S_3 - S_2) + \frac{S_0 F_2^{tot}}{V_2}, \quad (12)$$

$$\dot{S}_3 = \frac{m}{V_3} (S_1 - S_3) + \frac{S_0 (F_1 - F_2^{tot})}{V_3}, \quad (13)$$

$$\dot{S}_4 = \frac{m}{V_4} (S_2 - S_4), \quad (14)$$

where  $F_1$  and  $F_2^{tot}$  are multiplied by  $S_0$  for conversion to a salt flux.

The meridional volume transport  $m$  evolves according to the north-south temperature gra-

dient and salinity gradient:

$$m(t) = k (\beta (S_2(t) - S_1(t)) - \zeta (T_2(t) - T_1(t))), \quad (15)$$

where  $\zeta$  and  $\beta$  are thermal and haline expansion coefficients and  $k$  is a hydraulic constant. Note that the system of differential equations (7)-(14) is non-linear, because  $m$  also depends on temperature and salinity.

All parameters of the THC model, which is here presented in a condensed form, are derived by Zickfeld et al. (2004) from simulations with an Earth-system model of intermediate complexity.

The variable that the THC model hands over to DICE, the meridional volume transport  $m$ , is an essential part of the new climate damage function. In this model, the damage cost function consists of two parts: the first part relates to the (non-catastrophic) damages caused by global warming and the second and new part refers to the costs incurred by a substantial slow-down of the THC. For the first part, we adopt the same quadratic functional form as in DICE. Note that the commonly used cost function in DICE only encapsulates the impacts by increasing temperatures, but it also specified to capture all resulting losses comprising those of the non-catastrophic and the catastrophic impacts. As our model framework already accounts for one of these catastrophes explicitly, we choose to recalibrate the damage function in DICE according to the non-catastrophic damage estimates given by Cline (1992). After employing a two-point fit to these estimates, we obtain the following damage cost function:

$$D(\Delta T, m) = 0.000976(\Delta T)^2 + f_{THC}(m) \quad (16)$$

in percentage of global GDP, where  $f_{THC}(m)$  refers to the yet unspecified damage cost function of a slow-down in the THC. As with all climate damage functions, developing  $f_{THC}(m)$  is a challenging task. This damage function needs to mirror the sudden cost explosion incurred by the tipping point in the THC system. We therefore postulate that the damages are (close to) zero as long as  $m$  is higher than some critical flow strength. It can be shown analytically that this critical value is less than half of today's value of 22.6 Sv (Zickfeld and Bruckner (2003) and Zickfeld et al. (2004)). Zickfeld and Bruckner (2003) and Zickfeld et al. (2004) also account for transient effects and conclude that 10 Sv is a reasonable estimate for this tipping point value.<sup>5</sup> When  $m$  approaches this threshold, the damage function  $f_{THC}(m)$  abruptly increases up to  $d_m$ ,  $0 < d_m < 1$ , a fraction of global GDP for all lower values of  $m$ . That all lower values of  $m$  incur almost the same damage costs is demonstrated by simulations using the FUND model in Link and Tol (2011). A specification complying with

---

<sup>5</sup>Strictly speaking, the location of the tipping point depends on (amongst other factors) the rate of temperature increase. Still, for a variety of different experiments with THC models, the critical weakening appears on average at approximately 10 Sv.

the above description is:

$$f_{THC}(m) = \frac{d_m}{\pi} \left( \frac{\pi}{2} + \text{ArcTan}[100(10.5 - m)] \right) \quad (17)$$

where  $\pi$  is the mathematical constant  $\approx 3.14159$ . Figure 3 illustrates this specific function for the yet to be calibrated value  $d_m$ .

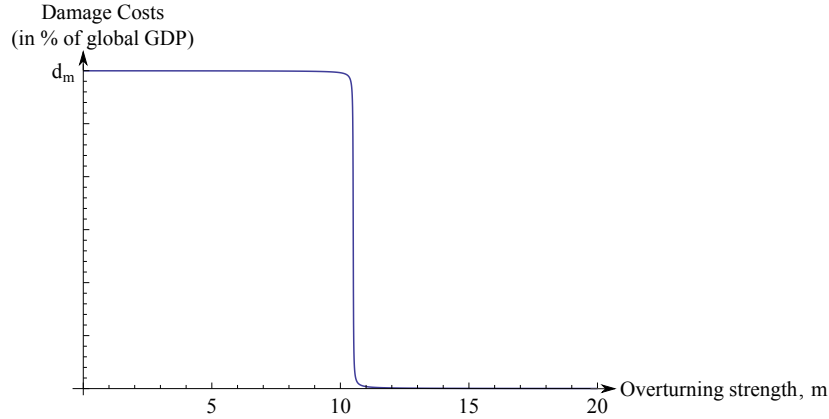


Figure 3: The damage cost function  $f_{THC}(m)$  for any  $d_m$ .

As the weakening of the THC not only depends on the levels of temperature but also on the rate of temperature change, the climate damage cost function in equation (17) is also determined by the rate of climate change, which cannot be captured by the commonly used damage function in DICE and DICE-CJL.

As the analytical solution is not available for this welfare optimization problem, the solution is numerically approximated using the General Algebraic Modeling System (GAMS, McCarl (2013)). Please refer to the Appendix for further information.

In the next section, we turn to investigate this model and its implications concerning the decision maker's risk preferences.

### 3 Numerical Simulation

Before investigating the simulations of the model described above, we direct special attention to the calibration of some crucial parameters. First, we specify the exact nature of the modelled uncertainty concerning the climate sensitivity value. Then, we proceed with calibrating the parameters which exhibit immense uncertainty. These include the global damage costs,  $d_m$ , caused by a critical weakening of the THC and the value of the North Atlantic hydrological sensitivity,  $h_2^{tot}$ . Afterwards, we explain our choices for the rate of pure time preference and the intertemporal elasticity of substitution. After providing a thorough analysis of the baseline case and a comprehensive sensitivity analysis, we direct our attention to the effects of not having unlimited abatement options. Further information is provided in the

Appendix. Appendix A tests whether our model can reproduce the simulations undertaken by Ackerman et al. (2013) and illuminates the differences between these models. Appendix B discusses the caveats of the numerical solution of the problem. Appendix C summarizes the entire calibration.

The five possible states of nature are given by a selection of climate sensitivity values and their associated probability. Figure 4 illustrates the proposal of Ackerman et al. (2013) for how these values can be derived. Assuming the probability distribution function (PDF) given by Roe and Baker (2007), they partition the cumulative probability function (CDF) into five unequal intervals spanning 50, 40, 5, 3 and 2 per cent of the whole interval, respectively. The five states of nature are then derived by taking the midpoints of these intervals as the percentiles of the distribution. For instance, the midpoint of the 50 per cent interval is 25 per cent, the 25th percentile climate sensitivity value is 2.43. The probability of 2.43 being the true state of nature is then assumed to be given by the length of the interval, i.e. 50 per cent. Taking everything together, the following climate sensitivity values evolve as the possible states of nature  $S$  with a probability  $p$ :  $S_1 = 2.43$ ,  $p_1 = 0.5$ ;  $S_2 = 3.76$ ,  $p_2 = 0.4$ ;  $S_3 = 6.05$ ,  $p_3 = 0.05$ ;  $S_4 = 8.2$ ,  $p_4 = 0.03$ ;  $S_5 = 16.15$ ,  $p_5 = 0.02$ .

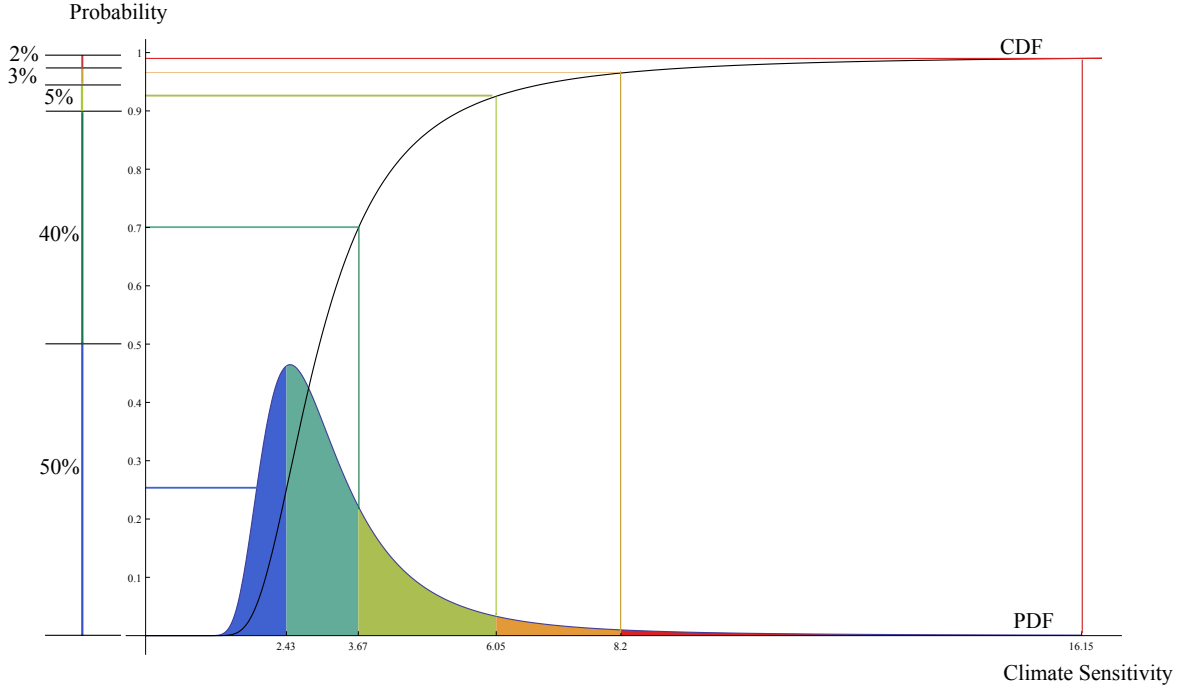


Figure 4: Partitioning of the probability distribution for climate sensitivity

The assessments of the potential costs,  $d_m$ , are an area of active research and mainly include changes in regional temperatures and in precipitation patterns, a reduced carbon uptake in the Atlantic and a rise in the sea level. However, they ignore several other possible impacts, such as a deterioration in food quality (Kuhlbrodt et al. (2009) and Zickfeld et al. (2007)). Furthermore, only few monetized estimates are available in the literature. Only accounting

for the effects on regional temperatures, Tol (1998) finds that the losses may range from zero to three per cent of global GDP. More recently, however, Link and Tol (2011) provide an estimate of 0.5 per cent of global GDP. The partial coverage of the impacts may systematically underestimate the actual costs. Accordingly, significantly higher cost estimates cannot be ruled out due to further consequences such as strain resulting from geopolitical tensions (Mastrandrea and Schneider (2001)). On the basis of these studies, we consider a baseline calibration of  $d_m = 0.03$  a reasonable choice and test the sensitivity of the results with respect to  $d_m$  being 0.015 and 0.045 later on.

The North Atlantic hydrological sensitivity,  $h_2^{tot}$ , is one of the factors that determines the magnitude of the changes in  $m$ . As explained by equation (6), the hydrological sensitivity  $h_2^{tot}$  measures the additional increase in freshwater for a warming of one degree Celsius, including the meltwater runoff from the Greenland ice sheet and the meltwater from the Arctic sea ice. The quantification of  $h_2^{tot}$  involves great uncertainties as it rests on largely differing estimates of evaporation, precipitation, river runoff and meltwater volumes (Rahmstorf and Ganopolski (1999) and Zickfeld et al. (2004)). The identified estimates for  $h_2$  range from  $0.01 \text{ Sv } ^\circ\text{C}^{-1}$  to  $0.053 \text{ Sv } ^\circ\text{C}^{-1}$  (Manabe and Stouffer (1994) and Zickfeld et al. (2004)). The contribution of the Greenland ice sheet to  $h_2^{MW}$  is found to be between  $0.002 \text{ Sv } ^\circ\text{C}^{-1}$  and  $0.01 \text{ Sv } ^\circ\text{C}^{-1}$  and the melting of the sea ice adds approximately  $0.0125 \text{ Sv } ^\circ\text{C}^{-1}$  (Zickfeld et al. (2004)). Taking all these estimates into account, we choose to set  $h_2$  to  $0.03 \text{ Sv } ^\circ\text{C}^{-1}$  and  $h_2^{MW}$  to  $0.0225 \text{ Sv } ^\circ\text{C}^{-1}$ , adding up to  $h_2^{tot}$  at  $0.0525 \text{ Sv } ^\circ\text{C}^{-1}$ . With the diverging estimates in mind, Zickfeld et al. (2004) also test the THC dynamics for alternative values ranging from  $0.013 \text{ Sv } ^\circ\text{C}^{-1}$  up to  $0.06 \text{ Sv } ^\circ\text{C}^{-1}$ . They find significant differences in the evolution of the overturning strength and thereby illustrate the key role of the uncertainty regarding  $h_2^{tot}$  in the predictions. Consequently, we also allow for alternative values being  $h_2^{tot} = 0.04$  and  $h_2^{tot} = 0.06$ .

The question of how to calibrate the parameters in the Epstein-Zin utility specification in the context of a climate policy decision remains unanswered. It is impossible directly to retrieve the data describing the preferences for such a decision. Furthermore, it is not clear which market is subject to a similar multitude of risks and could thus serve to provide helpful evidence for a reasonable calibration. One approach, which is also followed by Ackerman et al. (2013), is to obtain the parametrization from the studies of Bansal and Yaron (2004) and Vissing-Jørgensen and Attanasio (2003), which estimate the intertemporal elasticity of substitution (IES) at 1.5 and the coefficient of relative risk aversion at 9.5 to 10. We borrow these values for our study as well. Later on, the sensitivity with respect to the IES parameter is examined by implementing the value of 2.

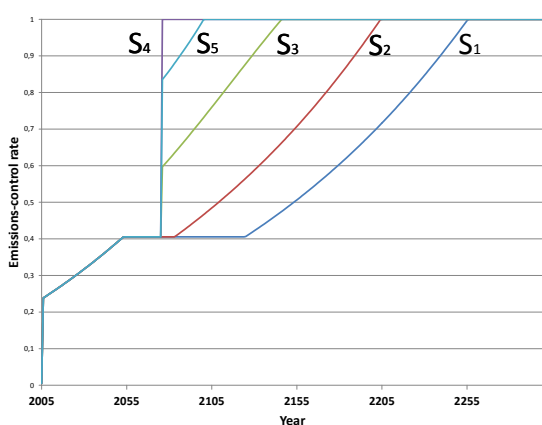
To evaluate the effects of risk aversion on near-term policy, we compare the results for two specifications of the risk aversion parameter  $\alpha$ . The first, with  $\alpha$  being 2, describes a mildly risk-averse policy maker. The alternative specification of  $\alpha = 10$  is – as already mentioned

– taken from the above-mentioned studies. The comparison of the resulting policy scenarios gives evidence of the sensitivity of the results with respect to the degree of risk aversion.

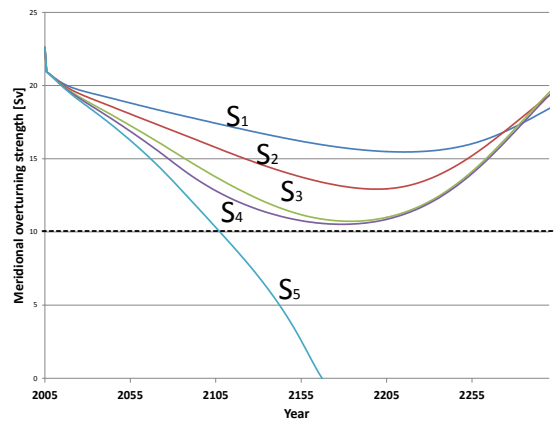
All figures in what follows provide this comparison by contrasting the results for the optimal emission control path and the implied dynamics of the THC. The simulations shown at the top of each figure rest on the underlying assumption of  $\alpha$  being 2, while those at the bottom result from  $\alpha$  being 10. The general idea behind the graphs on the left (i.e. all graphs (a) and (c)) displaying the optimal reduction policy is as follows. As described in section 2, the policy maker decides on one policy path until 2075, the year in which the true climate sensitivity value is revealed. This policy path factors in all possible climate sensitivity values and their associated probabilities. Depending on which value is disclosed, the policy maker chooses one of the five policy paths for the time afterwards. The consequences for the overturning strength, illustrated for all five climate sensitivity values, are presented by the graphs on the right-hand side (i.e. figures (b) and (d)).

The simulations in Figure 5 are obtained by employing the baseline calibration. Figure 5a shows that it is optimal to start with an emission control rate of 23.8 per cent and gradually to increase efforts up to 40.55 per cent in 2075. If a climate sensitivity value of  $S_1 = 2.43$  is disclosed, it is optimal to maintain this control rate for several more decades. For all other values, efforts have to be increased – the more so the higher this value turns out to be. In the case of  $S_5$ , this would even imply a sudden jump to the control rate of 83.54 per cent. Yet, these massive but late efforts in  $S_5$  cannot prevent critical weakening, which occurs when crossing the dashed line at 10 Sv. Accordingly, we may note the following. Apparently, it is not optimal to increase near-term abatement to avoid a collapse in  $S_5$ . A possible explanation may be that the probability of being in state 5 is too small to exert a significant effect on near-term policy. Although in 2075 it is too late to prevent the THC collapse, the increased efforts may be justified in view of the non-catastrophic damage costs (i.e. the first term in equation (16)). Furthermore, these efforts may help to delay the collapse. In all other states, the late policy efforts succeed in safeguarding the THC. Hence, the policy maker has enough time to wait until 2075. Then, the policy maker can adjust policy and can prevent shut-down in the more likely states. Interestingly, the late policy efforts in  $S_4$  are even more intense than in  $S_5$ . To understand this result, we forced the control rate in  $S_5$  to jump to 100 per cent after learning. This simple test revealed that the collapse in  $S_5$  can only be delayed but not prevented by these late efforts. In this situation, optimal policy after 2075 is about weighing the costs against the benefits of postponing the collapse. In contrast, in  $S_4$ , emission control is capable of preventing the critical weakening. Consequently, the differences in the benefits of emission control give the reason for a higher control rate in  $S_4$  than in  $S_5$ .

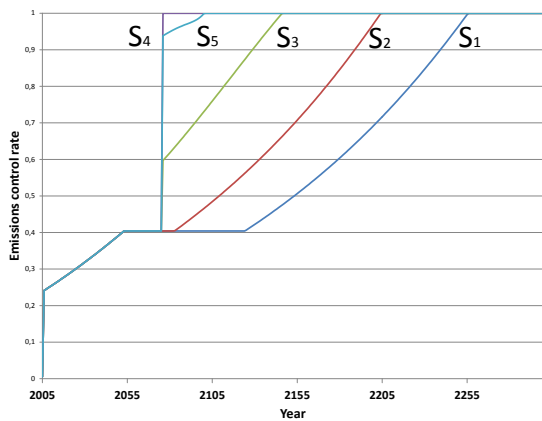
The optimal policy for a more risk-averse decision maker is presented in Figure 5c. It is striking that this policy path is highly similar to that described before for a mildly risk-averse decision maker. The critical weakening in  $S_5$  occurs almost at the same time. Hence,



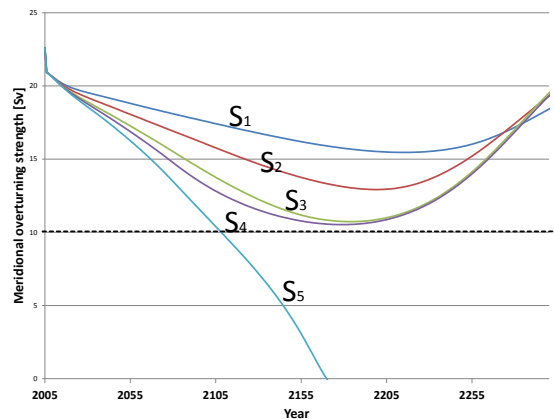
(a) Optimal emission reduction policy



(b) Possible trajectories of the overturning strength



(c) Optimal emission reduction policy



(d) Possible trajectories of the overturning strength

Figure 5: Comparison of the effects of alternative assumptions concerning risk aversion:  $\alpha = 2$  (top) and  $\alpha = 10$  (bottom)

the assumption of higher risk aversion neither changes the near-term policy in a qualitatively meaningful way nor does it affect the THC dynamics to any great extent. It only increases the late efforts in  $S_5$ . We may draw the conclusion that an even more risk-averse policy does not rate the possible, albeit not very likely, detrimental impacts by  $S_5$  as sufficiently high to make more decisive efforts in the near term. Next, we explore the sensitivity of this result with respect to a change in some key parameters. Table 3 summarizes the variations in the parameters described above.

The effects of a lower hydrological sensitivity,  $h_2^{tot}$ , are demonstrated in Figure 6. A lower  $h_2^{tot}$  value means a smaller freshwater input per degree of warming, which translates into a reduced rate of weakening. Accordingly, compared to Figure 5a, Figure 6a exhibits a lower control rate of 34.02 per cent in 2075. The curves illustrating the policy response after 2075 also indicate less stringent efforts. Only in  $S_5$  does the control jump to a higher rate, this

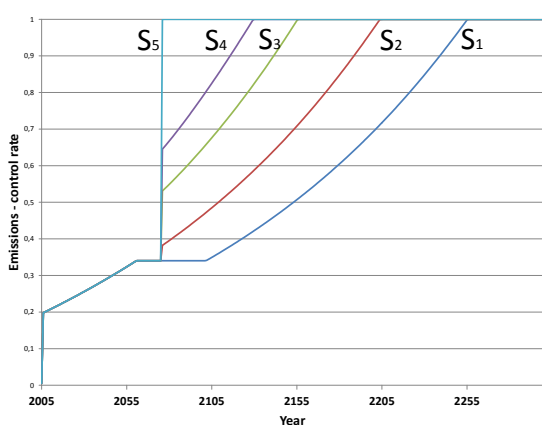
Model Run	$h_2^{total}$	$d_m$ in %	IES	Year of Learning	Figure
Baseline	0.0525	3.0	1.5	2075	5
Lower $h_2$	0.04	3.0	1.5	2075	6
Higher $h_2$	0.06	3.0	1.5	2075	7
Lower $d_m$	0.0525	1.5	1.5	2075	8
Higher $d_m$	0.0525	4.5	1.5	2075	9
Higher IES	0.0525	3.0	2	2075	10
Early Learning	0.0525	3.0	1.5	2050	11
Late Learning	0.0525	3.0	1.5	2100	12

Table 1: Sensitivity runs

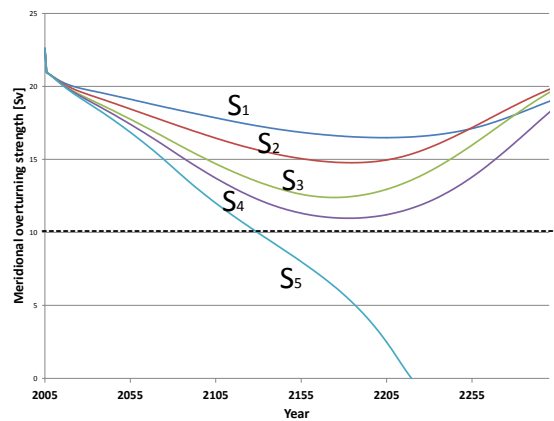
being 100 per cent. A possible reason is that compared to the baseline case, it proves easier to delay the collapse even further. In total, the shut-down occurs almost 50 years later in Figure 6b than in Figure 5b. A higher risk aversion again appears to be of subordinate importance for near-term policy. In Figure 6c, the control rate in 2075 is only 0.17 percentage points higher than in Figure 6a. In addition, this policy does not have a visible effect on the THC dynamics. This result makes sense, as a lower hydrological sensitivity makes the THC dynamics more stable and thus the risk of a possible collapse is smaller. This offers even more leeway to postpone decisive action.

The accelerated weakening of the THC caused by a higher hydrological sensitivity,  $h_2^{tot}$ , is counteracted by more abatement in the near term, as demonstrated in Figure 7a. In all possible states  $S_1$ – $S_5$ , the risk of a collapse increases and thus requires that more emission reduction is undertaken. Having implemented this high control rate, the policy maker does not have to adjust the abatement policy for many decades if  $S_1$  turns out to be true. Even in  $S_2$ , this rate is maintained for some time. As in Figure 5, there are higher efforts in  $S_4$  than in  $S_5$ . Although, the higher  $h_2^{tot}$  leads to a critical weakening in  $S_4$  and in  $S_5$ , it is reasonable to cut emissions sooner after learning in  $S_4$ . The reason is given by Figure 7b. Emission reduction helps the THC to recover after some time in  $S_4$ , while in  $S_5$  it will irreversibly collapse. Although the risk of a complete shut-down is more pronounced in this scenario, higher risk aversion does not lead to greater near-term efforts. Consequently, the probability of  $S_4$  or  $S_5$  turning out to be true is too low to affect policy until 2075. For all other states, the policy maker has all means to react in time and to prevent the critical weakening. Hence, risk aversion does not lead to qualitatively different results.

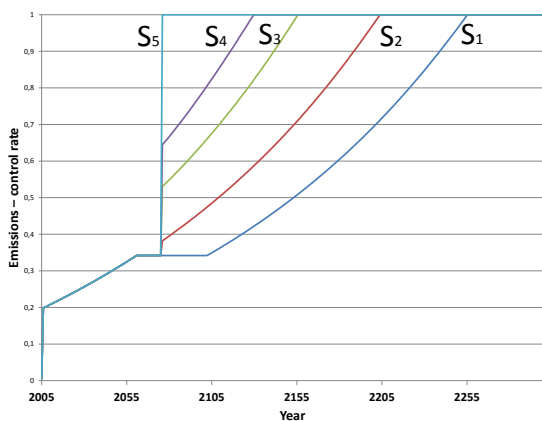
Figure 8a tells us how policy changes if the damages from a THC collapse are less severe. It is not surprising that policy until 2075 is a little less stringent than illustrated in Figure 5a. Although damages are lower, efforts are increased after 2075 and in four of the five cases the irreversible collapse can be prevented. In  $S_4$ , the THC weakens critically but soon recovers. As shown by Figure 8c, risk aversion has some, albeit small, effect on near-term policy. The



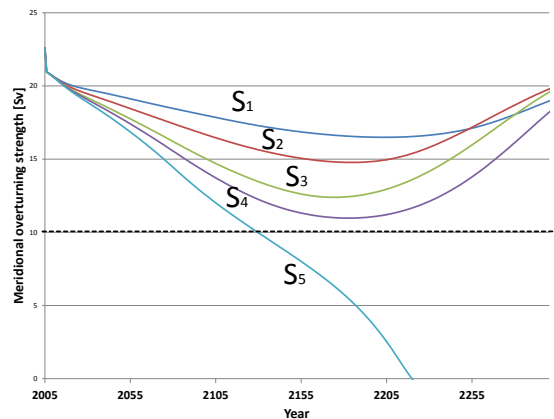
(a) Optimal emission reduction policy



(b) Possible trajectories of the overturning strength



(c) Optimal emission reduction policy



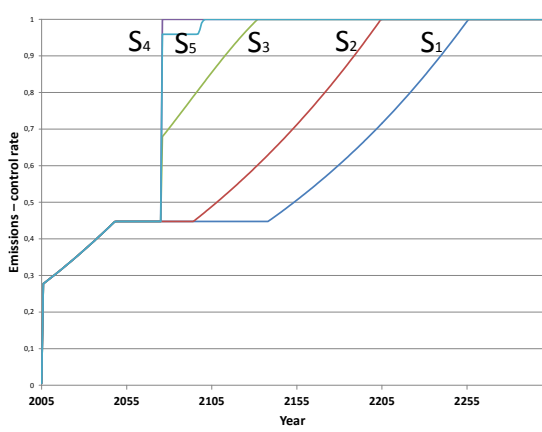
(d) Possible trajectories of the overturning strength

Figure 6: Sensitivity of the results with respect to a lower hydrological sensitivity value i.e.  $h_2^{total} = 0.04$

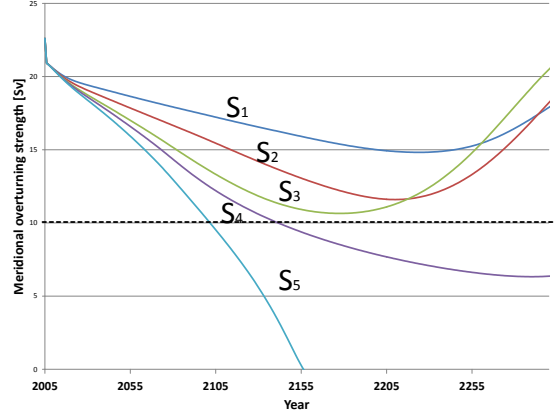
increased efforts prevent  $S_4$  from slowing down beyond 10 Sv.

Next we explore the implications of higher THC-specific damage cost case as depicted in Figure 9. Intuitively, the optimal control must be at least as strong as that in the baseline case. In Figure 5a, the decision maker successfully reduces the risk of the potential breakdown and prevents the strong weakening in the states of nature  $S_1$ – $S_4$ , while taking the risk of the critical weakening in the low probability state  $S_5$ . Consequently, the optimal abatement trajectory in the near term is virtually the same as in Figure 5a. Yet, in state  $S_5$ , higher damage costs imply a more stringent policy after learning of 13.1 percentage points. As a result, 96.64 per cent of emissions are under control by the year 2076. To reduce the long-term risk further is not economically justified and thus higher risk aversion does not have a significant effect.

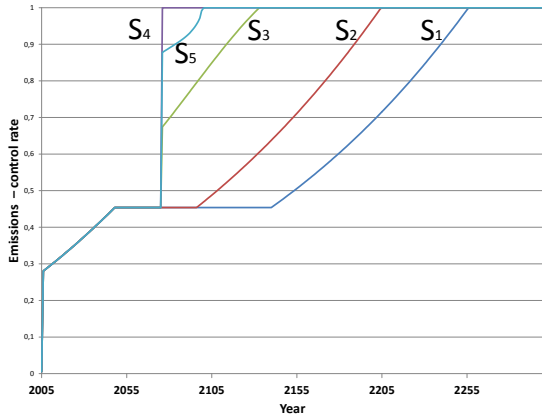
The effects of an increase in the intertemporal elasticity of substitution are demonstrated in Figure 10. As expected, a higher IES leads to stronger abatement efforts in the near term



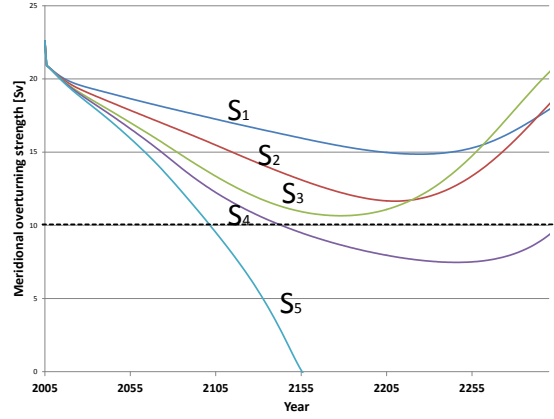
(a) Optimal emission reduction policy



(b) Possible trajectories of the overturning strength



(c) Optimal emission reduction policy

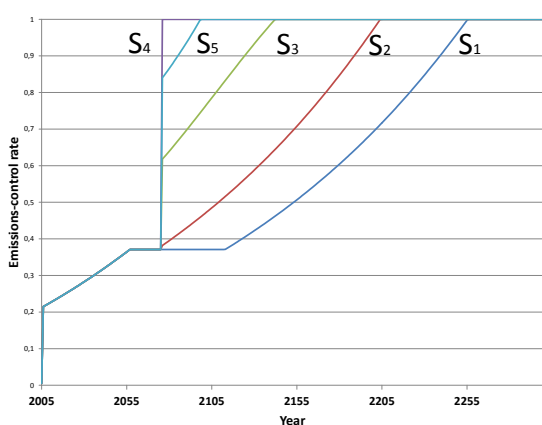


(d) Possible trajectories of the overturning strength

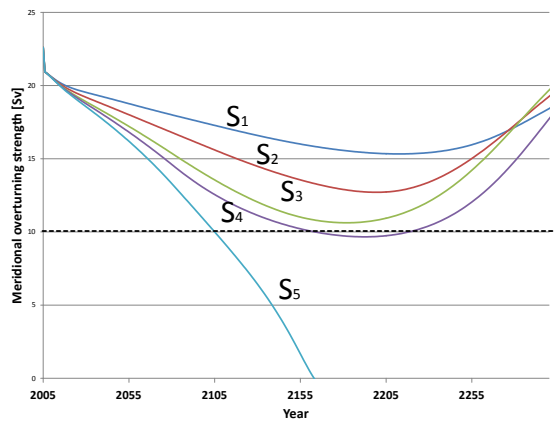
Figure 7: Sensitivity of the results with respect to a higher hydrological sensitivity value, i.e.  $h_2^{total} = 0.06$

(cf. Figure 10a with Figure 5a). Under higher risk aversion, the control rate is almost the same. As before, the collapse is only prevented in the more likely states,  $S_1$ – $S_4$ . Although a lower discount rate does not alter the implied THC dynamics significantly, it has an effect on the policy path because the future damages are discounted to a lesser extent.

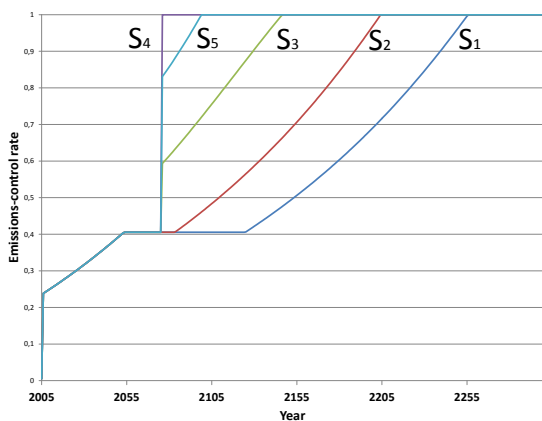
Figure 11 depicts the case in which uncertainty resolves rather early, i.e. in 2050. It is interesting to see that the opportunity to implement a perfectly tailored solution rather early in time does not significantly alter the outcome for the THC dynamics. Comparing Figure 11b with Figure 5b shows that the critical weakening in  $S_5$  occurs almost at the same time. The early stringent efforts in  $S_5$  as described by Figure 11a only help to delay the total collapse by approximately 50 years. Owing to the knowledge of early learning, the level of risk aversion is expected to have even less effect than in the baseline case. As indicated by Figure 11c, a higher level of risk aversion increases the abatement efforts until 2050 only



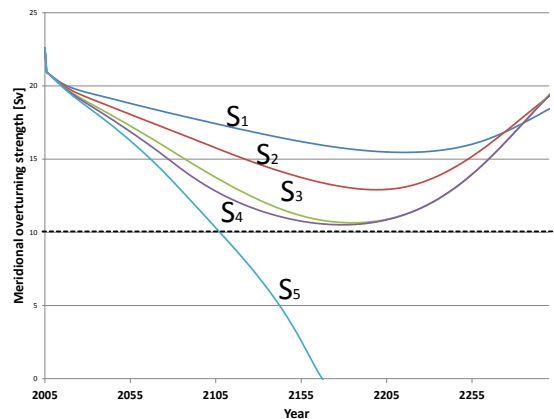
(a) Optimal emission reduction policy



(b) Possible trajectories of the overturning strength



(c) Optimal emission reduction policy

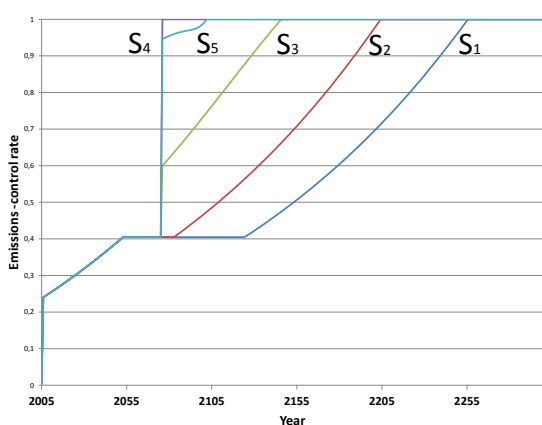


(d) Possible trajectories of the overturning strength

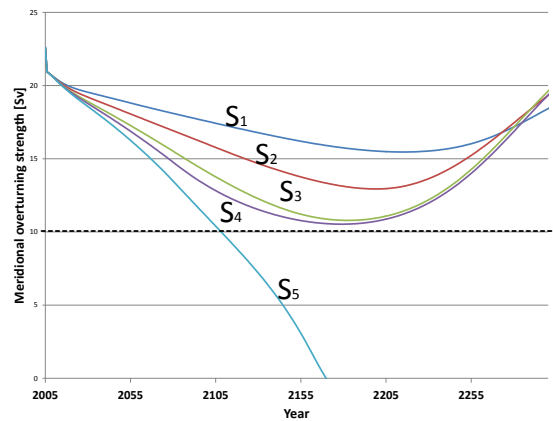
Figure 8: Sensitivity of the results with respect to lower damage costs caused by a THC collapse, i.e.  $d_m = 0.015$

negligibly. Risk aversion only has an effect on policy after learning if the worst case,  $S_5$ , turns out to be true.

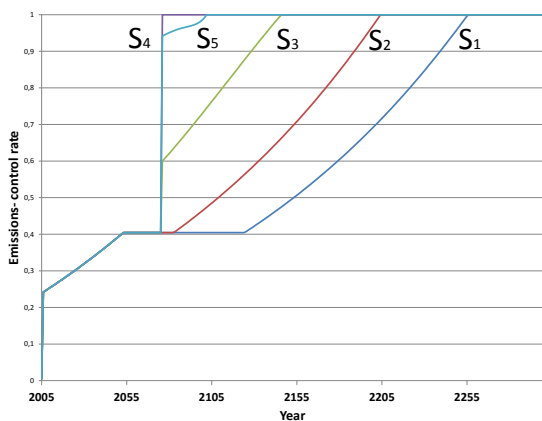
How do policy and the THC dynamics change if learning is known to occur later? Figure 12 assumes that the year of learning is 2100. Figure 12b demonstrates that the THC weakens strongly in state  $S_4$ , but it does recover after several decades. The implementation of abruptly increased efforts after 2100 is not timely enough as the collapse in  $S_5$  and the extreme weakening in  $S_4$  are already locked in. This is evidence of how limited the strategy to react after learning can be, as it may increase the likelihood of tipping points. It also emphasizes the importance of early learning. Interestingly, a more risk-averse policy maker increases the near-term efforts only slightly. The reason for this is that the two states,  $S_4$  and  $S_5$ , are only weighted by 5 per cent in the welfare optimization calculus. The simulation gives evidence that this probability is too small to have a significant impact on decision making towards longer term impacts.



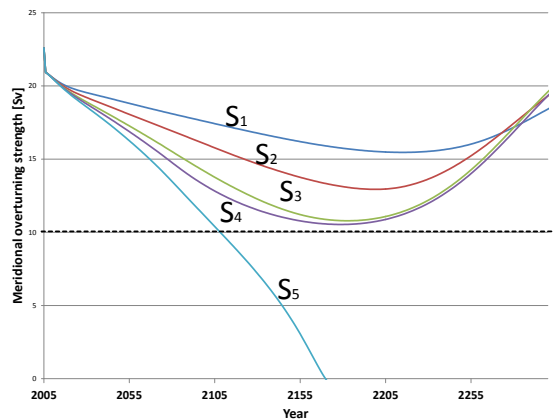
(a) Optimal emission reduction policy



(b) Possible trajectories of the overturning strength



(c) Optimal emission reduction policy



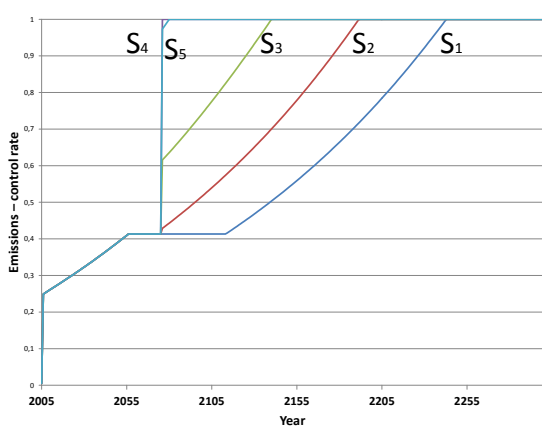
(d) Possible trajectories of the overturning strength

Figure 9: Sensitivity of the results with respect to higher damage costs caused by a THC collapse, i.e.  $d_m = 0.045$

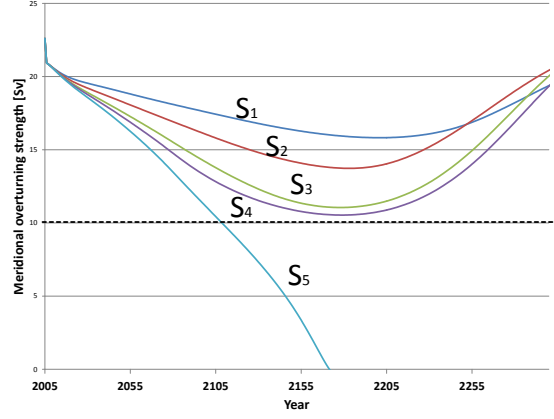
Please note that in the first set of simulations (Figures 5–12) no restrictions for the maximum change in the annual emission reduction rate are assumed. The policy maker can instantaneously raise the emission reduction rate up to 100 per cent if needed. However, the assumption that the technology to stop emitting immediately after 2075 will be available may be criticized as a naive faith in technology. What happens to policy and to the implications in relation to risk aversion if the abatement options are more restrictive than in the original DICE model? In the remainder of this section, we examine how limits on the speed of policy adjustment affect near-term optimal policy. This is expressed by an additional constraint in the model:

$$\mu(t+1) \leq \mu(t) + l, \quad (18)$$

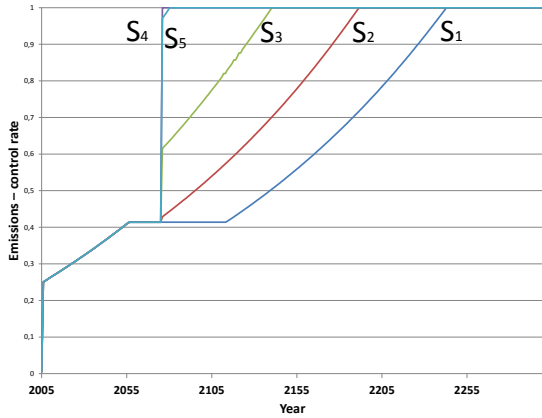
where  $\mu$  is the emission reduction rate and  $l$  the maximum possible increase in abatement



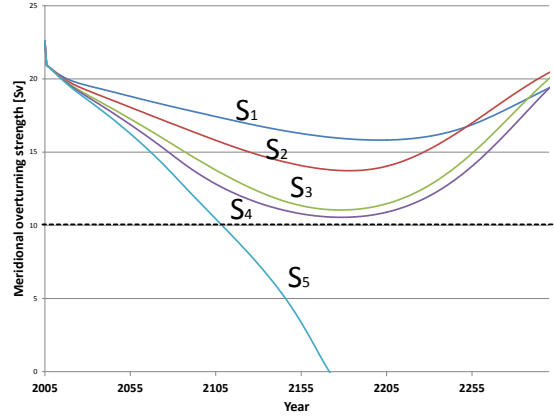
(a) Optimal emission reduction policy



(b) Possible trajectories of the overturning strength



(c) Optimal emission reduction policy



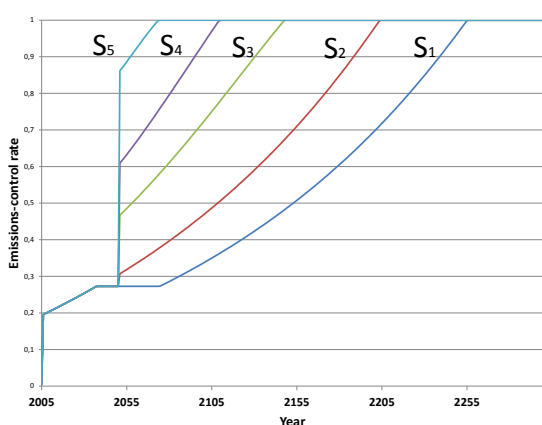
(d) Possible trajectories of the overturning strength

Figure 10: Sensitivity of the results with respect to a higher IES of 2%

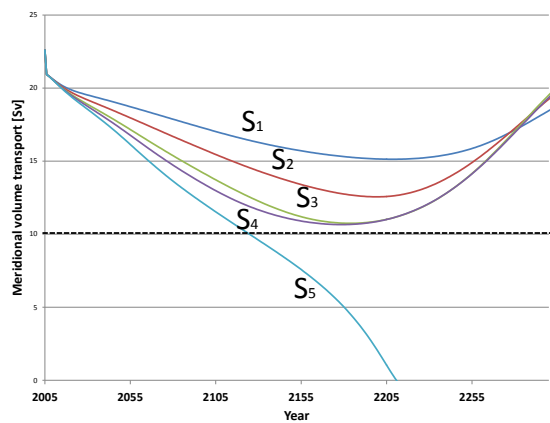
each year. We consider two cases for  $l$  with the purpose of showing the effects of differing limitations in the abatement options:  $l_1 = 0.5$  and  $l_2 = 0.25$ . Please note that these numbers are chosen to provide a good understanding of how  $l$  may influence the model. They are not derived from studies in the respective field of research.

Figure 13 provides insights into how the model responds to a 50 percentage limit on the maximum possible increase in abatement. This limit is not very restrictive considering that starting from the business as usual scenario, the policy maker needs two periods instead of one period to curb emissions fully. Comparing Figure 13a to the corresponding baseline case, Figure 5a, we see that the limit causes the near-term abatement path to increase. In 2075, the control rate is 0.4077, which is 0.22 percentage points higher than in the baseline case. Risk aversion does not amplify this effect on the near-term policy significantly as visualized by Figure 13c.

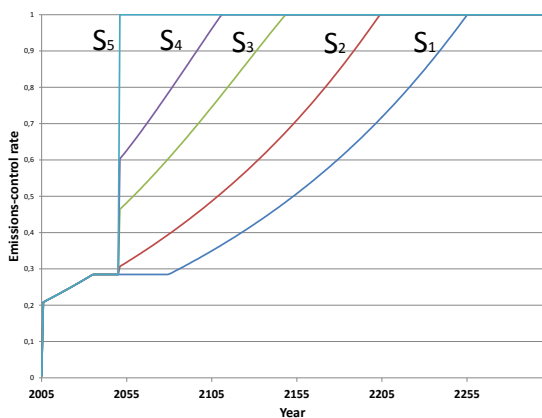
In Figure 14, more restrictive abatement options are assumed with  $l = 0.25$ , knowing that it



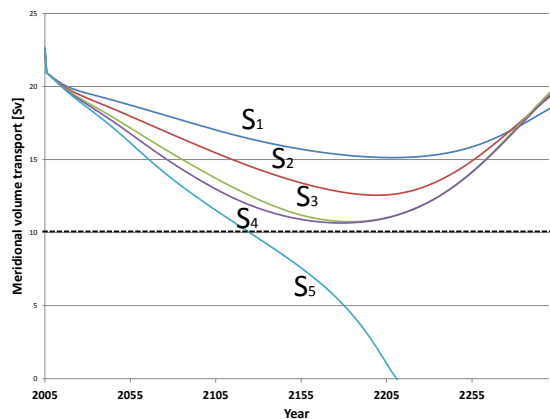
(a) Optimal emission reduction policy



(b) Possible trajectories of the overturning strength



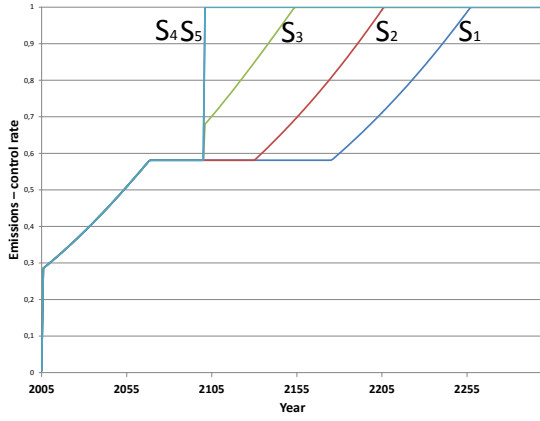
(c) Optimal emission reduction policy



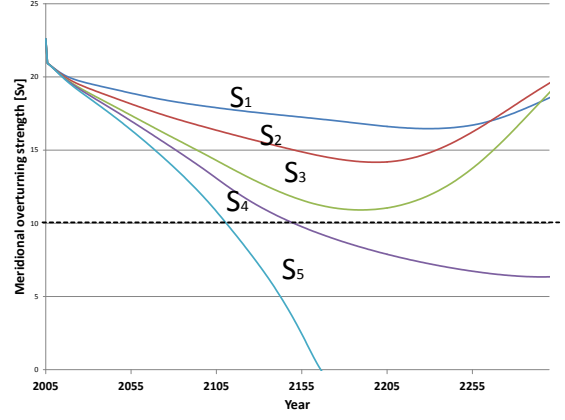
(d) Possible trajectories of the overturning strength

Figure 11: Sensitivity of the results with respect to learning early, i.e. in 2050

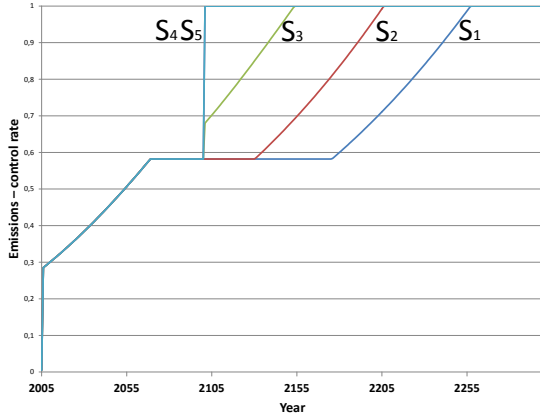
takes slightly longer to adjust policy after learning. The limitations in the policy adjustment after 2075 increase the near-term efforts by one percentage point. The qualitative behavior of the THC is not significantly different from that in Figure 5b. This makes sense as the lower the annual change in emissions control, the higher the rate of temperature increase and thus the higher rate of weakening. In other words, in overestimating the speed of success, we may in turn underestimate the near-term abatement necessary to maintain the stability of the North Atlantic meridional overturning circulation.



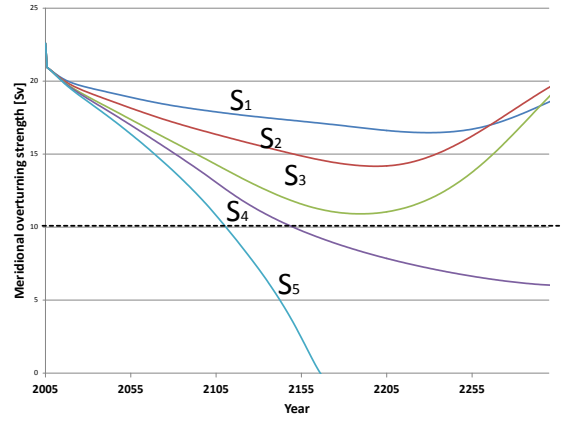
(a) Optimal emission reduction policy



(b) Possible trajectories of the overturning strength

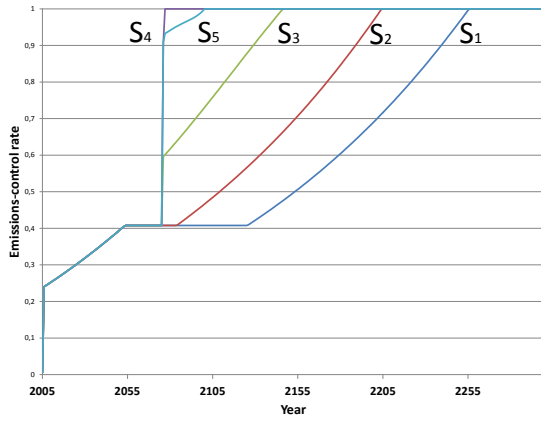


(c) Optimal emission reduction policy

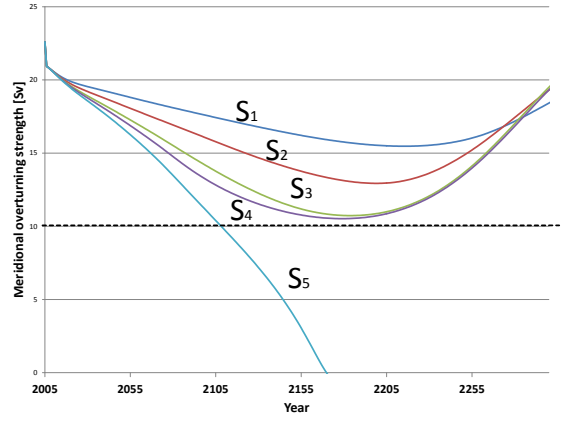


(d) Possible trajectories of the overturning strength

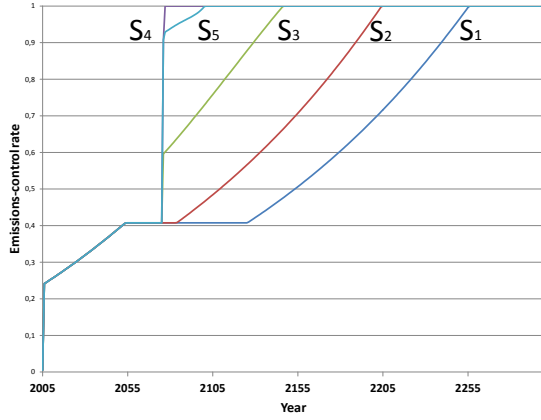
Figure 12: Sensitivity of the results with respect to learning later, i.e. in 2100



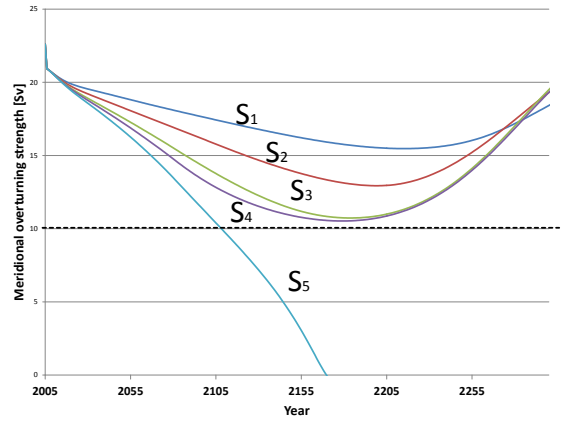
(a) Optimal emission reduction policy



(b) Possible trajectories of the overturning strength

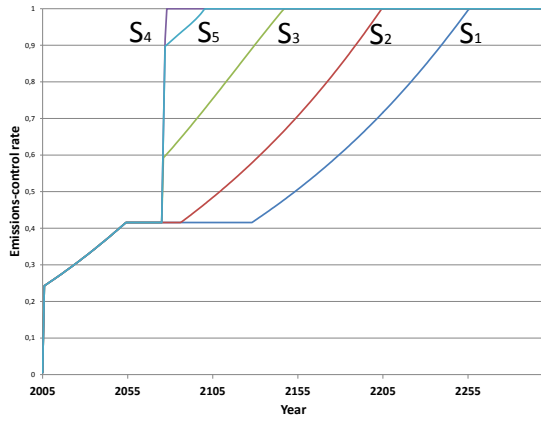


(c) Optimal emission reduction policy

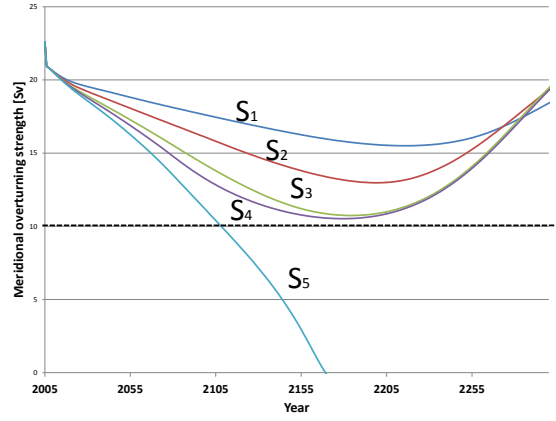


(d) Possible trajectories of the overturning strength

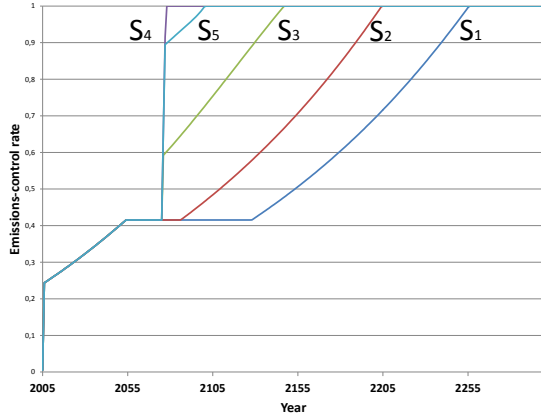
Figure 13: Sensitivity of the results with respect to limited abatement options, i.e.  $l = 0.5$



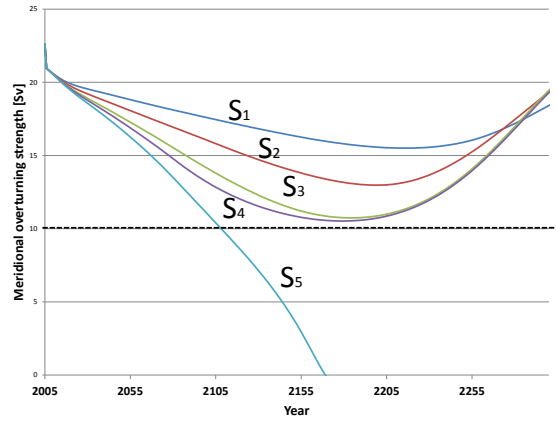
(a) Optimal emission reduction policy



(b) Possible trajectories of the overturning strength



(c) Optimal emission reduction policy



(d) Possible trajectories of the overturning strength

Figure 14: Sensitivity of the results with respect to limited abatement options, i.e.  $l = 0.25$

## 4 Conclusion

In the most common IAMs, higher risk aversion does not imply stronger optimal climate policy efforts in the next few decades. This result appears counterintuitive as in particular the looming climate catastrophes pose one of the greatest threats of climate change. Accordingly, optimal policy should be susceptible to differing concerns regarding these risks. One cause of this problem is the assumption of the CRRA utility specification, as it leads to an interwoven representation of the decision maker’s preferences for time and risk. However, as demonstrated by Ackerman et al. (2013), the puzzle is not solved by replacing the CRRA utility by a specification that disentangles these preferences, as shown for the Epstein-Zin utility specification. The reasons that have been conjectured draw upon the model’s strong assumptions on abatement and the insufficient representation of catastrophic risk. Accordingly, if the policy maker has the means at his or her disposal to avoid all risk after learning, risk aversion is irrelevant for near-term climate policies.

To advance the investigation of this puzzle, we fully couple a similar version to the Epstein-Zin-DICE model employed by Ackerman et al. (2013) with a dynamic model of the Atlantic thermohaline circulation. The dynamic model of the THC overturning strength promises a more refined representation of climate change risk.

Despite the model extension, the sensitivity of the optimal near-term policy with respect to risk aversion is qualitatively similar to the results obtained by Ackerman et al. (2013). However, in contrast to Ackerman et al. (2013), the extended model allows us to monitor the response of a specific climate system at risk. As a result, we can infer how the associated risk of its vulnerability evolves. Hence, we can pinpoint the characteristics of this specific risk element that may lead to the insensitivity result.

Consistent with Ackerman et al. (2013), we recognize that risk aversion is of little importance if the policy maker is assumed to have all means at his or her disposal to adjust emission control instantaneously to the desired rate after learning. Late but fine-tuned efforts suffice to prevent the THC from collapsing in the more likely climate sensitivity values. The risk that could influence near-term optimal policy is only caused by a very high climate sensitivity value. In this value the near-term optimal policy efforts by a mildly risk-averse decision maker are too late to prevent the THC from shutting down. However, the value accounts only for 2 per cent in welfare optimization. The simulations prove that even a more risk-averse policy maker does not place that much importance on the risk associated with this value and near-term abatement is only slightly increased.

A thorough sensitivity analysis shows that this result for risk aversion is highly robust to a wide choice of values within reasonable ranges. Even if learning occurs very late, risk aversion does not increase policy efforts significantly and it does not help prevent the THC from collapsing in a further high climate sensitivity value.

Finally, we direct attention to the reasons for this little effect of risk aversion. For the likely climate sensitivity values, this tipping point is rather distant in the future, which allows late action. Moreover, as indicated by the sensitivity analysis, the ramifications caused by a collapse are great but not great enough to justify taking precautionary measures to reduce the risk in the unlikely states that would put too high a burden on the economy. In sum, there are three reasons for the small effect of risk aversion in this fully coupled extended the DICE model: the long period until the risk materializes, the relatively low costs incurred by a weakening and the low probability associated with the risk.

## Appendices

### A Robustness check

As this research clearly relates to the study by Ackerman et al. (2013), we here clarify whether our model could reproduce their results. We may demonstrate this for their  $E_1$  model run. We keep the default damage function of the original DICE model, i.e. we set  $a_2 = 0.0028388$  and ignore the feedback from the THC module to DICE. Please note that there are different assumptions in the two models. First, we apply the updated transient temperature change equation from Cai et al. (2012a), which is shown to result in a lower optimal emissions control rate. In contrast to the equation in Nordhaus (2008) and Ackerman et al. (2013), in the updated equation the global mean temperature is not affected by future atmospheric carbon concentration. Second, we refine the discretization of the time steps from 10 to 1 years. Comparing the results of the two models, we recognize that the assumptions in our model offer a more differentiated view of the optimal policy. While in Ackerman et al. (2013) only one policy path is given for three very different climate sensitivity values,  $S_3$ – $S_5$ , our model sees the actual difference and produces two distinct control trajectories.

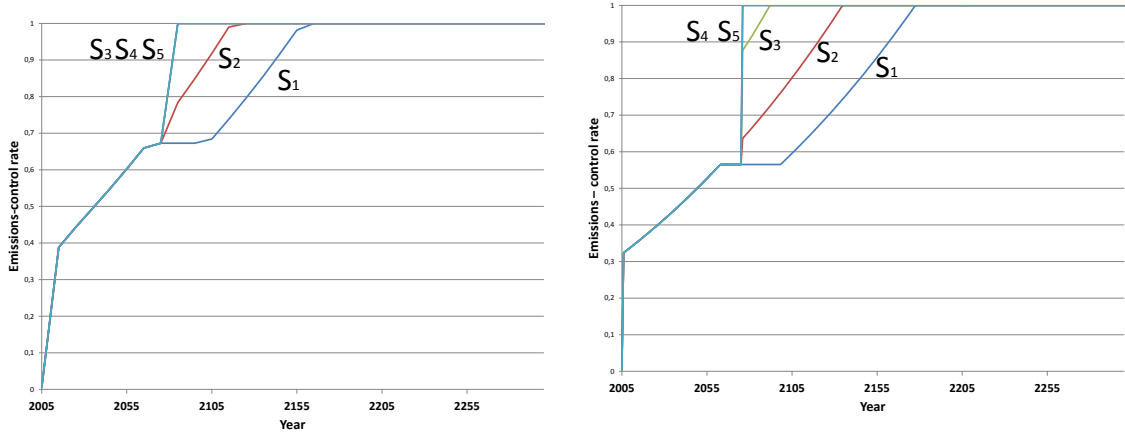


Figure 15: Comparison of the optimal emission reduction policies generated by the two models, the EZ-DICE model by Ackerman et al. (2013) (left) and the IAM in our paper, in which the feedback from the THC to DICE model is ignored (right).

## B Numerical Solution

Having a non-convex optimization problem poses the challenge of finding the global optimum. The utility maximization problem we tackle exhibits two solution options: to trigger the circulation breakdown and accept unavoidable future damages or to initiate stringent control in the near-term and prevent the catastrophe. Each option has the control path that ensures locally maximum utility. The numerical procedure has to provide an answer to the question: which of these two candidates is the global optimum? The answer is given in accordance with exogenously defined preferences of the global decision maker and the benefits of avoiding the damages from the THC collapse. The global solvers available in GAMS are rather restricted with respect to the allowable size of the model. Instead we implement a multi-start heuristic approach to the CONOPT3 solver in the GAMS environment using different initial values and scales. The algorithm in CONOPT3 is based on a generalized reduced gradient (GRG) technique<sup>6</sup> which relies on the gradient information and has a comparative advantage in handling large models such as ours. The three pillars of successfully applying the solver to a non-linear problem are the initial values, scaling and bounds. Furthermore, the solver is designed to operate with the smooth functions only, which is why throughout the model design process, e.g. developing the damage cost function  $f_{THC}$ , we verify this property to ensure convergence.

<sup>6</sup>The GRG is a generalization of the reduced gradient (RG) technique which allows nonlinear constraints. The key idea behind the GRG is to transform the constrained problem to bound constrained and thus reduce the number of independent variables. The further search is performed in the direction of the gradient of the superbasic variables. There are many possible GRG algorithms and CONOPT3 identifies the most appropriate for the particular problem setting. Please refer to Abadie and Carpentier (1969) for the concept of GRG and to Drud (1992) for its implementation in CONOPT.

## C Calibration

Parameter	Value
Relaxation temperature of the box 1, $T_1^*$	$6.6^\circ C$
Relaxation temperature of the box 2, $T_2^*$	$2.7^\circ C$
Relaxation temperature of the box 3, $T_3^*$	$11.7^\circ C$
Volume of the box 1, $V_1$	$1.1 \times 10^{17} m^3$
Volume of the box 2, $V_2$	$0.4 \times 10^{17} m^3$
Volume of the box 3, $V_3$	$0.68 \times 10^{17} m^3$
Volume of the box 4, $V_4$	$0.05 \times 10^{17} m^3$
Thermal coupling constant, box 1	0.059
Thermal coupling constant, box 2	0.059
Thermal coupling constant, box 3	0.1777
Northern Hemisphere temperature conversion constant, $p_{NH}$	1.07
Southern Hemisphere temperature conversion constant, $p_{SH}$	0.93
South Atlantic temperature downscaling constant, $p_1$	0.86
North Atlantic temperature downscaling constant, $p_2$	1.07
Tropics temperature downscaling constant $p_3$	0.79
Hydrological sensitivity, $h_1$	$-0.005 \text{ Sv } ^\circ C^{-1}$
Present day temperature of the box 1, $T_1(1)$	$6.5^\circ C$
Present day temperature of the box 2, $T_2(1)$	$4.7^\circ C$
Present day temperature of the box 3, $T_3(1)$	$11.4^\circ C$
Present day volume transport, $m(1)$	22.6 Sv
Thermal expansion coefficient, $\zeta$	$1.7 \times 10^{-4} ^\circ C^{-1}$
Haline expansion coefficient, $\beta$	$8 \times 10^{-4} psu^{-1}$
Empirical flow constant, $k$	$25.4 \times 10^{17} m^3 a^{-1}$
Reference salinity, $S_0$	35 psu

Table 2: The THC model parameters. (Zickfeld et al. (2004))

For the DICE-CJL model parameters, please refer to Cai et al. (2012a) as well as Cai et al. (2012b) and the accompanying website.

Parameter	Value
Rate of pure time preference, $\rho$	0.0113
Intertemporal elasticity of substitution, IES	1.5 per cent
Learning date	year 2075
North Atlantic hydrological sensitivity, $h_2^{tot}$	0.0525 Sv $^{\circ}C^{-1}$
THC collapse specific damages, $d_m$	0.03, fraction of global GDP
Initial emissions-control rate, $\mu(1)$ (Kyoto agreement)	0.005

Table 3: Parameters crucial to the present study

## References

- Abadie, Carpentier, 1969. Optimization. Academic press, Ch. Generalization of the Wolfe Reduced Gradient Method to the case of Nonlinear Constraints, pp. 37–47.
- Ackerman, F., DeCanio, S., Howarth, R., Sheeran, K., 2009. Limitations of integrated assessment models of climate change. *Climatic Change* 95 (3-4), 297–315.
- Ackerman, F., Stanton, E. A., Bueno, R., 2013. Epstein-Zin utility in DICE: Is risk aversion irrelevant to climate policy? *Environmental and Resource Economics* 56 (1), 73–84.
- Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance* 59 (4), 1481–1509.
- Barro, R. J., 2013. Environmental protection, rare disasters, and discount rates. Working Paper 19258, National Bureau of Economic Research.
- Cai, Y., Judd, K. L., Lontzek, T. S., 2012a. Continuous-time methods for integrated assessment models. Working Paper 18365, National Bureau of Economic Research.
- Cai, Y., Judd, K. L., Lontzek, T. S., 2012b. Open science is necessary. *Nature Climate Change* 2, 299–299.
- Cai, Y., Judd, K. L., Lontzek, T. S., 2013. The social cost of stochastic and irreversible climate change. Working Paper 18704, National Bureau of Economic Research.
- Cass, D., 1965. Optimum growth in an aggregative model of capital accumulation. *The Review of Economic Studies* 32 (3), 233–240.
- Cline, W. R., 1992. The Economics of Global Warming. No. 39 in Peterson Institute Press: All Books. Peterson Institute for International Economics.
- Crost, B., Traeger, C. P., 2010. Risk and aversion in the integrated assessment of climate change. CUDARE Working Paper Series 1104R, University of California at Berkeley, Department of Agricultural and Resource Economics and Policy.

- DeLong, J. B., Magin, K., 2009. The u.s. equity return premium: Past, present, and future. *The Journal of Economic Perspectives* 23 (1), pp. 193–208.
- Drijfhout, S., Weber, S., Swaluw, E., 2011. The stability of the moc as diagnosed from model projections for pre-industrial, present and future climates. *Climate Dynamics* 37 (7-8), 1575–1586.
- Drud, A., 1992. Conopt - a large-scale grg code. *ORSA journal on computing* 6, 207–216.
- Epstein, L. G., Zin, S. E., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica* 57 (4), pp. 937–969.
- Epstein, L. G., Zin, S. E., 1991. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy* 99, 263–286.
- Ganachaud, A., Wunsch, C., 2000. Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data. *Nature* 408, 453–457.
- Ha-Duong, M., Treich, N., 2004. Risk aversion, intergenerational equity and climate change. *Environmental and Resource Economics* 28 (2), 195–207.
- Hawkins, E., Smith, R. S., Allison, L. C., Gregory, J. M., Woollings, T. J., Pohlmann, H., de Cuevas, B., 2011. Bistability of the atlantic overturning circulation in a global climate model and links to ocean freshwater transport. *Geophysical Research Letters* 38 (10), L10605.
- Hope, C., 2006. The marginal impact of CO<sub>2</sub> from PAGE2002: An integrated assessment model incorporating the IPCC’s five reasons for concern. *Integrated Assessment* 6 (1), 19–56.
- IPCC, 2007. Summary for policymakers. In: Solomon, S., D., Qin, M., Manning, Z., Chen, M., Marquis, K., Averyt, M., Tignor, Miller, H. (Eds.), *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Jensen, S., Traeger, C., 2014. Optimal climate change mitigation under long-term growth uncertainty: Stochastic integrated assessment and analytic findings. *European Economic Review* forthcoming, –.
- Kaufman, N., 2012. The bias of integrated assessment models that ignore climate catastrophes. *Climatic Change* 110 (3-4), 575–595.

- Koopmans, T. C., 1963. On the concept of optimal economic growth. Cowles Foundation Discussion Papers 163, Cowles Foundation for Research in Economics, Yale University.
- Kopp, R. E., Hsiang, S. M., Oppenheimer, M., 2013. Empirically calibrating damage functions and considering stochasticity when integrated assessment models are used as decision tools. In: Impacts World 2013 Conference Proceedings.
- Kreps, D. M., Porteus, E. L., 1978. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* 46 (1), pp. 185–200.
- Kriegler, E., Hall, J. W., Held, H., Dawson, R., Schellnhuber, H. J., 2009. Imprecise probability assessment of tipping points in the climate system.
- Kuhlbrodt, T., Rahmstorf, S., Zickfeld, K., Vikeb  , F., Sundby, S., Hofmann, M., Link, P. M., Bondeau, A., Cramer, W., Jaeger, C., 2009. An integrated assessment of changes in the thermohaline circulation. *Climatic Change* 96 (4), 489–537.
- Lemoine, D., Traeger, C., 2014. Watch your step: Optimal policy in a tipping climate. *American Economic Journal: Economic Policy* 6 (1), 137–66.
- Lenton, T., Ciscar, J.-C., 2013. Integrating tipping points into climate impact assessments. *Climatic Change* 117 (3), 585–597.
- Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., Schellnhuber, H. J., 2008. Tipping elements in the Earth’s climate system. *Proceedings of the National Academy of Sciences of the United States of America* 105, 1786–1793.
- Link, P., Tol, R., 2011. Estimation of the economic impact of temperature changes induced by a shutdown of the thermohaline circulation: an application of FUND. *Climatic Change* 104 (2), 287–304.
- Lontzek, T. S., Cai, Y., Judd, K. L., 2012. Tipping points in a dynamic stochastic IAM. RDCEP Working Paper 12-03, RDCEP.
- Manabe, S., Stouffer, R. J., 1994. Multiple-century response of a coupled ocean-atmosphere model to an increase of atmospheric carbon dioxide. *Journal of Climate* 7, 5–23.
- Mastrandrea, M. D., Schneider, S. H., 2001. Integrated assessment of abrupt climatic changes. *Climate Policy* 1 (4), 433–449.
- Matei, D., Baehr, J., Jungclaus, J. H., Haak, H., M  ller, W. A., Marotzke, J., 2012. Multiyear prediction of monthly mean atlantic meridional overturning circulation at 26.5   n. *Science* 335 (6064), 76–79.

- McCarl, B., 2013. McCarl Expanded GAMS User Guide. GAMS Development Corporation.
- Nordhaus, W., 2008. A Question of Balance: Weighing the Options on Global Warming Policies. Yale University Press.
- Nævdal, E., Oppenheimer, M., 2007. The economics of the thermohaline circulation – a problem with multiple thresholds of unknown locations. *Resource and Energy Economics* 29 (4), 262 – 283.
- Pindyck, R. S., 2013. Climate change policy: What do the models tell us? Working Paper 19244, National Bureau of Economic Research.
- Rahmstorf, S., Ganopolski, A., 1999. Long-term global warming scenarios computed with an efficient coupled climate model. *Climatic Change* 43 (2), 353–367.
- Ramsey, F. P., 1928. A mathematical theory of saving. *The Economic Journal* 38 (152), 543–559.
- Roe, G. H., Baker, M. B., 2007. Why is climate sensitivity so unpredictable? *Science* 318 (5850), 629–632.
- Schneider, S. H., Thompson, L., 2000. Simple climate model used in economic studies of global change. In: DeCanio, S. J., Howarth, R. B., Sanstad, A. H., Schneider, S. H., Thompson, S. L. (Eds.), *New Directions in the Economics and Integrated Assessment of Global Climate Change*. Pew Center on Global Climate Change, Ch. 5, pp. 59–80.
- Stern, N., 2013. The structure of economic modeling of the potential impacts of climate change: Grafting gross underestimation of risk onto already narrow science models. *Journal of Economic Literature* 51 (3), 838–59.
- Stommel, H., 1961. Thermohaline convection with two stable regimes of flow. *Tellus* 13 (2), 224–230.
- Tallarini, T. D. J., 2000. Risk-sensitive real business cycles. *Journal of Monetary Economics* 45 (3), 507 – 532.
- Tol, R., 1998. Potential slowdown of the thermohaline circulation and climate policy. Discussion paper ds98/06, Institute for Environmental Studies Vrije Universiteit Amsterdam.
- Tol, R., 2009. An analysis of mitigation as a response to climate change. Discussion paper, Copenhagen Consensus on Climate.
- Vissing-Jørgensen, A., Attanasio, O. P., 2003. Stock-market participation, intertemporal substitution, and risk-aversion. *The American Economic Review* 93 (2), 383–391.

- Zickfeld, K., Bruckner, T., 2003. Reducing the risk of abrupt climate change: Emissions corridors preserving the atlantic thermohaline circulation. *Integrated Assessment* 4 (2), 106–115.
- Zickfeld, K., Bruckner, T., 2008. Reducing the risk of atlantic thermohaline circulation collapse: sensitivity analysis of emissions corridors. *Climatic Change* 91, 291–315.
- Zickfeld, K., Levermann, A., Morgan, M., Kuhlbrodt, T., Rahmstorf, S., Keith, D., 2007. Expert judgements on the response of the atlantic meridional overturning circulation to climate change. *Climatic Change* 82, 235–265.
- Zickfeld, K., Slawig, T., Rahmstorf, S., 2004. A low-order model for the response of the atlantic thermohaline circulation to climate change. *Ocean Dynamics* 54, 8–26.