NCLIM-15010097A

Stochastic integrated assessment of climate tipping points calls for strict climate policy

Thomas S. Lontzek[1]*, Yongyang Cai[2,3], Kenneth L. Judd[2,4], Timothy M. Lenton[5]*

[1]Department of Quantitative Business Administration, University of Zurich, 8008 Zürich, CH

[2]Hoover Institution, Stanford University, Stanford, CA 94305, USA

[3]Becker Friedman Institute, University of Chicago, Chicago, IL 60636, USA

[4]National Bureau of Economic Research, Cambridge, MA 02138, USA

[5]College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QE, UK

*Correspondence to: lontzek@gmail.com, t.m.lenton@exeter.ac.uk

Perhaps the most "dangerous" aspect of future climate change is the possibility that human activities will push parts of the climate system past "tipping points", leading to irreversible impacts[1]. The likelihood of such "large-scale singular events"[2] is expected to increase with global warming[1-3], but is fundamentally uncertain[4]. A key question is how should the uncertainty surrounding tipping events[1,5] affect climate policy? We address this using a stochastic integrated assessment model[6], based on the widely-used deterministic 'DICE' model[7]. The temperature-dependent likelihood of tipping is calibrated using expert opinions[3], which we find to be internally consistent. The irreversible impacts of tipping events are assumed to accumulate steadily over time (rather than instantaneously[8-11]), consistent with scientific understanding[1,5]. Even with conservative assumptions about the rate and impacts of a stochastic tipping event, today's optimal carbon tax is increased by ~50%. For a plausibly rapid, high-impact

tipping event, today's optimal carbon tax is increased by >200%. The additional carbon tax to delay climate tipping grows at only about half the rate of the baseline carbon tax. This implies that the effective discount rate for the costs of stochastic climate tipping is much lower than the discount rate[7,12,13] for deterministic climate damages. Our results support recent suggestions that the costs of carbon emission used to inform policy[12,13] are being underestimated[14-16], and that uncertain future climate damages should be discounted at a low rate[17-20].

Integrated assessment models (IAMs) are key tools to assist climate policy-making[7,12,13], which attempt to capture two-way interactions between climate and society. There is much debate over what discount rate to assume for evaluating future damages due to global temperature rise[17], which in turn partly determines how much we should be willing to pay now to avoid or delay those damages. The Stern Review[21] followed a prescriptive (and controversial[22-24]) approach; based on ethical arguments it assumed a near-zero rate for discounting the utility of future generations, implying a low discount rate for monetized damages of climate change and a high willingness to pay now. In contrast, studies using a descriptive approach[7,12,13] generally evaluate the costs of climate change using much higher market rates of return as discount rates. Most studies are deterministic, but uncertainty will also affect the rate at which future levels of climate damage are discounted[17-20]. Climate tipping points and their impacts are a key source of uncertainty, for several reasons[1,3,4]. Firstly, our knowledge of thresholds, in terms of e.g. regional warming, is imperfect, and the mapping from global temperature rise to regional thresholds is also uncertain. Secondly, even if we knew a tipping point precisely, stochastic internal variability in the climate system could trigger tipping at a range of times and corresponding global temperatures[4]. Several IAM approaches to model climate tipping points are fundamentally deterministic[8,9,14,25,26], whereas only a few studies include stochastic climate damages[10,11,27] (see Supplementary

Discussion). In common with deterministic IAMs, they generally assume[10,11] that the impacts of passing a tipping point are felt instantaneously, whereas in reality impacts will accumulate over time at a rate determined by the dynamics of the system that has been tipped[1]. One recent study[27] assumes that tipping instantaneously increases climate sensitivity or weakens carbon sinks, which then causes damages to accumulate at an increased rate, but this is scientifically questionable (see Supplementary Discussion) and leads to increased discounting of future damages[27].

Here, we examine how a more realistic treatment of stochastic climate tipping points affects the optimal policy choice, including the discount rate to evaluate future damages. Our stochastic integrated assessment model[6], DSICE (Fig. 1a), builds on the deterministic Dynamic Integrated Climate and Economy (DICE) model[7] (2007 version) as used in the 2010 U.S. federal assessment of the social cost of carbon[12]. The federal assessment[13] and the DICE model[28] have since been updated, in ways that tend to increase the estimated social cost of carbon (see Supplementary Methods). Hence the reader should focus on our relative changes in carbon tax due to stochastic climate tipping more than the absolute values.

DSICE uses a dynamic programming framework, representing the decision maker's uncertainty by a stochastic formulation of a tipping event as a Markovian process (see Methods and Supplementary Methods). Specifically, for a potential hazard event the model specifies a hazard rate – i.e. the conditional probability that a tipping point will be passed in a particular year given the temperature that year. The decision maker is assumed to use the hazard rates inferred from an expert elicitation study[3] (see Methods and Supplementary Methods). The average experts' hazard rate has a default value of 0.0025 $°C^{-1}$ yr^{-1} – e.g. if we observe 1°C of warming, the conditional probability of having a tipping event in that year is 0.25%, rising to 0.5% yr^{-1} for 2°C of warming. Following the expert elicitation[3], the tipping event could be one out of five candidates: reorganization of the Atlantic meridional overturning circulation; irreversible melt of the Greenland ice sheet; collapse of the West Antarctic ice sheet; dieback of the Amazon rainforest; or an increase in the amplitude of the El Niño Southern Oscillation. We conservatively assume that (i) whatever the tipping event is, it leads to only modest damages – our default setting is a 10% reduction in global GDP, and (ii) these damages take significant time to unfold (Fig. 1b) – with a default setting of 50 years (appropriate e.g. for Amazon rainforest dieback). Incorporating this stochastic potential tipping event into the DSICE model, the resulting cumulative probability of tipping is ~2.5% in 2050, ~13.5% in 2100, and ~48% (i.e., as likely as not) in 2200 (see Supplementary Results), in good agreement with the expert elicitation results[3].

Despite our conservative default assumptions, the prospect of an uncertain future tipping point causes an immediate increase in the initial (2005) carbon price (Fig. 2, blue line) by ~50%, from \$36.7 tC^{-1} to \$55.6 tC^{-1} (all prices are in 2005 US\$, multiply by 1.16 for 2013 US\$). The relatively low carbon price when the tipping point is ignored, and its high average growth rate of 1.68% yr^{-1} (from ~\$36.7 tC^{-1} in 2005 to \$173 tC^{-1} in 2100: Fig. 2, black line), is the response to the steadily increasing, deterministic effect of rising temperature on economic output. It reflects the DICE preferences of discounting future welfare at a high rate. In contrast, the expected additional carbon tax to address the tipping point threat (difference between black and blue lines in Fig. 2) grows at roughly half the average rate (0.81% yr^{-1}) of the baseline DICE carbon tax (Fig. 3). Such a flat carbon tax path is also obtained when the discount rate is prescribed to be lower (as in e.g. the Stern Review[20]). Thus, despite assuming the same dynamic preferences of discounting welfare of future generations as Nordhaus[7], our model indicates that the appropriate discount rate for climate tipping damages is a low one.

This can be understood by considering the expected returns on mitigation investment. Tipping points add a source of risk to the economic system, which increases the variance of future output. Hence mitigation expenditures have two effects on economic output. First, they increase expected output (by reducing expected damages). Second, they reduce the variance of output, further increasing social welfare. This means decisions on capital investment and mitigation expenditures will face different criteria: Increasing the capital stock in the DICE model will increase future expected output, and the marginal benefit from investment today is discounted at the market interest rate. Increasing mitigation expenditures will increase future expected output (again discounted at the market interest rate), but will also reduce the variance of future output. Therefore, mitigation expenditures to address stochastic damages will exceed the level justified by the discounted impact on expected output[17,19,20]. This implies a discount rate that is less than the interest rate. It explains why the increase in the carbon tax from tipping events exceeds that from the change in future expected output.

The optimum way of dealing with the threat of a tipping point event also resembles characteristics of an insurance policy. Insurance purchases have a negative rate of return since insurance premiums are much higher than the expected loss. The expected additional carbon price thus balances discounting of the future with the desire for insurance, resulting in its slower growth rate. It can be thought of as a premium that is levied upon society with the purpose of delaying potential damage from the tipping event.

Previous deterministic IAM studies[14,25,26,29] have suggested that increasing the convexity of the damage function in the DICE model could represent the characteristics of a tipping point. As a comparison exercise we studied the implications for climate policy of doubling or tripling the exponent of the damage function. Unsurprisingly, these deterministic approaches enhance the growth rate of the carbon price (implying a higher discount rate) (Fig. 3, red lines), whereas our stochastic treatment decreases it (Fig. 3, blue line). Hence, existing studies[16,26-29] that adjust the shape of a deterministic damage function qualitatively fail to capture the implications of stochastic tipping points.

Candidate tipping points differ in their intrinsic timescales and impacts[1,5]. Hence, in a sensitivity study (Fig. 4), we considered tipping processes that take 5, 50 (default), 100, and 500 years to fully unfold, with final stage impacts of 2.5%, 5%, 10% (default), and 20% damage to output. We also looked at how a higher hazard rate affects the optimal climate policy. This gives a total of 32 combinations, each of which can be thought of as hypothetically representing the characteristics of some tipping event. The additional carbon price significantly decreases with increasing transition time (Fig. 4a) suggesting that previous studies[10,11] (see Supplementary Discussion), assuming an instantaneous full impact of climate tipping, bias the carbon price upward. The additional carbon price also increases with increasing damage and likelihood of the tipping point event (Fig. 4a). As the final stage damage doubles, the additional carbon price also roughly doubles. Furthermore, a higher hazard rate amplifies the effect of shorter transition scales on the additional carbon price.

The additional carbon tax delays the expected occurrence of the climate tipping point (Fig. 4b), in our default scenario by 20 years (from year 2214 to 2234). This expected delay time increases with increased damage, shorter transition periods, and with higher likelihood of tipping, to more than a century in our extreme cases (Fig. 4b). The expected additional carbon tax (in $/tC) correlates with the length of the expected delay (in years), such that each dollar added to the carbon tax correlates with a delay of the tipping event by a year.

The growth rate of the additional carbon price is relatively insensitive to varying damage level or transition time, ranging over 0.43%-0.96% yr^{-1} in our sensitivity analysis (Fig. 4c). This is 40-70% less than the growth rate of the baseline carbon price (1.68% yr^{-1}) in the deterministic model without tipping.

Actual candidate tipping elements in the climate system[1] can be tentatively related to modeled combinations of hazard rate, tipping duration, and final damages (Fig. 4d), based in part on previous reviews of the literature[1,5]. This is necessarily somewhat subjective.

Nevertheless, it serves to qualitatively illustrate that the optimal policy response for different specific climate tipping points could differ profoundly.

In conclusion, the optimal policy in response to the threat of a stochastic, irreversible tipping point differs substantially from the policy response to the deterministic effect of temperature on output. The damages associated with the stochastic possibility of a future climate tipping point should be discounted at a low rate[17]. This calls for a higher carbon price and increased efforts to mitigate emissions now – without even considering other co-benefits of mitigation[30], such as decreased air pollution and greater energy security. Thus, when appropriately treating the intrinsic uncertainty in the climate system – in this case the stochastic nature of future climate tipping points – a strict climate policy can emerge from a pure market-based approach. It does not have to be based on moral judgments about sustainability and the wellbeing of future generations[21] – although these are, of course, legitimate and important concerns.

Methods Summary

We use DSICE[6], a multidimensional stochastic integrated assessment model (IAM) of climate and the economy, based on the DICE model[7]. DICE has been applied in numerous studies, e.g.[9,14,26], and the main drivers of its behavior have been analyzed[7]. DSICE computes the optimal, global greenhouse gas emission reduction. Higher emission control at present mitigates the damage from climate change in the future but limits consumption and/or capital investment today. The global economy (the social planner) is set to weigh these costs and benefits of emission control to maximize the expected present value of global social welfare. DSICE includes the possibility of a climate tipping point with potential damages to economic output. The occurrence of a climate tipping point is modeled by a Markov process (with a hazard rate) and its timing is not known at times of decisions. Because DSICE is a stochastic model, it can compute the optimal policy response, i.e.: a tax on carbon emissions to address the uncertain climate tipping event. See Supplementary Methods for a full model description.

The hazard rate for a tipping event represents the conditional probability that a tipping point will occur in a particular year given the actual degree of global warming in that year (above year 2000). Previous work[3] from a range of experts has elicited imprecise cumulative probabilities for passing five different tipping points under three different temperature corridors up to the year 2200. Each temperature corridor spans an uncertainty range, and together they range over 0-8 °C warming (above year 2000) depending on the year and the scenario. Here, we calibrate the hazard rate for the tipping event by reverse engineering the contemporaneous conditional probability of tipping from the cumulative probabilities from the expert elicitation study[3]. See Supplementary Methods for full details of the hazard rate calibration.

References

1 Lenton, T.M. *et al.*, Tipping Elements in the Earth's Climate System. *Proceedings of the National Academy of Sciences* 105 (6), 1786-1793 (2008).

2 IPCC, Climate Change 2014: Impacts, Adaptation and Vulnerability. Summary for Policymakers. 2014.

3 Kriegler, E., Hall, J.W., Held, H., Dawson, R., & Schellnhuber, H.J., Imprecise probability assessment of tipping points in the climate system. *PNAS* 106 (13), 5041-5046 (2009).

4 Lenton, T.M., Early warning of climate tipping points. *Nature Climate Change* 1, 201-209 (2011).

5 Lenton, T.M. & Ciscar, J.-C., Integrating tipping points into climate impact assessments. *Climatic Change* 117 (3), 585-597 (2013).

6 Cai, Y., Judd, K.L., & Lontzek, T.S., The Social Cost of Stochastic and Irreversible Climate Change. *National Bureau of Economic Research Working Paper Series* No. 18704 (2013).

7 Nordhaus, W.D., *A Question of Balance: Weighing the Options on Global Warming Policies*. (Yale University Press, 2008).

8 Mastrandrea, M.D. & Schneider, S.H., Integrated assessment of abrupt climatic changes. *Climate Policy* 1 (4), 433-449 (2001).

9 Kosugi, T., Integrated Assessment for Setting Greenhouse Gas Emission Targets under the Condition of Great Uncertainty about the Probability and Impact of Abrupt Climate Change. *Journal of Environmental Informatics* 14 (2), 89-99 (2009).

10 Gjerde, J., Grepperud, S., & Kverndokk, S., Optimal climate policy under the possibility of a catastrophe. *Resource and Energy Economics* 21 (3-4), 289-317 (1999).

11 Castelnuovo, E., Moretto, M., & Vergalli, S., Global Warming, Uncertainty and Endogenous Technical Change. *Environmental Modeling & Assessment* 8 (4), 291-301 (2003).

12 Interagency Working Group on Social Cost of Carbon, Social Cost of Carbon for Regulatory Impact Analysis - Under Executive Order 12866. United States Government, 2010.

13 Interagency Working Group on Social Cost of Carbon, Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis. United States Government, 2013.

14 Ackerman, F. & Stanton, E.A., Climate Risks and Carbon Prices: Revising the Social Cost of Carbon *Economics: The Open-Access, Open Assessment E-Journal* 6 (2012-10), 1-25 (2012).

15 van den Bergh, J.C.J.M. & Botzen, W.J.W., A lower bound to the social cost of CO2 emissions. *Nature Clim. Change* 4 (4), 253-258 (2014).

16 Revesz, R.L. *et al.*, Improve economic models of climate change. *Nature* 508, 173-175 (2014).

17 Arrow, K. *et al.*, Determining Benefits and Costs for Future Generations. *Science* 341 (6144), 349-350 (2013).

18 Barro, R.J., Rare Disasters, Asset Prices, and Welfare Costs. *American Economic Review* 99 (1), 243-264 (2009).

19 Gollier, C., *Pricing the Planet's Future: The Economics of Discounting in an Uncertain World*. (Princeton University Press, Princeton, NJ, 2012).

20 Pindyck, R.S. & Wang, N., The Economic and Policy Consequences of Catastrophes. *American Economic Journal: Economic Policy* 5 (4), 306-339 (2013).

21 Stern, N., *The Economics of Climate Change: The Stern Review*. (Cambridge University Press, Cambrige, 2006).

22 Dasgupta, P., Commentary: The Stern Review's Economics of Climate Change. *National Institute Economic Review* 199, 4-7 (2007).

23 Nordhaus, W.D., A Review of the *Stern Review on the Economics of Climate Change*. *Journal of Economic Literature* 45 (3), 686-702 (2007).

24 Weitzman, M.L., A Review of the Stern Review on the Economics of Climate Change. *Journal of Economic Literature* 45 (3), 703-724 (2007).

25 Weitzman, M.L., GHG Targets as Insurance Against Catastrophic Climate Damages. *Journal of Public Economic Theory* 14 (2), 221-244 (2012).

26 Ackerman, F., Stanton, E.A., & Bueno, R., Fat tails, exponents, extreme uncertainty: Simulating catastrophe in DICE. *Ecological Economics* 69 (8), 1657-1665 (2010).

27 Lemoine, D. & Traeger, C., Watch Your Step: Optimal Policy in a Tipping Climate. *American Economic Journal: Economic Policy* 6 (1), 137-166 (2014).

28 Nordhaus, W., *The Climate Casino: Risk, Uncertainty, and Economics for a Warming World*. (Yale University Press, 2013).

29 Nordhaus, W.D., An Analysis of the Dismal Theorem. Cowles Foundation for Research in Economics, Yale University, 2009.

30 McCollum, D. *et al.*, Climate policies can help resolve energy security and air pollution challenges. *Climatic Change* 119 (2), 479-494 (2013).

Acknowledgments. We thank Kenneth Arrow, Buz Brock, Larry Goulder and participants of the 2014 Workshop on the Economics of Complex Systems at the Beijer Institute of Ecological Economics for comments. Y.C., K.L.J, and T.S.L. were supported by NSF (SES-0951576). T.S.L was also supported by the Züricher Universitätsverein, the University of Zurich, and the Ecosciencia Foundation. T.M.L. was supported by a Royal Society Wolfson Research Merit Award and the European Commission (ENV.2013.6.1-3). Part of this study was done while T.S.L. was visiting the Hoover Institution. Supercomputer support was provided by Blue Waters (NSF awards OCI-0725070 and ACI-1238993, and the state of Illinois), and by NIH (1S10OD018495-01).

Author contributions. Y.C., K.L.J and T.S.L developed the model with input from T.M.L.. Y.C. and K.L.J. developed the computational method and Y.C. developed the code. All authors analyzed the results. T.S.L. and T.M.L. took the lead in writing the paper with inputs from Y.C and K.L.J.

Correspondence and requests for materials should be addressed to lontzek@gmail.com or t.m.lenton@exeter.ac.uk

Figure captions

Fig. 1. **Schematic of the DSICE model. a.** The forward-looking decision maker (social planner) chooses mitigation and consumption to maximize the sum of discounted expected utilities over some time horizon. Increased mitigation must be traded off against consumption and savings. Global warming adversely impacts the economy and increases the probability of a tipping point with additional irreversible economic impacts. **b**. The length of the pre-tipping phase is stochastic, and its likelihood depends on global warming. Once tipping is triggered, damages increase linearly over a specified transition time (5-500 years here) to a specified final level (2.5-20% of World GDP here).

Fig. 2. Optimal carbon tax path. Gray-shaded area: Range of stochastic carbon tax paths from 10,000 simulations of the optimal model's solution. Black line: Expected carbon tax from the stochastic model (average of 10,000 simulations). Blue line: Optimal carbon tax from a deterministic version of the model in which the decision maker ignores the tipping point (consistent with the DICE model path).

Fig. 3. Growth rates of carbon tax. Black line: Baseline carbon tax in the deterministic DICE model (no tipping point). Blue line: Expected additional carbon tax when including a stochastic tipping point. Solid red line: Additional carbon tax when doubling the exponent of the damage function in the DICE model. Dashed red line: Additional carbon tax when further increasing the deterministic damage function to 6^{th}-order. Solid green line: Additional tax in a deterministic setting where the potential damage from tipping is represented by a

deterministic damage function with an additional component which is the expected damage path of the stochastic model.

Fig. 4. Sensitivity analysis. Sensitivity of DSICE model results to varying the likelihood (hazard rate), transition time, and final impact of the tipping event: **a.** Expected additional carbon tax ($\$ tC^{-1}$) in year 2005. **b.** Expected delay (yr) of the tipping event. **c.** Average (2005-2100) annual growth rate ($\% yr^{-1}$) of the expected additional carbon tax. **d.** Illustrative categorization of elements that could be tipped: Arctic summer sea-ice (ASI), Greenland ice sheet (GIS), West Antarctic ice sheet (WAIS), Atlantic meridional overturning circulation (AMOC), El Niño Southern Oscillation (ENSO), Indian summer monsoon (ISM), West African monsoon (WAM), Amazon rainforest (AMAZ), boreal forests (BOFO).

Fig 1a

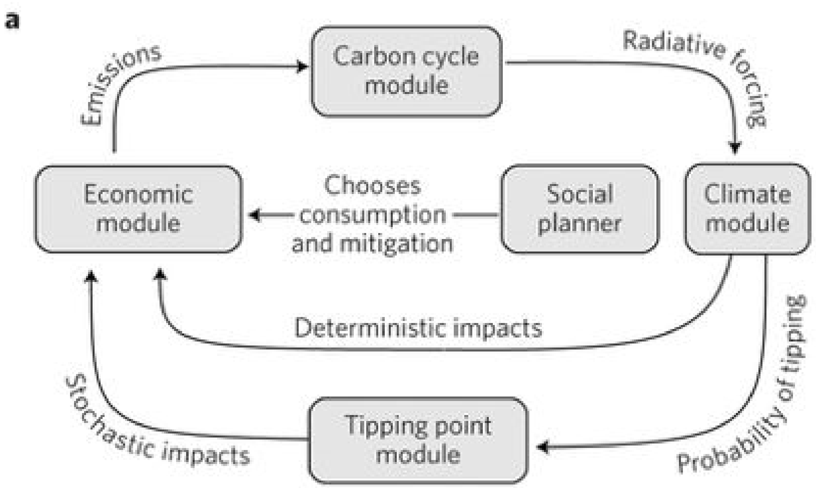

a
Emissions
Carbon cycle module
Radiative forcing
Economic module
Chooses consumption and mitigation
Social planner
Climate module
Deterministic impacts
Probability of tipping
Tipping point module
Stochastic impacts

Fig 1b

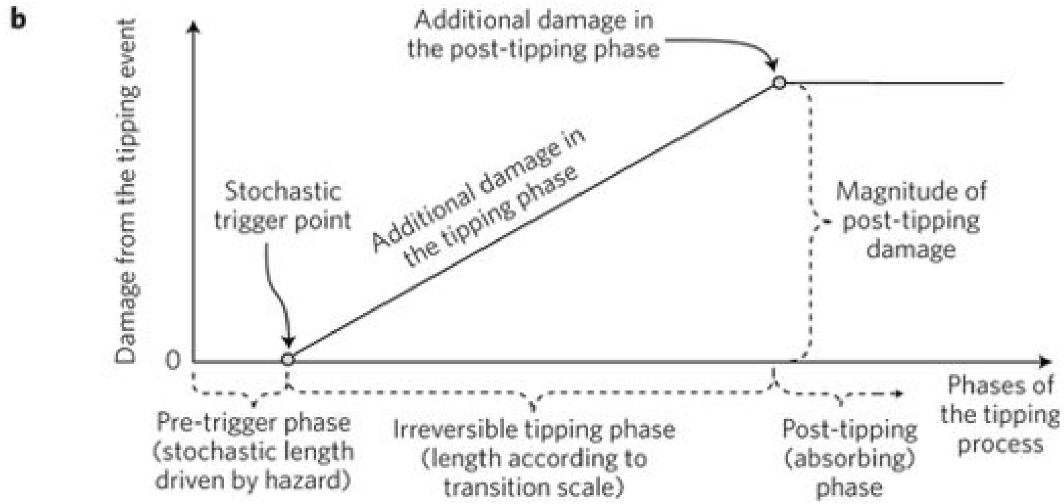
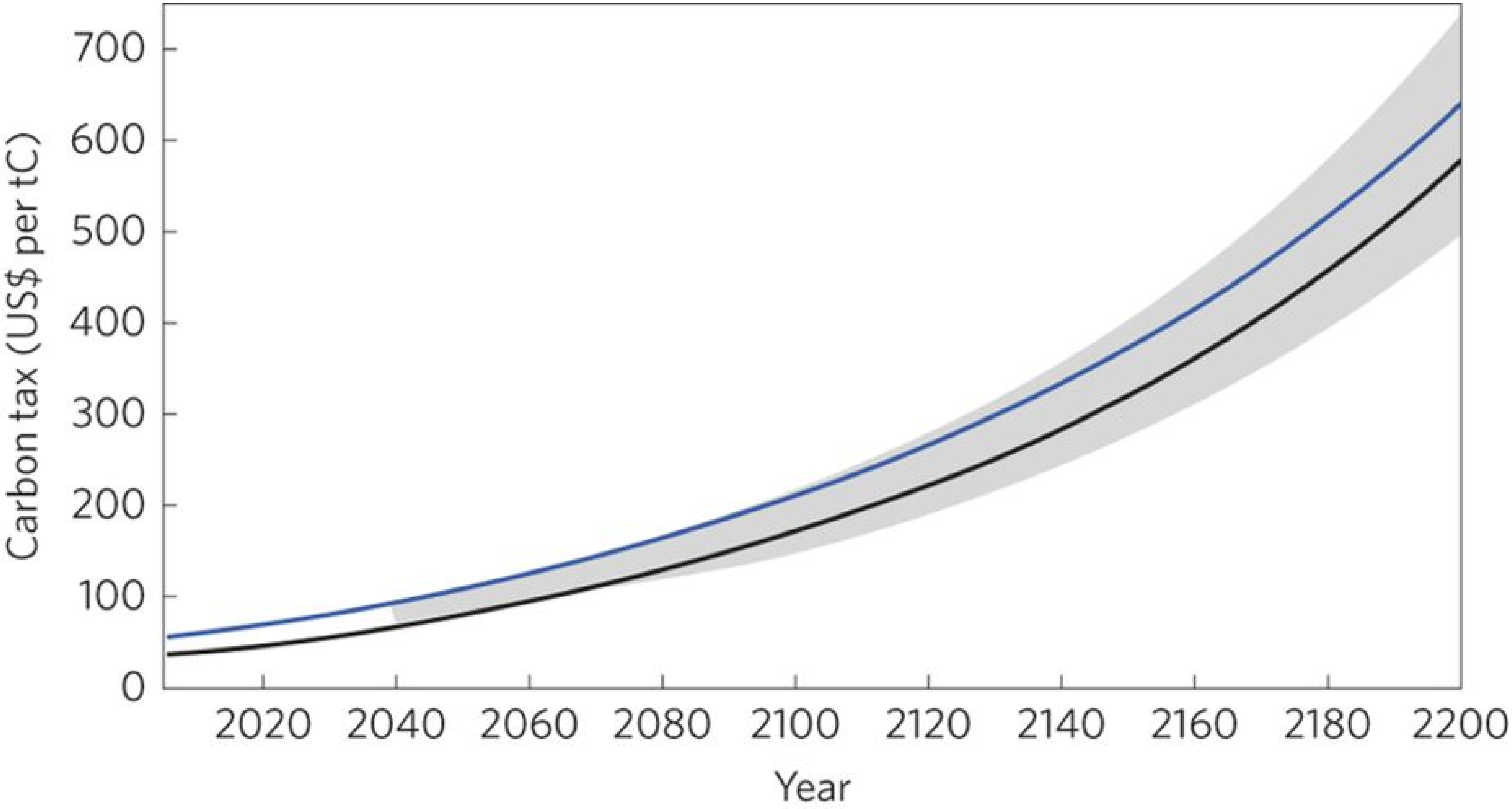

b
Damage from the tipping event
Additional damage in the post-tipping phase
Stochastic trigger point
Additional damage in the tipping phase
Magnitude of post-tipping damage
0
Phases of the tipping process
Pre-trigger phase (stochastic length driven by hazard)
Irreversible tipping phase (length according to transition scale)
Post-tipping (absorbing) phase

Fig 2



700
600
500
400
300
200
100
0
Carbon tax (US$ per tC)
2020
2040
2060
2080
2100
2120
2140
2160
2180
2200
Year

Fig 3

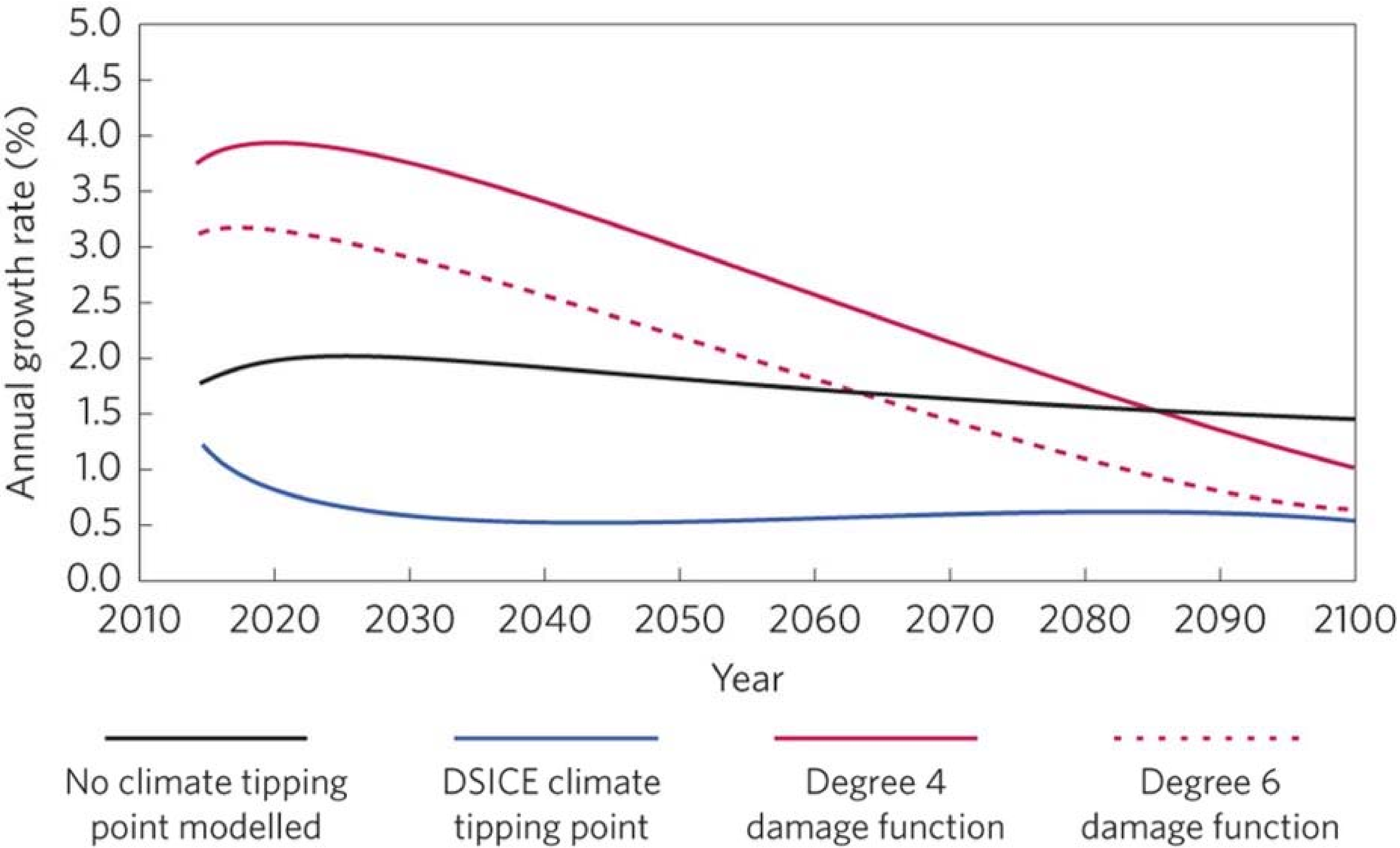
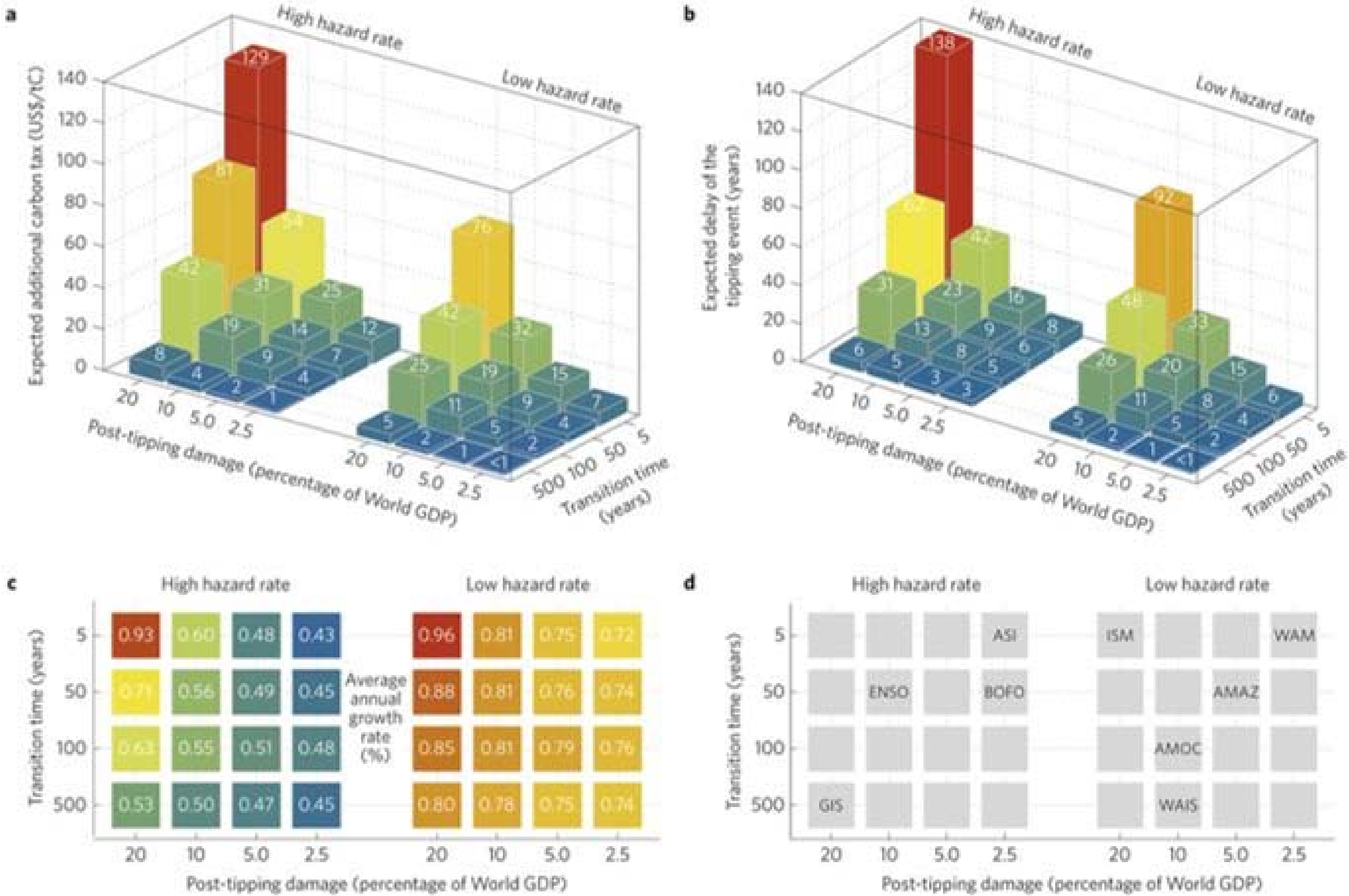

5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Annual growth rate (%)
2010
2020
2030
2040
2050
2060
2070
2080
2090
2100
Year
No climate tipping point modelled
DSICE climate tipping point
Degree 4 damage function
Degree 6 damage function

Fig 4


a
High hazard rate
Low hazard rate
Expected additional carbon tax (US$/tC)
Post-tipping damage (percentage of World GDP)
Transition time (years)
b
Expected delay of the tipping event (years)
c
Average annual growth rate (%)
d
ASI
ISM
WAM
ENSO
BOFO
AMAZ
AMOC
GIS
WAIS

Supplementary Information

Stochastic integrated assessment of climate tipping points calls for strict climate policy

Thomas S. Lontzek[1]*, Yongyang Cai[2,3], Kenneth L. Judd[2,4], Timothy M. Lenton[5]*

[1]Department of Quantitative Business Administration, University of Zurich, 8008 Zürich, CH

[2]Hoover Institution, Stanford University, Stanford, CA 94305, USA

[3]Becker Friedman Institute, University of Chicago, Chicago, IL 60636, USA

[4]National Bureau of Economic Research, Cambridge, MA 02138, USA

[5]College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QE, UK

*Authors for correspondence. E-mail: lontzek@gmail.com or t.m.lenton@exeter.ac.uk

Contents

Supplementary Methods

Description of DSICE: The stochastic integrated assessment model (IAM) used in the present analysis is an adapted version of the DSICE framework[1,2]. The latter builds on the DICE-CJL model[3], which itself is a numerically stable version of DICE-2007[4] with a flexible time-period length. The DICE model has been applied in numerous studies and the main drivers of its behavior have been studied extensively. Besides those associated with risk and uncertainty, the model parameters used for our analysis are calibrated to the same levels as those used in DICE-2007. Hence the deterministic form of DSICE (without stochastic tipping) is almost identical to DICE-2007. The few differences are due to the adaptation of the DICE code to dynamic programming to facilitate solving a stochastic version. These few differences include a flexible time step (DICE has 10 years, DSICE has 1 year), a terminal value function for the optimization procedure (DICE has none) and a fix that future levels of atmospheric carbon do not induce global warming today (DICE-2013[5] has that fix as well). Generally, a multidimensional stochastic IAM (e.g., the one in the present study) requires heavy computation, but DSICE makes it tractable by using an efficient and accurate dynamic programming algorithm[6]. Here, we present the details of the equations of the model used in this study, with the parameter values given in Table S1.

Like DICE, DSICE computes the time paths of the optimal greenhouse gas emission reduction for the world. Stringent emission control at present mitigates the damage of climate change in the future while reducing income today; the global economy (the social planner) is set to weigh these costs and benefits of emission control. Uncertainty (stochasticity) of climate change effects is included in such a way that the global economy makes emission control decisions by expecting future developments of climate and the economy that are not precisely known at the times of decisions. The model finds the levels of global emission control that maximize the expected present value of global social welfare.

The impact of global warming on the economy is reflected by a convex damage function of temperature in the atmosphere. This is a standard feature of the DICE model family. We modify the standard damage function by explicitly modeling the possibility of a climate shock (a tipping point) to account for the threat of abrupt and irreversible climate change. One important element of the tipping point module is the transition scale of the tipping process. For example, Lenton et al.[7] characterize the transition scales of tipping for various climate elements. While some elements are believed to exhibit a rapid tipping transition of about ten years (e.g., Arctic summer sea ice), other elements are believed to exhibit a slower transition of more than 300 years (e.g., the Greenland ice sheet). As we have argued, we do not concentrate on a specific tipping element but we find the one-out-of-five example in Kriegler et al.[8] appealing. Therefore, our tipping point can represent any of the five tipping elements in Kriegler et al.[8]. In this study, we investigate transition scales of 5, 50, 100, and 500 years to reach a new equilibrium of the tipping element.

Mathematically, the tipping point module is described as follows: after the tipping event is triggered, a persistent climate impact state, J_t, will increase continuously from a minimal level (i.e., $J_t = 0$) to some maximum level ($\bar{J} > 0$), implying that $J_{t+1} = \min\{J_t + \Delta_t, \bar{J}\}$, where Δ_t is the incremental impact level from stage t to $t+1$. In our computations we make the continuous tipping process discrete with the time steps. We use I_t as the indicator function to denote the pre-tipping state of the world as $I_t = 0$ and the post-tipping state of the world as $I_t = 1$, where I_t is a jump process with the Markovian hazard rate. The latter is endogenous with respect to the contemporaneous level of global average atmospheric temperature T_t^{AT}. The transition function for I_t from stage t to stage $t+1$ is $I_{t+1} = g^I(I_t, T_t^{AT}, \omega_t^I)$, where ω_t^I is a random process. Therefore, $J_{t+1} = \min\{J_t + \Delta_t, \bar{J}\} I_t$ and the impact factor on the economy becomes

$$\Omega(T^{AT}, J, I) = \frac{1 - IJ}{1 + \pi_2 (T^{AT})^2}$$

where T^{AT} is the average global atmospheric temperature and π_2 is a coefficient in the damage function.. We specify the probability transition matrix of the tipping process at time t as

$$\begin{bmatrix} 1 - p_t & p_t \\ 0 & 1 \end{bmatrix},$$

where its (i, j) element is the transition probability from state i to j for I_t, and $p_t = 1 - \exp\left(-b_1 \max\{0, T_t^{AT} - 1\}\right)$, where b_1 is called the hazard rate factor.

We assume that the damage function affects final output and consequently the accumulation of capital k_t, which transits to the next period, is formulated in a standard fashion:

$$k_{t+1} = (1 - \delta)k_t + Y_t(k_t, T_t^{AT}, \mu_t, J_t, I_t) - c_t,$$

where μ_t is the emission control rate, δ is the annual depreciation rate, c_t is consumption, and Y_t denotes the stochastic production function. It follows that

$$Y_t(k_t, T_t^{AT}, \mu_t, J_t, I_t) = \left(1 - \theta_{1,t}\mu_t^{\theta_2}\right) A_t k_t^{\alpha} l_t^{1-\alpha} \Omega(T_t^{AT}, J_t, I_t),$$

where l_t is exogenous labor supply and A_t is the exogenous productivity level. Furthermore, $\theta_{1,t}\mu_t^{\theta_2}$ accounts for the costs of mitigation as a fraction of output. Given the production function, annual total carbon emissions are given by

$$\varepsilon_t(k_t,\mu_t) = \sigma_t(1-\mu_t)A_t k_t^{\alpha} l_t^{1-\alpha} + E_t^{Land},$$

where σ_t denotes an exogenous carbon intensity of output, and E_t^{Land} is an exogenous rate of emissions from biological processes. Furthermore, we employ the standard separable utility function in the DICE-2007class of models[4], which is

$$u(c_t,l_t) = \frac{(c_t/l_t)^{1-\psi}}{1-\psi}\, l_t\ ,$$

Here, ψ represents the risk aversion parameter.

The structure of the carbon cycle in this study is adapted from the DICE-2007 model[4]. The carbon cycle components are modeled by a three-box module with

$$\boldsymbol{M}_t = (M_t^{AT}, M_t^{UP}, M_t^{LO})^{\top},$$

representing carbon concentrations in the atmosphere (M_t^{AT}), the upper oceans (M_t^{UP}) and the lower oceans (M_t^{LO}). The transition system of the carbon concentration from year t to year $t+1$ is

$$\boldsymbol{M}_{t+1} = \Phi^M M_t + (\varepsilon_t(k_t,\mu_t), 0, 0)^{\top},$$

with the carbon cycle transition matrix given by

$$\boldsymbol{\Phi}^M = \begin{bmatrix} 1-\phi_{12} & \phi_{12} & 0 \\ \phi_{12} & 1-\phi_{21}-\phi_{23} & \phi_{32} \\ 0 & \phi_{23} & 1-\phi_{32} \end{bmatrix},$$

where the coefficient ϕ_{ij} is the rate at which carbon diffuses from carbon stock i to carbon stock j, for $i,j \in \{M^{AT}, M^{UP}, M^{LO}\}$. The carbon concentrations in the atmosphere affect the global average surface temperature via radiative forcing:

$$F_t(M^{AT}) = \eta \log_2\left(M^{AT}/M_{PI}^{AT}\right) + F_t^{EX},$$

where η is an exogenous forcing parameter, M_{PI}^{AT} is the preindustrial carbon concentration in the atmosphere and F_t^{EX} denotes exogenous radiative forcing. Here, we also make use of the DICE-2007 two-box model for the climate[4]. The global mean temperature is represented by a two-layer model,

$$\mathbf{T}_t = (T_t^{AT}, T_t^{LO})^{\top}$$

representing the average temperature in the atmosphere (T_t^{AT}) and the lower oceans (T_t^{LO}). The transition system of the global average temperature from year *t* to year $t+1$ is

$$\mathbf{T}_{t+1} = \boldsymbol{\Phi}^{\top}\mathbf{T}_t + (\xi_1 F_t(M_t^{AT}), 0)^{\top},$$

with

$$\Phi^T = \begin{bmatrix} 1-\varphi_{21}-\xi_2 & \varphi_{21} \\ \varphi_{12} & 1-\varphi_{12} \end{bmatrix},$$

where ξ_1 is a conversion parameter, the coefficient φ_{ij} is the heat diffusion rate from temperature stock i to temperature stock j, for $i,j \in \{T^{AT}, T^{OC}\}$ and ξ_2 is the rate of atmospheric temperature change by infrared radiation to space.

The social planner's goal is to maximize the expected sum of present-discounted welfare over a specific time horizon, taking into account (i) that the release of carbon into the atmosphere has a deterministic and reversible adverse impact on future economic productivity, and (ii) that enhanced global warming induces a higher probability of a stochastic and irreversibly cascading tipping point event leading to permanent damages. The stochastic optimization problem of the social planner becomes

$$\max_{c_t,\mu_t} \mathbb{E}\left\{\sum_{t=0}^{\infty} e^{-\rho t} u(c_t, l_t)\right\}$$

$$k_{t+1} = (1-\delta)k_t + Y_t(k_t, T_t^{AT}, \mu_t, J_t, I_t) - c_t$$

$$\mathbf{M}_{t+1} = \mathbf{\Phi}^M \mathbf{M}_t + (\varepsilon_t(k_t, \mu_t), 0, 0)^\top$$

$$\mathbf{T}_{t+1} = \mathbf{\Phi}^T \mathbf{T}_t + (\xi_1 F_t(M_t^{AT}), 0)^\top$$

$$J_{t+1} = g^J(J_t, I_t)$$

$$I_{t+1} = g^I(I_t, T_t^{AT}, \omega^I)$$

where $\mathbb{E}$ represents the expectation operator and ρ signifies the utility discount rate. There are seven continuous state variables: the capital stock k, the three-dimensional carbon system $\mathbf{M}$, the two-dimensional temperature vector $\mathbf{T}$, and J, which represents the damage factor of the tipping point event. Furthermore, I is the discrete shock to the climate.

The Bellman equation cannot be solved analytically and must be solved numerically. We approximate the infinite horizon model by a 600-year-horizon problem with a given terminal value function approximating the summation of discounted expected utilities from period 600 onwards. The main idea behind the solution method is that we recursively compute the optimal amount of consumption and mitigation in each period as a function of the state space, including the capital stock, the carbon cycle, the climate and the productivity state. Thus, we start at the last decision period (period: 599) and use value function iteration and approximation methods to obtain the value function (maximand) for any combination of possible (bounded) state space in that period. After solving the problem in the second-last period, we move one period backwards in time to the second-last-decision period (period: 598) and proceed with the same approach until we reach the initial time (period: 1). For a more detailed discussion this procedure see e.g., Cai and Judd[6].

In accordance with the DICE model[4] in which a time period lasts 10 years, we use the following functional forms and parameter values for our model in which we assume that each period lasts one year. The exogenous population path is given by

$$L_t = 6514e^{-0.035t} + 8600(1 - e^{-0.035t})$$

The deterministic productivity level A_t equals

$$A_t = A_0 \exp\left(\alpha_1(1 - e^{-\alpha_2 t})/\alpha_2\right),$$

where α_1 is the initial growth rate and α_2 is the decline rate of the growth rate. Additional equations for σ_t (carbon intensity of output.), $\theta_{1,t}$ (mitigation cost coefficient), $E_{Land,t}$ (annual carbon emissions from biological processes), and $F_{EX,t}$ (exogenous radiative forcing) are given by

$$\sigma_t = \sigma_0 \exp\left(-0.0073(1 - e^{-0.003t})/0.003\right)$$

$$\theta_{1,t} = 1.17\sigma_t(1 + e^{-0.005t})/2\theta_2)$$

$$E_{Land,t} = 1.1e^{-0.01t}$$

$$F_{EX,t} = \begin{Bmatrix} -0.06 + 0.0036t, & if\ t \leq 100 \\ 0.3 & otherwise \end{Bmatrix}$$

The following Table S1 specifies all variables and parameter values of the model.

$t \in \{0,1,\dots,600\}$	time in years (t represents year t + 2005)
$\psi = 2$	risk aversion parameter
$\rho = 0.015$	discount factor
A_t	productivity trend at time t , $A_0 = 0.0272$
K_t	capital at time t (in $ trillions), $K_0 = 137$
$\alpha = 0.3$	output elasticity of capital
$\alpha_1 = 0.0092$	initial growth rate of the productivity trend
$\alpha_2 = 0.001$	decline rate of the growth rate of the productivity trend
$\delta = 0.1$	annual depreciation rate
M_t^{AT}	carbon concentration in atmosphere (billion tons) at time t. $M_0^{AT} = 808.9$
M_t^{UO}	carbon concentration in upper ocean (billion tons) at time t. $M_0^{UO} = 1255$
M_t^{LO}	carbon concentration in lower ocean (billion tons) at time t. $M_0^{LO} = 18365$
T_t^{AT}	average surface temperature change from 1900 ($\circ\ C$) at time t. $T_0^{AT} = 0.7307$
T_t^{OC}	average ocean temperature change from 1900 ($\circ\ C$) at time t. $T_0^{OC} = 0.0068$
$\pi_2 = 0.0028388$	damage factor parameter
$\theta_2 = 2.8$	damage factor parameter
$\sigma_0 = 0.13418$	initial technology factor
$\varphi_{12} = 0.019$	rate of carbon flux: atmosphere to upper ocean
$\varphi_{23} = 0.0054$	rate of carbon flux: upper ocean to lower ocean
$\varphi_{21} = 0.01$	rate of carbon flux: upper ocean to atmosphere ocean
$\varphi_{32} = 0.00034$	rate of carbon flux: lower ocean to upper ocean
$\xi_1 = 0.037$	temperature transition parameter
$\xi_2 = 0.047$	rate of atmospheric temperature decrease due to infrared radiation to space
$\phi_{12} = 0.01$	rate of heat diffusion: atmosphere to ocean
$\phi_{21} = 0.0048$	rate of heat diffusion: ocean to atmosphere
$\eta = 3.8$	radiative forcing parameter
$M_{PI}^{AT} = 596.4$	preindustrial atmospheric carbon concentration
$\underline{T^{AT}} = 1$	surface temperature with zero probability of tipping
$b_1 \in [0.0025,0.0045]$	hazard rate factor

Table S1: Parameters of the DSICE Model.

Differences in DICE-2013: Since calibrating the DSICE model against DICE-2007, the underlying model has been updated[5] to DICE-2013, with several changes: Exogenous processes have been updated, including a higher growth rate of total factor productivity, a lower rate of decarbonisation of the economy and newer estimates of population and exogenous radiative forcing. Ocean absorption of carbon is now lower for higher degrees of global warming. The deterministic damage function has been increased by about 25% to account for additional damages. Perhaps most importantly, the rate of relative risk aversion has been reduced (from 2 to 1.45), which is justified[5] by the desire to match empirical data on market interest rates. This is a key driver of an increase in the carbon tax in DICE-2013 relative to DICE-2007, which is also increased by a much higher projection of world population.

Calibration of the Hazard Rate: We calibrate the hazard rates for climate tipping events using the results of the expert elicitation study by Kriegler et al.[8]. In that study, experts were offered the opportunity to give imprecise probability ranges for the likelihood of triggering 7 different tipping points under 3 different temperature corridors, by the year 2200. A total of 52 experts responded across the 7 tipping points, and their names and affiliations are given in Table S2 of Kriegler et al.[8]. For each tipping point, the number of self-selected experts varied from 9-22 (their identities are listed in Table S1 of Kriegler et al.[8]). For 2 tipping point scenarios (dieback of boreal forests, and decline of ocean carbon sink) the number of experts willing to give imprecise probability statements was <10 and those results were excluded from the main analysis of Kriegler et al.[8]. Imprecise probability statements for the 5 remaining tipping points are summarized in Fig. 1 of Kriegler et al.[8]. Individuals sometimes gave large imprecise probability ranges for a particular tipping point and temperature corridor, and different experts sometimes disagreed considerably. These are reasons to support modeling a tipping point event in a (stochastic) Markovian fashion.

Figure S1 replicates the three different corridors for the evolution of temperature to 2200 from Kriegler et al.[8] (shaded gray). We also show the computed average temperature corridor and the benchmark deterministic DSICE temperature path without tipping. The latter is well represented by the low temperature corridor (the left part of Fig. S1) for this century and by the lower half of the medium corridor (the center part of Fig. S1) from 2100 onwards.

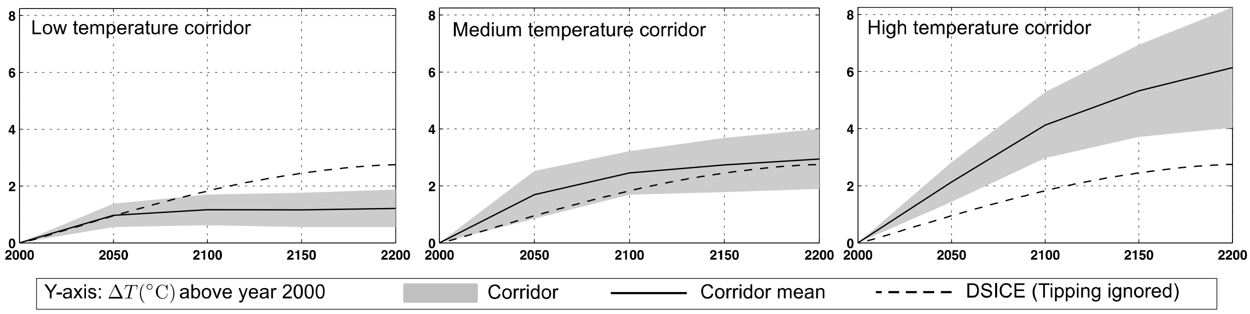


Fig. S1. Temperature corridors from Kriegler et al.[8]

Given the temperature corridors in Fig. S1, the climate experts expressed their beliefs about the occurrence of tipping point events in a probabilistic fashion for each of those temperature corridors. We use the experts' subjective beliefs (within a range of 0-8 Celsius of global warming) to calibrate the probability of a tipping point event. Our calibration method involves the following steps. First, with the three specified temperature corridors from Kriegler et al.[8] we perform a polynomial approximation on the lower bound, the mean, and the upper bound of each of these corridors. This leaves us with nine temperature paths. For our default case, we use the average of the medium temperature corridor, implying a warming of about 3°C by 2200 (relative to 2000 levels). Second, we specify the general relationship between the hazard rate and the contemporaneous temperature by $h_t = b_1(T_t\text{-}b_0)$ if $T_t>b_0$ (here b_0 =1). Third, using the approximated time path for the temperature scenario and integrating the hazard rate up to time t gives the cumulative hazard (H_t) with the probability that a tipping has occurred up to time t being $1 - \exp(-H_t)$.

At this point we make use of the range of cumulative probabilities of tipping at 2200 from the expert elicitation study[8]. For example, for the medium temperature corridor, for which temperature at 2200 ranges from 2°C to 4°C, the range of expert-assigned probabilities is between 37% and 100%. It extends downward to 16% when the experts are weighted according to their level of expertise. We do not consider the weighting results on the lower end. However, we also cut off the top 9.45% of the remaining range, leaving us with a probability range of 37% to 90.55%. In the next step, we assign the probability range to the temperature corridor, resulting in matched probabilities of 37% for 2°C, 90.55% for 4°C, and 63.78% (the average probability) for 3°C (the average temperature). The latter combination is used for our default case.

We first derive the hazard rate for the mean of the medium temperature corridor (which gives ~3 °C warming in 2200), and calibrated the cumulative probability of tipping using the average experts' results for the probability of tipping (at least 1 of 5 elements) given a 3 °C warming by 2200. With these numbers the hazard rate can be inferred by solving for the constant b_1 and assuming that b_0 =1. This implies that the probability of observing a tipping point today is zero. For our default case we obtain a hazard rate factor of ~ b_1 = 0.0025 which means if we observe 1 °C warming the probability of having a tipping event in that year is ~0.25%, and if we later observe 2 °C warming the probability rises to ~0.5% in that year (if tipping has not already occurred). We also study other combinations of probabilities and temperature paths. We proceed in the same fashion for the low and high temperature. Looking across the different temperature corridors (Table S2), the experts are internally (logically) consistent, with the average experts' hazard rate factor b_1 ~0.003 under the low temperature corridor and b_1 ~0.0025 under the medium and high temperature corridors. Furthermore, the difference in inferred hazard rates among experts narrows as the degree of global warming increases – i.e., experts' beliefs on tipping point probabilities converge.

Temperature Corridor	Low			Medium			High		
Temp. in 2200 Above 2000	0.56	1.21	1.87	1.89	2.95	4	4.04	6.14	8.25
Expert Type	opt.	Ø	pes.	opt.	Ø	pes.	opt.	Ø	pes.
Inferred hazard rate	0.001	0.003	0.0053	0.0018	0.0025	0.0041	0.0025	0.0025	0.0033

Table S2. Inferred hazard rate factors of triggering a tipping point event based on data from expert elicitations[8]. "Expert type" resembles the range of elicited probabilities of a tipping point event (any one of five) among experts (opt.: representative optimistic expert, pes.: representative pessimistic expert and Ø: average expert)

Our hazard rate formula implies a linear relationship between global warming and the hazard rate. While in principle any other relationship can be used, we found that if using a linear hazard rate, the experts' beliefs for different temperature corridors as reported in the elicitation study[8] exhibit a high degree of consistency. Table S2 summarizes the inferred hazard rates for the nine different temperature paths.

One alternative method to our hazard rate approach would be to formulate a tipping point event as a problem with an unknown threshold. However, a threshold formulation implies by definition that a tipping point event cannot occur in phases of global cooling. Given the fact that there are indeed some years in which the average global temperature decreases, we consider such a threshold formulation as unnecessarily limiting. Nevertheless, our specification is compatible with a threshold problem for any monotonically increasing temperature path. Our specification of tipping and its dependence on temperature is agnostic. Some experts might actually have had in mind a threshold model. Even for those experts, the hazard rate simply represents the conditional probability that at this temperature a climate element is passing its threshold. Consequently, with the hazard rate approach we can handle many different interpretations of the nature of tipping points. Furthermore, our assumption of the uncertainty faced by the decision maker is equivalent to assuming that the decision maker considers the experts' beliefs, which are reflected by hazard rates. Therefore, we abstain from arguing that the evolution of tipping elements is deterministic or stochastic. We rather assume that the decision maker has a lack of knowledge about that evolution and that the decision maker's uncertainty can be represented by a stochastic formulation.

Supplementary Results

Comparison of DSICE with MAGICC6 using RCP scenarios: The DSICE model[1,2] used in this study builds on the DICE-2007 model[4] in which the carbon cycle and the climate module are represented by a three-box and a two-box model respectively. This is considerably less than what is typically assumed in more complex climate models. Nevertheless, DICE-2007[4] is the only forward-looking model that has been used by the recent U.S. Government Interagency Working Group on Social Cost of Carbon for the analysis of optimal carbon taxation[9]. Here, we briefly study how the DSICE carbon and climate module compares to a more complex model. Fig. S2 illustrates the comparison of the benchmark deterministic DSICE carbon-climate module with MAGICC6[10], a more advanced and widely accepted box model. We compare the DSICE and MAGICC6 (default setting) global average temperature for four IPCC RCP scenarios until 2100. As Fig. S2 shows, DSICE delivers a good match of MAGICC6. For the highest exogenous emission scenario (RCP8.5) DSICE is off by 0.8°C at year 2100. However, here we used the benchmark MAGICC6 specification and our results are still within the range of temperature paths that result under alternative model settings in MAGICC6. Other specifications within MAGICC6 - which we do not show here - make us conclude that even 4°C of global warming by 2100 might be reasonable. At the same time, the deterministic (no tipping point) variant of DSICE used in this study delivers a temperature path very close to that of the RCP4.5 run (second from the left).

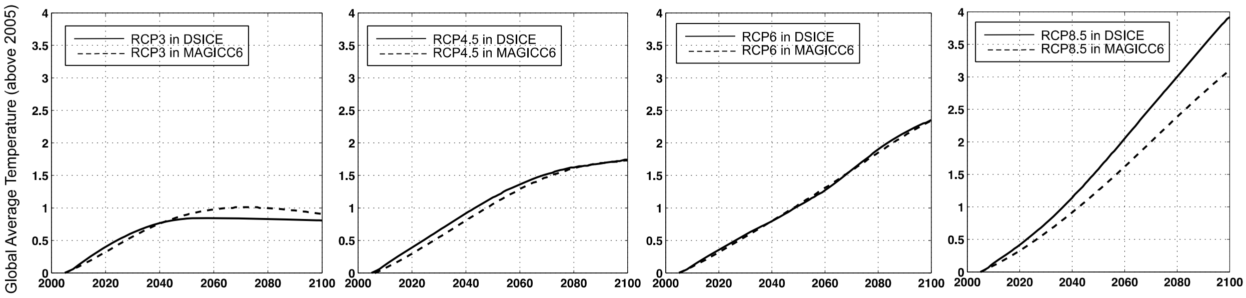


Fig. S2. Temperature responses to four RCP emission scenarios

DSICE model (default parameter specification): The right plot in Fig. S3 shows how our benchmark calibrated hazard rate translates into the cumulative probability of triggering the tipping point event. The cumulative probabilities are obtained by simulating 10,000 stochastic time paths of the model's optimal solution. Our benchmark case translates into a probability of about 97.5% that a tipping point event will occur after the year 2050 and a probability of about 86.5% that the tipping point event will not be triggered by the year 2100. These numbers compare well with those from Kriegler et al.[8] considering that our optimal temperature path lies clearly in the lower part of the medium temperature corridor (see Fig. S1).

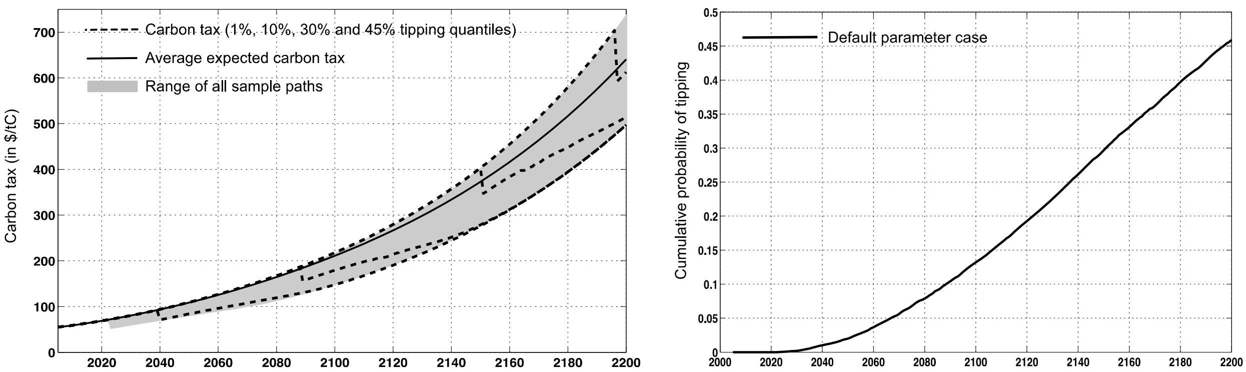


Fig. S3. Results of the stochastic optimization model DSICE. (**left**) - The optimal total carbon tax - statistical analysis of 10,000 Monte Carlo paths. (**right**) - Cumulative probability of triggering a tipping point event (average of 10,000 Monte Carlo paths).

To illustrate the results of the optimal carbon tax when the decision maker does account for the tipping point event, we perform Monte Carlo simulations using the optimal solution to the dynamic programming problem, generating 10,000 possible stochastic realizations of the model. The dashed lines in Fig. S3 denote the 1%, 10%, 30% and 45% quantiles of triggering the tipping point process and the gray area represents the range of all sample paths. For instance, 2022 is the earliest year in our 10,000 simulations at which a tipping point process has been triggered. By the year 2038, 1% of all Monte Carlo runs have triggered a tipping point process (left dashed line), by 2086 in 10% of our sample paths a tipping point process has been triggered, by 2150 in about 30% of our sample paths a tipping point process has been triggered and by 2195 in about 45% of our (right dashed line). These numbers can also be read off the right panel in Fig. S3. Note that this implies that it is as likely as not that a tipping point process will be triggered by the year 2200, with its full impacts taking about 50 years to fully unfold. Yet, as our results suggest, the optimal pre-trigger carbon tax for our default parameter specification (as specified by the upper envelope of the shaded area) increases strongly by about 50% when compared to a model version in which the tipping point process is ignored in the decision making process (see Fig. 4 in the main text).

One possible explanation for this significant increase in the optimal carbon tax lies in the Markovian structure of the tipping point event which allows for triggering that event much sooner, such as e.g., in 2038 as in our 1% quantile (left dashed line). Also, as our calculations show, this additional and substantial carbon taxation leads to an expected delay of the tipping point trigger of 20 years in our benchmark case from year 2214 to 2234 (see Fig. 4 in the main text).

Two other important observations are warranted. First, after the tipping point process has been triggered the optimal carbon tax drops significantly. This is because the uncertainty about the tipping point event has disappeared and no additional policy is required to address the tipping point externality, since the latter has already occurred.

Second, note the convex shape of the average expected carbon tax (solid black line), which reproduces the carbon tax ramping structure of the DICE-2007 model[4]. This ramping occurs because damages rise gradually with global warming and all future values are discounted. The convex shape is thus an artifact of the DICE-2007 model. Consequently, the general pattern of the carbon tax in the DSICE model will also exhibit this ramping since the deterministic damage structure from DICE-2007 is retained. The 2005 carbon tax in a purely deterministic version of our model, in which climate tipping is ignored is about \$36.7 and it increases to about \$173 by the end of this century, an average annual growth rate of about 1.68%. However, as Fig. 3 in the main text depicts, our benchmark model calibration produces a quite flat expected additional carbon tax to delay the triggering of the tipping point event. In fact, its average growth rate until 2100 is only about 0.81%. This is less than half the average growth rate of the deterministic carbon tax component (which is 1.68%).

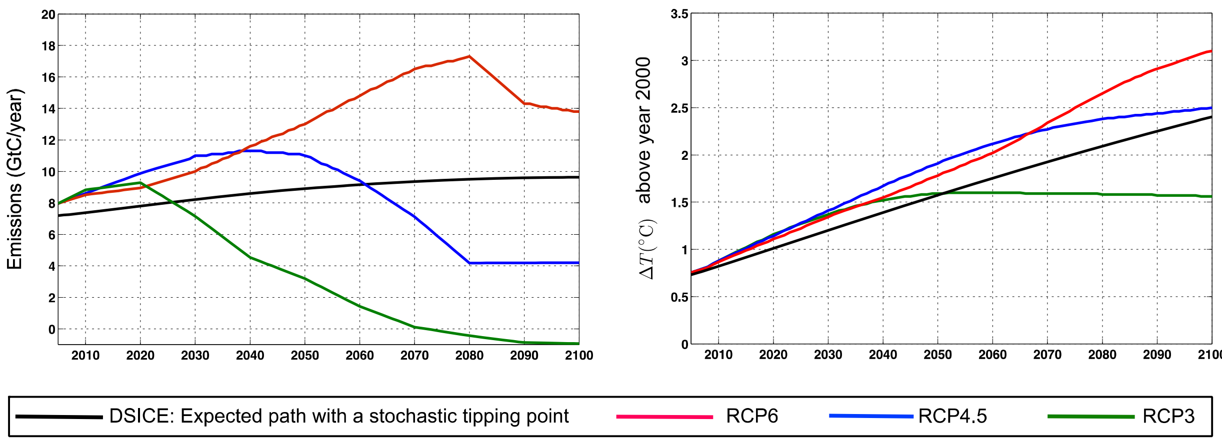


Fig. S4. Comparison of our model (default case) with RCP emission and temperature scenarios

Paths of the total carbon tax for selected cases of the tipping point parameters: In Fig. S5 we present the time paths of the total expected carbon tax (average of 10,000 simulated stochastic paths) for our default case and compare its paths resulting from some alternative parameterizations of the tipping point events (e.g. higher damage, higher hazard, lower transition time). Our default specification of the tipping point events results in today's optimal carbon tax of about \$55.6 tC^{-1} and the optimal carbon tax in year 2100 is about \$210.6 tC^{-1} under this scenario. Higher hazard rates (0.0045 instead of 0.0025), lower transition times (5 years instead of 50 years) and higher damage levels (20% instead of 10%) increase strongly today's optimal tax level, as quantified in Fig. 4 of the main text. The 2100 levels of the carbon tax are also much higher. In general, the paths in panel A indicate that the growth rate of the additional carbon tax is similar in these cases, as shown in Fig. 3 of the main text. In panel B, we compare our default parameter case to the most extreme parameter case in our sensitivity (20% damage, 5 years transition and 0.0045 hazard). In addition, we also show a case for which the damage is increased to 30% (red line). The latter case increases today's optimal total carbon tax to about \$280 tC^{-1}.

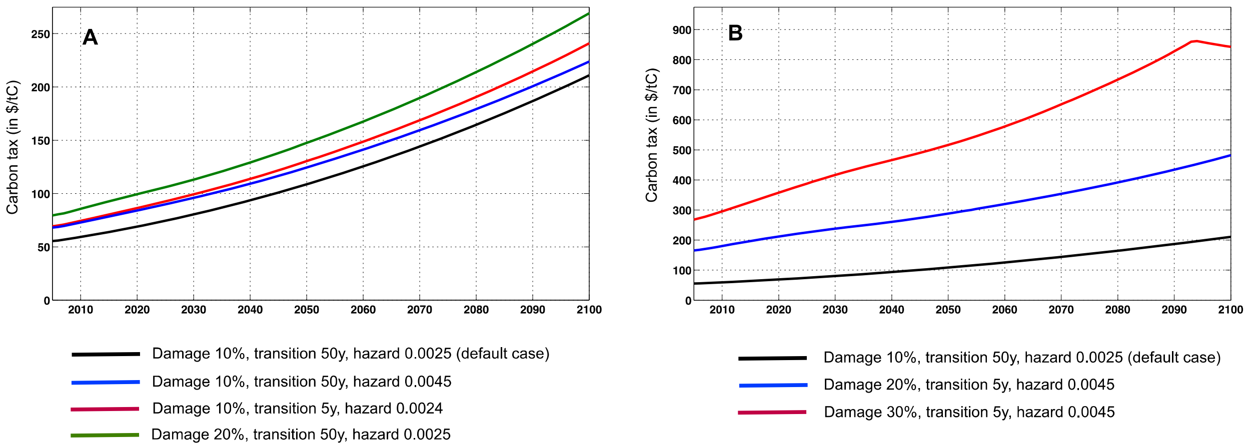


Fig. S5. Expected total carbon tax paths for alternative specifications of the tipping point parameters.

Supplementary Discussion

Comparison of our approach to other studies: The analysis performed in this study is based on three key novel elements: (a) we make the scientifically-grounded assumption that the effects of passing a tipping point are not instantaneous, they accumulate over time at a rate dictated by the transition time of the tipping element in question (in common with other studies, these impacts are irreversible); (b) we calibrate the hazard rate of a climate tipping point using the results of the most recent expert elicitation by Kriegler et al.[8]; (c) we are able to solve a higher-dimensional system of equations under stochastic uncertainty than in other studies and are thus able (in the deterministic special case) to reproduce the DICE-2007 model as used in the original US government social cost of carbon estimates.

Previous studies generally assume that tipping points are instantaneous, which is unrealistic because no part of the Earth system can transition instantaneously. Furthermore, most previous studies treat tipping points deterministically while only a few treat them stochastically. To explain how our work differs from existing studies we focus on those studies that may appear similar to our own, in particular, numerical integrated assessment model studies, especially those using variants of the DICE model, and those implementing a stochastic treatment of a tipping point/catastrophe. Here we review the scientific basis for the functions chosen to represent the likelihood and impacts of a tipping point/catastrophe, and the methodology used.

Earlier expert elicitation: Many existing studies use results for the likelihood of a climate catastrophe from an earlier expert elicitation by Nordhaus[11]. In that study experts provided a probability of a "high-consequence" outcome defined as a 25% loss in GDP indefinitely (akin to a permanent Great Depression), under scenarios of 3°C temperature rise by 2090, 6°C by 2175, or 6°C by 2090. The results differed considerably according to the background of participants and between the scenarios. An often-used result is the mean (across 19 participants) 4.8% probability of a catastrophe under 3°C in 2090. However, for this scenario, the 3 natural scientist participants gave a mean 12.3% probability whereas the 8 non-environmental economists gave a 0.4% probability. Mean probabilities across all participants for the other scenarios were 12.1% under 6°C by 2175 and 17.5% under 6°C by 2090 (suggesting sensitivity to the rate of warming). Some confusion has stemmed from statements of generally lower probabilities subsequently given by Nordhaus & Boyer[12] referring to the same expert elicitation. For example, they[12] state a mean probability of 0.6% for the 3°C warming in 2090 scenario (rather than 4.8%), and a 3.4% probability under 6°C warming in 2175 (rather than 12.1%). Nordhaus & Boyer[12] then double their stated probabilities based on concern at the time about a collapse of the thermohaline circulation (THC), which is Nordhaus' model for the catastrophe, although this was not made explicit to the participants. In Nordhaus & Boyer[12] the damage due to the catastrophe is increased (by 20%) to 30% permanent loss in income based on increased concern about the impacts of THC collapse.

Stochastic IAM studies: Gjerde et al.[13] was one of the first numerical integrated assessment model studies of a stochastic climate catastrophe, which assumed the hazard rate to be a convex function of temperature (above preindustrial) with power 1.5, calibrated on Nordhaus'[11] 4.8% probability of catastrophe under 3°C warming in 2090. Three catastrophe scenarios were considered, one where utility drops instantaneously and irreversibly to zero, one where it drops to 1990 levels, and one where it drops to zero but the hazard rate is independent of temperature (exogenous). The reduction of utility to e.g. a constant 1990 level implies that the economic damage due to a particular tipping point is relatively small if it happens soon and can grow hugely the further in the future it occurs. A percentage decrease in utility would have been more defensible. More fundamentally, with stochasticity only applied to the utility function – and not changing the climate or economic system – this is a certainty-equivalent method, so it can be easily transformed to be a deterministic optimal control problem that can be solved by an optimization solver directly.

A subsequent study by Castelnuovo et al.[14], in turn building on a working paper by Bosello & Moretto[15], follows Gjerde et al.[13] in assuming that an uncertain level of global temperature triggers an irreversible drop in utility to 1990 levels (or to zero, as a sensitivity study). Again, these studies only apply stochasticity to the utility function making it a certainty-equivalent method, which can be transformed to a deterministic optimal control problem that can be solved by an optimization solver directly. Castelnuovo et al.[14] use the ETC-RICE model, whereas Bosello & Moretto[15] use different integrated assessment models. Both endogenous and exogenous hazard rates are considered and in the endogenous case the catastrophe requires temperature to increase in a given year (which is overly-restrictive). Large reductions in emissions are predicted as optimal, but they don't depend on whether the irreversible drop in utility is to 1990 levels or to zero – probably because the drops in utility is already very large, as by the time the catastrophe happens, GDP has risen far above 1990 levels. In one variant of the endogenous case of Castelnuovo et al.[14], a dependence of the hazard rate on the rate of temperature change is included and calibrated on Nordhaus'[11] 4.8% probability of a catastrophe under 3°C warming in 2090. The probability is assumed to be 0% for the 1990 climate state. It is further assumed to be <0.1% in 2010-2020, but this is erroneously justified with reference to Bentley[16] who argues specifically that <0.1% is the probability of a West Antarctic Ice Sheet collapse over the next *two centuries*. This erroneous assumption requires a (therefore unjustified) probability of catastrophe that is non-linear with time and steeply rising late in this century. The authors acknowledge that their chosen values have "weak scientific foundations".

Recently, Lemoine & Traeger[17] added a representation of tipping points to a stochastic-dynamic IAM based on a simplified, 3-dimensional version of the 6-dimensional DICE-2007, developed in a working paper by Crost & Traeger[18]. Neither the carbon cycle nor ocean temperature is included in the model – there is just atmospheric carbon and atmospheric temperature – with 3°C climate sensitivity.

Lemoine & Traeger[17] use a very different formulation of a climate ‘tipping point’ to ours. Their ‘tipping point’ cannot occur if temperature falls from one year to the next – which seems overly restrictive. The ‘tipping point’ can be one of two types (leading to two different versions of the model) – an instantaneous, irreversible increase in climate sensitivity (from 3°C to 4, 5 or 6°C), or an instantaneous, irreversible weakening of carbon sinks (by 25, 50 or 75%), which in turn feeds into the standard DICE formulation of damages (quadratically increasing with temperature). This is different from a tipping point that directly and irreversibly decreases GDP. Neither of the formulations of a tipping point suggested by Lemoine & Traeger[17] is particularly scientifically plausible.

To support the argument for an abrupt increase in climate sensitivity, Lemoine & Traeger[17] discuss two examples of positive climate feedbacks. However, positive feedbacks are never instantaneously switched on – instead they may get progressively stronger as temperature increases – so an instantaneous ‘tipping’ formulation is qualitatively wrong. Furthermore, the actual examples given are implausible as sources of strong and rapid positive feedback. The first example given is warming mobilizing large methane stocks resulting in further warming that mobilizes further methane stocks. Current assessments[19] are that this feedback has modest strength on a century timescale, potentially enhancing warming by up to ~10%. Furthermore, the feedback is roughly proportional to warming – it does not suddenly switch on. The second example given is the retreat of land ice sheets lowering surface albedo and thus increasing warming, but this is an inherently slow and weak positive feedback because the timescale of ice sheet retreat is centuries to millennia, and they cover a relatively small area.

Lemoine & Traeger[17] support their argument for an instantaneous weakening of carbon sinks, with reference to positive climate feedbacks on ocean and land carbon storage. Again, such feedbacks do not suddenly switch on. Instead, they may increase in strength as temperature increases, for example if Amazon rainforest dieback is triggered, providing an additional source of CO_2.

Finally, Lemoine & Traeger[17] assume that the hazard rate of tipping increases linearly with temperature. However, in sharp contrast to our study (where climate tipping is always stochastic), the baseline assumption in Lemoine & Traeger[17] is that tipping is certain at 4.27°C (with a sensitivity range of 3-9°C), which is not supported by available scientific information. Aggregated imprecise probabilities from expert elicitation by Kriegler et al.[8] suggest tipping is more likely than not in a 4-8°C long-term warming scenario, but not certain.

Our model is also so far the only stochastic integrated assessment model, which embodies the full 6-dimensional DICE model on short time-steps with a computational structure that is able to deal with a more realistic formulation of climate tipping points. Future research with DSICE will also embody ideas from ongoing studies by e.g. Martin & Pindyck[20] on multiple catastrophes or van der Ploeg & de Zeeuw[21] on different shapes of the hazard function for climate tipping. The study by van der Ploeg & de Zeeuw[21] is the first to investigate the effects of climate tipping on precautionary capital accumulation, suggesting that climate tipping requires precautionary accumulation of capital together with an additional price on carbon emissions. Our model, in part captures this effect indirectly, but we do not focus on the decomposition explicitly.

Deterministic IAM studies and critiques: Studies that treat tipping points as deterministic are qualitatively different to our approach. Assuming deterministic (i.e. perfect) knowledge of the likelihood of catastrophe is not scientifically reasonable given that even for a deterministic tipping point, stochastic internal variability in the climate system will tip the system at an uncertain time, before the deterministic tipping point is reached. This intrinsic uncertainty is compounded by uncertainty in our present knowledge about the location of tipping points.

Kosugi[22] offers a deterministic version of the DICE-2007 model with no stochasticity and instantaneous, irreversible impacts of a tipping point, giving results in terms of emissions paths not social cost of carbon. The author explores a range of parameter settings for the probability of abrupt change (under 2.5°C rise in 2090 from 1900, using default 1.2%, and range 0-30%), the exponent of the hazard function (either 2 or 12), and the impact of abrupt change (default 30%, range 0-100%), where the first two parameters jointly determine the intercept of the hazard function. The default settings of 1.2% probability and 30% impact are taken from Nordhaus & Boyer[12] but (as noted above) are inconsistent with those originally given by Nordhaus[11]. Setting this discrepancy aside, the implied hazard rate from Nordhaus'[11] experts appears to be near quadratic, so the 12th power explored by Kosugi[22] must be viewed as an extreme sensitivity test. The default value of instantaneous 30% reduction in GDP (taken from Nordhaus & Boyer[12]) is also very high. In the sensitivity analysis, an upper bound of 30% probability of abrupt change in 2090 is considered, which when combined with a 12th power hazard rate gives a 100% probability (i.e. certainty) of catastrophe early in the 21st century, at around 3°C warming. This is far above results from expert elicitation[8].

Weitzman[23] offers a deterministic, analytical model study, which elegantly critiques the assumption of a quadratic damage function, as assumed in DICE and many other models – showing that such a function is relatively insensitive to very high warming levels and is outweighed by discounting. Weitzman[23] focuses on the possibility that climate sensitivity could be very high (in other words, there are strong positive feedbacks in the climate system as a whole). Weitzman[23] considers two alternative fat-tailed probability density functions for climate sensitivity, and introduces an additional damage function that is approximately a 7th power of temperature, giving 50% loss in output at 6°C and 99% loss in output at 12°C warming. With this formulation, Weitzman[23] shows that a normally risk-averse agent should be willing to pay a significant amount now to reduce the likelihood of a catastrophic outcome a long way into the future. Weitzman[23] demonstrates this by having all the damaging effects of temperature be stored up and then felt instantaneously, 150 years from now. Despite strong discounting (6% interest rate) the fat-tailed probability of high temperature outcomes combined with the high damages from high temperatures still register now. The approach of Weitzman[23] could be described as considering an 'impacts tipping point' under smooth climate change, especially in his 150-year scenario.

Ackerman & Stanton[24] use a deterministic version of the DICE model and do not explicitly consider climate-tipping points. They build on Weitzman[23] by including as an option his 7th power damage function, also experimenting with combining it with Hanemann's[25] estimate of greater damages than Nordhaus under low warming (2.5°C) – this damage function can be seen as an impacts 'catastrophe' centered on 6°C warming. The analysis also looks at two discount rates (1.5% and 3%) and considers average (<3°C) or 95th percentile (7.14°C) climate sensitivity. For various combinations of these assumptions Ackerman & Stanton[24] obtain very high values of the social cost of carbon. They conclude that given the possibility of a catastrophic outcome an insurance approach is needed which lowers the deterministic probability of catastrophe by reducing emissions to zero as swiftly as is feasible. In our model, we take the original parameter values from DICE2007 – if we applied the same sensitivity analysis on the parameters we would also get much higher results for the social cost of carbon.

As part of our research, we studied the deterministic approximations of a climate tipping point as advocated by e.g. Weitzman[23] and Ackerman & Stanton[24] – and described in the previous paragraphs. As Fig. 3 in the main text shows, these deterministic approximations which attempt to represent the potential damages from climate tipping by increasing the convexity of the damage function fail to capture the flattening of the growth rate of the carbon tax, which occurs when the tipping point is formulated in a stochastic manner. Instead they tend to have the opposite effect.

References

1 Cai, Y., Judd, K.L., & Lontzek, T.S., DSICE: A dynamic stochastic integrated model of climate and economy. *The Center for Robust Decision Making on Climate and Energy Policy Working Paper* No. 12-02 (2012).

2 Cai, Y., Judd, K.L., & Lontzek, T.S., The Social Cost of Stochastic and Irreversible Climate Change. *National Bureau of Economic Research Working Paper Series* No. 18704 (2013).

3 Cai, Y., Judd, K.L., & Lontzek, T.S., Continuous-Time Methods for Integrated Assessment Models. *National Bureau of Economic Research Working Paper Series* No. 18365 (2012).
4 Nordhaus, W.D., *A Question of Balance: Weighing the Options on Global Warming Policies*. (Yale University Press, 2008).
5 Nordhaus, W., *The Climate Casino: Risk, Uncertainty, and Economics for a Warming World*. (Yale University Press, 2013).
6 Cai, Y. & Judd, K.L., Stable and efficient computational methods for dynamic programming. *Journal of the European Economic Association* 8 (2-3), 626-634 (2010).
7 Lenton, T.M. *et al.*, Tipping Elements in the Earth's Climate System. *Proceedings of the National Academy of Sciences* 105 (6), 1786-1793 (2008).
8 Kriegler, E., Hall, J.W., Held, H., Dawson, R., & Schellnhuber, H.J., Imprecise probability assessment of tipping points in the climate system. *PNAS* 106 (13), 5041-5046 (2009).
9 Interagency Working Group on Social Cost of Carbon, Social Cost of Carbon for Regulatory Impact Analysis - Under Executive Order 12866. United States Government, 2010.
10 Meinshausen, M., Raper, S.C.B., & Wigley, T.M.L., Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6 - Part 1: Model description and calibration. *Atmos. Chem. Phys.* 11 (4), 1417-1456 (2011).
11 Nordhaus, W.D., Expert Opinion on Climatic Change. *American Scientist* 82, 45-51 (1994).
12 Nordhaus, W.D. & Boyer, J., *Warming the World. Models of Global Warming*. (MIT Press, Cambridge, Massachusetts, USA, 2000).
13 Gjerde, J., Grepperud, S., & Kverndokk, S., Optimal climate policy under the possibility of a catastrophe. *Resource and Energy Economics* 21 (3-4), 289-317 (1999).
14 Castelnuovo, E., Moretto, M., & Vergalli, S., Global Warming, Uncertainty and Endogenous Technical Change. *Environmental Modeling & Assessment* 8 (4), 291-301 (2003).
15 Bosello, F. & Moretto, M., Dynamic uncertainty and global warming risk. Fondazione Eni Enrico Mattei (FEEM) Working Paper 80.99. FEEM, 1999.
16 Bentley, C.R., Rapid Sea-Level Rise Soon from West Antarctic Ice-Sheet Collapse? *Science* 275 (5303), 1077-1078 (1997).
17 Lemoine, D. & Traeger, C., Watch Your Step: Optimal Policy in a Tipping Climate. *American Economic Journal: Economic Policy* 6 (1), 137-166 (2014).
18 Crost, B. & Traeger, C., Risk and aversion in the integrated assessment of climate change. CUDARE Working Paper No. 1104. University of California eScholarship, 2011.
19 Schaefer, K., Lantuit, H., Romanovsky, V.E., G., S.E.A., & Witt, R., The impact of the permafrost carbon feedback on global climate. *Environmental Research Letters* 9, 085003 (2014).
20 Martin, I.W.R. & Pindyck, R.S., Averting Catastrophes: The Strange Economics of Scylla and Charybdis. *National Bureau of Economic Research Working Paper Series* No. 20215 (2014).
21 van der Ploeg, F. & de Zeeuw, A., Climate Policy and Catastrophic Change: Be Prepared and Avert Risk. OxCarre Research Paper 118, 2013.
22 Kosugi, T., Integrated Assessment for Setting Greenhouse Gas Emission Targets under the Condition of Great Uncertainty about the Probability and Impact of Abrupt Climate Change. *Journal of Environmental Informatics* 14 (2), 89-99 (2009).
23 Weitzman, M.L., GHG Targets as Insurance Against Catastrophic Climate Damages. *Journal of Public Economic Theory* 14 (2), 221-244 (2012).
24 Ackerman, F. & Stanton, E.A., Climate Risks and Carbon Prices: Revising the Social Cost of Carbon *Economics: The Open-Access, Open Assessment E-Journal* 6 (2012-10), 1-25 (2012).
25 Hanemann, W.M., What is the economic cost of climate change? Department of Agricultural and Resource Economics, UC Berkeley, 2008.