# Tipping Points and Ambiguity in the Integrated Assessment of Climate Change*

Derek M. Lemoine† & Christian Traeger§

† Department of Economics, University of Arizona
McClelland Hall 401, PO Box 210108, Tucson, AZ, 85721-0108, USA
dlemoine@berkeley.edu

§ Department of Agricultural & Resource Economics, University of California, Berkeley
207 Giannini Hall #3310, Berkeley, CA 94720-3310, USA
traeger@berkeley.edu

International Energy Workshop 2011

This Version: July 3, 2011, First Version: October 2010

**Abstract:** The threat of crossing tipping points in the climate system often serves as an argument for more stringent greenhouse gas emission reductions. We introduce such regime shifts into a recursive relative of the DICE integrated assessment model for establishing optimal climate policies. A tipping point's effects on system dynamics are irreversible and are triggered by crossing an unknown temperature threshold. The tipping point's probability and timing are therefore endogenously determined by the chosen emission policy. The policymaker can display ambiguity aversion in assessing tipping point uncertainty, which recognizes that the probability distribution for the temperature threshold is not known with confidence. Our simulations show that tipping points can increase the near-term social cost of carbon by 50% under reasonable assumptions. Regime shifts that directly increase temperature or damages for a given $CO_2$ concentration have a stronger impact on optimal policy than do regime shifts that increase the atmospheric lifetime of $CO_2$. The possibility of a tipping point is more important for the social cost of carbon than is ambiguity about the temperature threshold.

**JEL Codes:** Q54, D81, Q48, Q00, D61, D62, D90, C63

**Keywords:** climate change, tipping points, thresholds, ambiguity aversion, uncertainty, integrated assessment, risk aversion, dynamic programming

*We are grateful to Larry Karp, Eric Nævdal, and Benjamin Crost for valuable comments.

## 1 Introduction

The threat of climate tipping points motivates calls for aggressive near-term emission reductions to prevent global average temperature from increasing by 2°C relative to pre-industrial levels (e.g., Hansen et al., 2008; Ramanathan and Feng, 2008; Rockström et al., 2009). Tipping points—or irreversible changes in the climate system caused by increasing carbon dioxide ($CO_2$) and temperature—are poorly understood, difficult to model, and of increasing concern (Alley et al., 2003; Overpeck and Cole, 2006; Smith et al., 2009). By connecting today's climate policy decisions to expectations of future wealth and warming, integrated assessment models (IAMs) estimate the social cost of carbon for use in evaluating emission policies and in cost-benefit analyses. However, because they have not included tipping point possibilities in their system dynamics, these IAMs might consistently underestimate the social cost of carbon.

We answer three questions about the implications of tipping points for IAMs' policy outcomes. First, how do different kinds of tipping points affect the social cost of carbon and optimal abatement policy? Second, how sensitive is optimal policy to prior beliefs about a temperature threshold's location? Third, is aversion to ambiguity about the threshold's location important for the social cost of carbon? We find that tipping points which increase the response of temperature to $CO_2$ are more important than tipping points which increase either the damages incurred by high temperatures or the quantity of non-$CO_2$ greenhouse gases. The least important of the studied tipping points increases the atmospheric lifetime of $CO_2$. Optimal policy is not highly sensitive to the form of the threshold distribution, and it avoids the temperature trigger if it is known to be at 2.5°C or above. Aversion to the ambiguity arising from lack of knowledge about the location of the temperature threshold only mildly affects near-term optimal policy.

To answer these questions, we extend a standard IAM to include endogenous regime shifts, learning about the threshold that triggers a regime shift, and a welfare evaluation based on the smooth ambiguity model. Our first extension incorporates tipping points into a recursive version of the Dynamic Integrated model of Climate and the Economy (DICE), a welfare-optimizing IAM (Nordhaus, 2008). A tipping point occurs upon crossing some temperature threshold. Each tipping point irreversibly changes the climate system from its conventional representation in DICE to a new regime that depends on the type of tipping point under consideration. Emission decisions affect whether tipping points occur by determining the temperature expected in each period. The decision-maker anticipates how it would choose emissions and consumption in the post-threshold world. The timing, probability, and welfare consequences of a regime switch are endogenous because they depend on the policies chosen before and after the threshold occurs.

We explore four tipping points (Table 1), whose effects are illustrated by the thick arrows in Figure 1.[2] The first tipping point changes the response of temperature to increased $CO_2$ concentrations. More specifically, it increases climate sensitivity, which is the equilibrium warming produced by doubling $CO_2$ concentrations. This change would occur if land ice sheets begin to retreat on a shorter-than-expected timescale. The second tipping point affects the relation between

[2] We do not model tipping points that might increase welfare because beneficial tipping points are not much discussed in the climate science literature. An interesting analogue of a beneficial tipping point is a new technology that changes the dynamics governing abatement cost. We also do not model tipping points that bring the planet to a cold equilibrium (Zaliapin and Ghil, 2010).

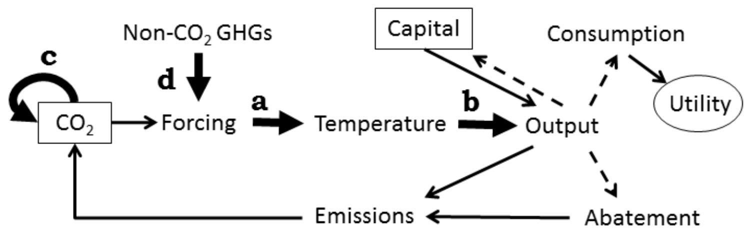



Figure 1: A simplified schematic of the modeled relation between the economy and the climate. Dashed arrows indicate the decision variables of consumption, investment, and abatement, and boxes indicate stock variables. Tipping points alter the relationships shown with thick arrows and are labeled as: a) increased climate sensitivity, b) increased convexity of damages, c) weakened $CO_2$ sinks, and d) increased non-$CO_2$ forcing.

Table 1: The four tipping points that can occur from crossing a temperature threshold. Each model run uses only one of these post-threshold regimes.

| Post-threshold regime | Threshold event |
| --- | --- |
| Climate sensitivity increased from 3°C to 6°C | Land ice sheets retreat faster than expected |
| Damages become a cubic function of temperature | West Antarctic ice sheet collapses |
| $CO_2$ decay reduced by 75% | $CO_2$ sinks weaken |
| Non-$CO_2$ forcing increased by 1.5 W $m^{-2}$ | Warming releases methane from hydrates |

temperature and output. It increases the convexity of the damage function, which determines the percentage of world GDP that is lost when temperature increases. The convexity of the damage function would be increased by abrupt events that raise sea level (as in the collapse of the West Antarctic or Greenland ice sheets) or that produce unexpectedly nonlinear responses in agricultural production (as in sudden shifts in rainfall patterns). The third tipping point reflects the possibility that carbon sinks weaken beyond the predictions of the simple carbon cycle model represented in DICE. These weakened sinks decrease the decay rate of $CO_2$, which in turn increases the time for which emitted $CO_2$ affects the atmosphere. Finally, the fourth tipping point produces a permanent increase in forcing from non-$CO_2$ greenhouse gases. This corresponds to a large, sustained release of methane from melting permafrost or subsea hydrates, which would raise greenhouse gas concentrations independently of further economic activity.

Most IAMs assume that the dynamics governing the climate and the economy evolve smoothly over time.[3] However, the tipping points described above can be relatively abrupt (Hansen et al., 2008; Lenton et al., 2008). The omission of tipping points has long been recognized as potentially important (e.g., Nordhaus, 1993; Hall and Behl, 2006). One form of sudden change is a generic catastrophic impact that reduces every future period's utility (Gjerde et al., 1999). In contrast,

[3]In DICE-2007, some abrupt damages are included in the calculation of willingness-to-pay used to parameterize the damage function (Nordhaus, 2008), but the dynamics of passing into catastrophic scenarios are not modeled.

we model sudden temperature-induced changes as altering underlying system dynamics. Further, unlike the more restricted policy setting of Lempert et al. (1994), we embed these shifts in a model of dynamic policy that optimizes welfare at each timestep. Both before and after crossing the threshold, our decision-maker optimizes consumption, $CO_2$ abatement, and, via residual output, investment.

Our second extension recognizes that the temperature threshold for a tipping point is unknown. The decision-maker learns about the temperature threshold by observing whether a threshold has or has not been crossed as the world reaches higher temperatures. The chosen emissions determine the probability that a tipping point occurs, and this probability itself depends on the temperature produced by previous emission decisions. Increasing global temperature produces a greater chance of a tipping point occurring when temperature is already high. Both the probability of a tipping point occurring and the decision-maker's knowledge of the probability distribution are endogenous in the sense that they depend on chosen emission levels. Similarly to our work, Keller et al. (2004) extended DICE to endogenously determine when a threshold is potentially crossed. They considered an abrupt collapse of the thermohaline circulation (or Gulf Stream) that permanently shifts the damage function. Our model differs by including pervasive temperature stochasticity and by having the decision-maker endogenously learn about the threshold's location through the chosen emission path.

Several studies have analyzed uncertainty in DICE by drawing model parameters from a distribution and then determining optimal policy under certainty for each realization (e.g., Nordhaus, 2008; Ackerman et al., 2010). They approximated the optimal policy under uncertainty as the average policy from the various deterministic model runs. The resulting policy is usually not the same as optimal policy under uncertainty (Newbold and Daigneault, 2009). Other models have optimized consumption in the face of persistent stochasticity but imposed exogenous greenhouse gas policies (Gerst et al., 2010). We instead convert DICE into a recursive dynamic programming model (compare Kelly and Kolstad, 1999; Leach, 2007; Crost and Traeger, 2010). This enables us to analyze optimal policies under uncertainty about temperature change and about the temperature threshold that triggers a regime shift. Each period's optimal policies reflect the possibility that the next period's temperature will not be as expected and, moreover, that a climate threshold will be crossed.

By extending a full numerical IAM, we estimate the effect of uncertain tipping points on the social cost of carbon. Other relevant work has investigated the effects of uncertain climate tipping points in less complex models. It has considered how the possibility of a climate catastrophe affects the optimal steady state $CO_2$ concentration (Tsur and Zemel, 1996), how the endogenous risk of especially high-damage outcomes affects irreversible investment in abatement capital (Fisher and Narain, 2003), and how hyperbolic discounting increases the salience of a low probability, distant catastrophe (Karp and Tsur, 2007). Another pair of studies considered optimal control problems in the presence of uncertain thresholds that cause a stream of disutility (Nævdal, 2006; Nævdal and Oppenheimer, 2007). Finally, if emissions cannot be fine-tuned but instead are controlled only by discrete policies of predefined magnitude, then the policymaker must choose when to adopt the emission policies. This timing depends in part on how their adoption affects the possibility of a catastrophe (Baranzini et al., 2003; Guillerminet and Tol, 2008).

Beyond the climate context, other work has considered how an endogenous, uncertain threshold that makes pollution stop decaying altogether can produce multiple optimal paths for consumption and environmental quality (Ayong Le Kama et al., 2011). This threshold's effect on system dynamics is similar to our tipping point with weakened $CO_2$ sinks. A separate effort analytically modeled a shift in pollutant loading upon crossing a stochastic, reversible threshold (Brozović and Schlenker, 2011). They found that precaution is non-monotonic in the variance of the distribution for the threshold level: precaution at first increases with the variance as pollution levels just below the expected threshold have a greater chance of crossing it, but precaution eventually decreases when the variance becomes especially high because more probability mass then falls on threshold values that are either too high or too low to warrant precaution.[4] In contrast, we consider a fixed, irreversible threshold and model learning about its location. The possibility of collapse or altered system dynamics also arises in the optimal use of renewable resources. Polasky et al. (in press) considered a regime shift that reduces the resource's growth function. They found that the pre-threshold policy becomes more precautionary when the possibility of the regime shift is endogenous. Our work also examines the implications of an endogenous regime shift that affects system dynamics, but we do so with a nonlinear utility function that allows for risk aversion and we explore several different types of regime shifts.

Our third extension of the standard IAM distinguishes different types of uncertainty when evaluating welfare. Specifically, our decision-maker can be more averse to poorly understood tipping point uncertainty than to better understood temperature risk. We use "objective risk" to describe uncertain outcomes whose probabilities can be gleaned from data. In our model, this describes temperature stochasticity. We use the term "subjective uncertainty" to describe uncertain outcomes when there is less information available for determining probabilities. This deficiency in probabilistic knowledge applies to tipping points, which are more poorly understood than other climate phenomena (Alley et al., 2003; Lenton et al., 2008; Ramanathan and Feng, 2008; Kriegler et al., 2009; Smith et al., 2009). Distinguishing between types of uncertainty reaches back to Keynes (1921: Chapter VI) and Ellsberg (1961), who added an additional confidence weight to probabilities in order to describe uncertainty more comprehensively. We employ the Klibanoff et al. (2005, 2009) smooth ambiguity model to capture the decision-maker's ambiguity attitude, or the decision-maker's attitude to varying confidence in probability judgments. The model is consistent with conventional decision-theoretic axioms, including time consistency and a minimally modified version of the von Neumann and Morgenstern (1944) axioms (Traeger, 2010). Lange and Treich (2008) applied the Klibanoff et al. (2005) model to analytically explore the implications of ambiguity about a climate damage parameter. Millner et al. (2010) applied the multiperiod evaluation function of Klibanoff et al. (2009) to evaluate exogenous consumption paths from DICE simulations under ambiguity about climate sensitivity. They explored the response of this evaluation to changes in ambiguity attitude. Finally, Hennlock (2009) applied robust control theory to a two-sector analytic version of DICE. The robust control model can be interpreted as the limiting case of extreme ambiguity aversion in our smooth ambiguity framework. Our work differs from these in numerically implementing a model of ambiguity aversion that is not only used to generate endogenous policy paths but also entails optimal updating of the ambiguous distribution.

[4]Because we have exogenous variables evolving in time, one additional difference between our model and that of Brozović and Schlenker (2011) is that the cost of the policies needed to stay just below a known threshold changes over time.

Section 2 explains how we extend DICE, emphasizing the inclusion of tipping points and the generalized welfare evaluation. In section 3, we discuss the details of our four tipping point scenarios. Section 4 presents the results for each tipping point with a known threshold, with an unknown threshold, and with ambiguity aversion. We conclude in section 5 with a discussion of the implications for climate science, economics, and policy. The two appendices describe the model calibration and equations.

## 2 Introducing tipping points and ambiguity aversion into DICE

Our infinite horizon model extends Crost and Traeger (2010), which is a stochastic dynamic programming relative of the DICE-2007 model by Nordhaus (2008). DICE is a Ramsey-Cass-Koopmans growth model that has an aggregate world economy interacting with a climate module. Gross economic output (or potential GDP) is determined by an endogenous capital stock, an exogenously growing labor force, and exogenously improving production technology. Gross output produces $CO_2$ emissions. Non-abated $CO_2$ emissions accumulate in the atmosphere and ultimately translate into global warming, which causes damage proportional to world output. The control variables of the model are abatement and consumption, and residual output not allocated to these two options becomes capital investment (Figure 1). We implement a recursive modeling structure for the following reasons. First, we can account for temperature stochasticity. Second, we can model endogenous regime shifts and learning about the temperature thresholds that trigger them. Third, we can employ a more comprehensive approach to welfare evaluation that distinguishes between aversion to temperature risk and aversion to ambiguous tipping point occurrences. In order to replicate DICE in a stochastic dynamic programming framework, we use the state variables of capital, the stock of $CO_2$ in the atmosphere, and time. But in order to avoid the curse of dimensionality, we approximate the carbon cycle and the cooling effect of ocean heat capacity as interpolated functions of the $CO_2$ stock and time rather than as additional state variables. Moreover, we translate DICE's intrinsic warming delay into increased strength of the cooling reservoirs.[5] The two appendices provide more details on the equations governing the economy and the climate system as well as on the calibration and interpolation procedures.

In the following, we explain the richer uncertainty and welfare structure of our model. We represent maximized aggregate welfare by the function $V_\psi(k_t, M_t, t)$, where $\psi$ indicates whether a threshold has been crossed. The state variables are capital $k_t$ (per effective unit of labor), the stock $M_t$ of $CO_2$ in the atmosphere, and time $t$.[6] Crossing the threshold causes an irreversible switch from a pre-threshold regime (indicated by $\psi_t = 0$) to a post-threshold regime (indicated by $\psi_t = 1$). A regime switch immediately after period $t$ means that $\psi_s = 0\ \forall s \leq t$ and $\psi_s = 1\ \forall s > t$. Welfare in a period is defined by aggregating immediate utility $u(c_t) = c_t^{1-\eta}/(1-\eta)$ from per capita consumption $c_t$ with the expectation of future welfare. The parameter $\eta$ is the Arrow-Pratt measure of relative risk aversion, and $\eta^{-1}$ gives the intertemporal elasticity of substitution. Thus,

[5] We are working on implementing our model with a temperature state variable to include DICE's delay equation for temperature. We currently adjust the transient feedback (or cooling reservoir) to calibrate our baseline model to DICE's optimized output.

[6] Including time as a state variable enables other variables to evolve exogenously.

$\eta$ captures the preference for consumption smoothing over time and over risk. Once the regime switch occurs, all that matters for the decision problem are the post-threshold equations of motion and the state variables; the pre-threshold value function $V_{\psi=0}(k_t, M_t, t)$ is irrelevant after crossing a threshold. The post-threshold value function $V_{\psi=1}(k_t, M_t, t)$ is defined recursively by optimizing over consumption $c_t$ and the abatement rate $\mu_t$ so as to maximize welfare under the constraints of the equations of motion governing the economy and the climate system described in Appendix 2. This gives the dynamic programming equation:[7]

$$\begin{aligned} V_{\psi=1}(k_t, M_t, t) &= \max_{c_t,\mu_t} \frac{c_t^{1-\eta}}{1-\eta} + \beta_t \, \mathrm{E}_t \left[ V_{\psi=1}(k_{t+1}, M_{t+1}, t+1) \right] \\ &= \max_{c_t,\mu_t} \frac{c_t^{1-\eta}}{1-\eta} + \beta_t \int V_{\psi=1}(k_{t+1}, M_{t+1}, t+1) \; d\mathbb{P} \;, \end{aligned} \tag{1}$$

with the random variable affecting the output constraint as described below. Future welfare is discounted with the factor $\beta_t$. The discount factor captures a rate of pure time preference of 1.5% as in DICE-2007. It also adjusts for population growth so that equation (1) effectively measures welfare as the population-weighted sum of instantaneous per capita utility.[8] As discussed below, we solve the dynamic programming problem for the value function $V_{\psi=1}$ numerically for four different post-threshold regimes. We use a function iteration approach employing Chebychev polynomials for the value function approximation at a set of Chebychev nodes in the three-dimensional state space.

Once the threshold has already been crossed, welfare in period $t+1$ is uncertain only because of climatic fluctuations affecting the realized temperature. Cumulative temperature change affects the total output available for allocation to consumption, abatement, and, as a residual, investment (Appendix 2). Temperature is determined by an independent, normally distributed multiplicative shock having probability measure $\mathbb{P}$. We calibrate the mean-1 shock to the years 1881-2010 in the NASA Goddard Institute for Space Studies (GISS) Surface Temperature Analysis dataset.[9] We take expected temperature in each year to be the mean of the surrounding 10 years' realized temperatures. The estimated standard deviation of the resulting time series of multiplicative shocks is 0.0061.[10] This multiplicative noise captures period-to-period temperature variability that makes extreme outcomes more likely as $CO_2$ increases.

Once we have solved for the post-threshold value function, we use it to solve the pre-threshold value function. We assume that the system passes from the pre-threshold regime ($\psi_t = 0$) into the post-threshold regime ($\psi_{t+1} = 1$) if the expected temperature change $E_t[T_{t+1}]$ relative to pre-industrial levels crosses a threshold.[11] We make the threshold depend on expected temperature

[7] As in Crost and Traeger (2010), the model is actually solved using a transformation mapping the infinite time horizon to the unit interval. Compare also Kelly and Kolstad (1999).

[8] We originally have an instantaneous utility function $L_t u(C_t/L_t)$ as in DICE. For convenience of representation and of numerical solution, we divide through with the population $L_t$ and represent welfare as a function of per capita consumption $c_t = C_t/L_t$. The discount factor then has to pick up the exogenous change in population on top of pure time preference.

[9] Available at http://data.giss.nasa.gov/gistemp/.

[10] We implement the continuous distribution numerically using a Gauss-Legendre quadrature rule with 8 nodes.

[11] The effect of making a tipping point lag the crossing of a temperature threshold would depend on when the

rather than on actual temperature realizations because, first, expected temperature captures the intuition that medium-term changes in temperature are more likely to trigger tipping points and, second, it saves a state variable.[12] When the threshold temperature is known to be $T^*$, we call it a "certain threshold". In general, however, the decision-maker does not know where the threshold lies, making the threshold temperature a random variable $\tilde{T}$. The probability distribution for $\tilde{T}$ follows from assuming that a tipping point is sure to occur by the time the world reaches a temperature $\bar{T}$ and that any temperature between present levels and $\bar{T}$ has an equal chance of being the threshold. In most runs, we assume that the time 0 (i.e., year 2005) expected value for the threshold is 2.5°C: $\mathrm{E}_0\ \tilde{T} = 2.5$°C, which combines with the assumed uniform distribution for $\tilde{T}$ to give $\bar{T} = 4.33$°C (Figure 2a). This value for $\mathrm{E}_0\ \tilde{T}$ is consistent with the political 2°C limits for avoiding dangerous anthropogenic interference. Further, in Smith et al. (2009), 2.5°C is in the upper end of the temperature region that produces significant risk of large-scale discontinuities and is just below the temperatures that produce severe risk. In the baseline optimal policy scenario without a threshold, temperature rises above 2.5°C in the year 2098 and never reaches 4.33°C. We vary $\bar{T}$ between 3°C and 5°C to assess the sensitivity of the results to the upper bound of the uniform distribution. The probability of crossing the threshold between periods $t$ and $t+1$ conditional on not having crossed the threshold by time $t$ is:

$$h(E_t[T_{t+1}]) = \max\left\{0, \frac{\min\{E_t[T_{t+1}], \bar{T}\} - E_{t-1}[T_t]}{\bar{T} - E_{t-1}[T_t]}\right\} . \tag{2}$$

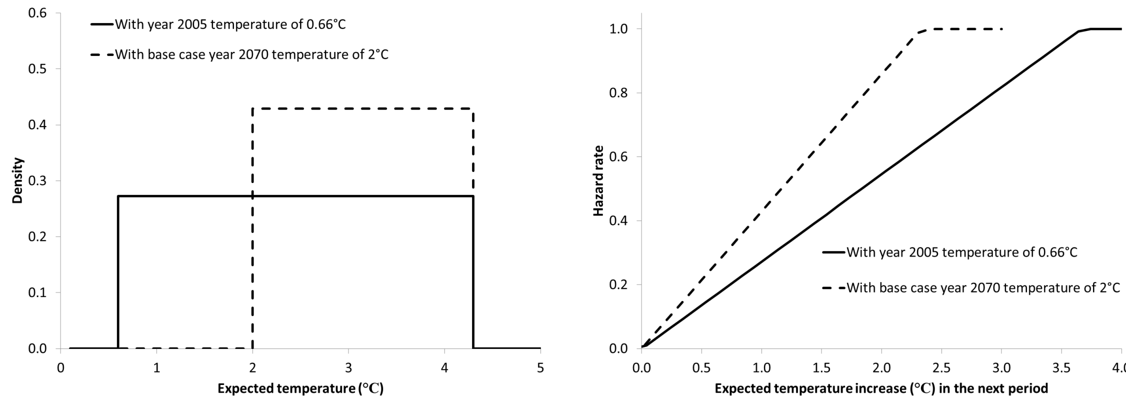
This expression gives the hazard rate for a contemplated temperature increase. The time $t$ expectation of temperature at time $t+1$ is conditional on the $CO_2$ stock and the transient feedback at time $t+1$ (see Appendix 1).[13] These are known at time $t$, and expectations are formed over the noise term generating stochastic temperatures. The hazard rate for further temperature increases is greater when the current temperature is greater (Figure 2b). However, the hazard rate as a function of changes in $CO_2$ does not always rise with the current $CO_2$ concentration. While greater current $CO_2$ corresponds to greater current temperature (raising the hazard rate), additional $CO_2$ has less of an effect on temperature when the $CO_2$ concentration is already high (decreasing the hazard rate). A period's hazard rate is endogenously determined by the decisions made given the realized state variables.

Expected welfare in the pre-threshold regime is the sum of the known current welfare and the discounted uncertain future welfare. If the threshold is not crossed during the next period, welfare

policymaker learns the threshold was crossed and on whether the tipping point's effects are irreversible once the threshold is crossed.

[12] In DICE-2007, the $CO_2$ stock increases monotonically until the model reaches a sufficiently high level of abatement. From this point on, the decay rate of $CO_2$ outweighs the flow of emissions, making the $CO_2$ stock decrease monotonically. With an increasing $CO_2$ stock, the probability of crossing the threshold is proportional to the difference between expected temperature in the next period and the expected temperature for the current period (as determined by the current $CO_2$ stock). For a decreasing $CO_2$ stock, the probability of crossing the threshold is 0. As long as expected temperature is a quasiconcave function of time, we do not need an additional state variable to keep track of the highest historic expected temperature. As in DICE-2007, $CO_2$ concentrations in our model follow a path that is quasiconcave in time, with the policymaker selecting the $CO_2$ path so that expected temperature is also quasiconcave in time.

[13] Our uncertain threshold is therefore distributed in the two-dimensional state space of time and the $CO_2$ stock (compare Nævdal, 2006; Nævdal and Oppenheimer, 2007).




(a) Probability distribution for $\tilde{T}$ (b) Hazard rate

Figure 2: As the time $t$ expected temperature increases without crossing a threshold, the probability distribution for the threshold level $\tilde{T}$ places more mass on temperatures yet to be reached. Each additional increase in temperature therefore also produces a greater risk of crossing the threshold.

is given again by the pre-threshold welfare function $V_{\psi=0}$ (evaluated at the next period's states and subject to the temperature uncertainty captured by $\mathbb{P}$). However, with a probability given by the hazard rate, next period's welfare is determined by the post-threshold regime in the function $V_{\psi=1}$. Therefore, expected welfare in the next period is the weighted average of welfare with and without crossing the threshold, where the weights follow from the hazard rate. This gives the pre-threshold dynamic programming equation:

$$\begin{aligned} V_{\psi=0}(k_t, M_t, t) &= \max_{c_t,\mu_t} u(c_t) + \beta_t \int f_{amb}^{-1}\bigg[[1 - h(E_t[T_{t+1}])] f_{amb}[V_{\psi=0}(k_{t+1}, M_{t+1}, t+1)] \\ &\qquad + h(E_t[T_{t+1}]) f_{amb}[V_{\psi=1}(k_{t+1}, M_{t+1}, t+1)]\bigg] d\mathbb{P} \\ &= \max_{c_t,\mu_t} \frac{c_t^{1-\eta}}{1-\eta} + \frac{\beta_t}{1-\eta} \int \bigg[[1 - h(E_t[T_{t+1}])] \big[(1-\eta) V_{\psi=0}(k_{t+1}, M_{t+1}, t+1)\big]^{\frac{1-\gamma}{1-\eta}} \\ &\qquad + h(E_t[T_{t+1}]) \big[(1-\eta) V_{\psi=1}(k_{t+1}, M_{t+1}, t+1)\big]^{\frac{1-\gamma}{1-\eta}}\bigg]^{\frac{1-\eta}{1-\gamma}} d\mathbb{P} \ . \end{aligned} \tag{3}$$

The function $f_{amb}$ captures smooth ambiguity aversion (Klibanoff et al., 2005, 2009), or intertemporal aversion to subjective uncertainty (Traeger, 2010). In our case, subjective uncertainty characterizes the chance that a tipping point is crossed. We adopt a power function $f_{amb}(V) = ((1-\eta)V)^{\frac{1-\gamma}{1-\eta}}$. In a one-commodity setting, $\gamma$ can be understood as a measure of Arrow-Pratt relative risk aversion with respect to subjective uncertainty (or to poorly understood uncertainty). As described earlier, the probability of crossing the threshold before the next period is generally less confidently known than is the distribution for temperature in a given period. For this reason, we have $\gamma \geq \eta$, allowing the decision-maker to be more averse to the chance of a threshold crossing (determined by $h(\cdot)$)

Table 2: The welfare specifications used to assess the effect of ambiguity aversion when tipping point probabilities are considered more subjective. All specifications employ an aversion to intertemporal substitution and objective risk of $\eta = 2$ as in Nordhaus (2008).

| Preference specification | Aversion to tipping point uncertainty ($\gamma$) |
|---|---|
| Ambiguity neutrality | 2 |
| Moderate ambiguity aversion | 9.5 |
| Strong ambiguity aversion | 50 |

than to the risk produced by not knowing the next period's temperature exactly (determined by $\mathbb{P}$). We solve for the pre-threshold value function employing the same method as described above for the post-threshold value function.

When $\eta = \gamma$, the policymaker is ambiguity-neutral and the welfare evaluation is as in DICE (Table 2). Runs with this parameterization do not disentangle preferences over the tipping point lottery from preferences over time and over the temperature change lottery. To assess the effect of ambiguity aversion, we consider a case with moderate ambiguity aversion close to the calibration result of Ju and Miao (2009) in the asset pricing context ($\gamma = 9.5$) and also a case with stronger ambiguity aversion ($\gamma = 50$).

# 3 Modeled tipping points

We now further describe the four tipping points listed in Table 1 and illustrated in Figure 1. Appendix 2 shows how optimal policy responds to the modeled changes in system dynamics. It helps to first introduce the equations describing equilibrium temperature change, damages, and the evolution of the $CO_2$ stock. Equilibrium temperature change $T_{equil}$ relative to pre-industrial levels is given by:

$$T_{equil} = s\frac{\ln(M_t/M_{pre}) + EF_t/5.35}{\ln 2} \ . \tag{4}$$

The parameter $s$ is climate sensitivity, $M_t$ and $M_{pre}$ are the time $t$ and pre-industrial $CO_2$ stocks, and $EF_t$ is the exogenous time $t$ non-$CO_2$ forcing. To obtain expected time $t$ temperature change $E_{t-1}[T_t]$, we adjust $T_{equil}$ for the cooling reservoir described in Appendix 1. Let $Y_{gross}$ be the total output produced by the time $t$ capital stock. Then the output available for allocation to consumption, abatement, and investment is $Y_{gross}/(1 + D)$. The function $D$ gives damages due to temperature change:

$$D = b_2 T_t^{b_3} \ . \tag{5}$$

The parameter $b_2$ equals 0.0028388 as in DICE-2007. The damage function in DICE is quadratic, giving $b_3 = 2$. Finally, the $CO_2$ stock $M_t$ evolves according to the following transition equation:

$$M_{t+1} = M_{pre} + \big(1 - \phi(M,t)\big)\big(M_t - M_{pre}\big) + e_t \ , \tag{6}$$

where $e_t$ gives time $t$ emissions. The function $\phi(M,t)$ is defined in Appendix 1. It determines the decay of the $CO_2$ stock towards pre-industrial levels. This decay is a function of time and the $CO_2$

stock and is based on values inferred from the output of DICE-2007.

The first tipping point occurs if we observe land ice sheets to be retreating over shorter-than-expected timescales in response to an experienced increase in temperature. Common definitions of climate sensitivity assume that these "slow feedbacks" do not affect temperature on relevant timescales. If they do begin to operate, they amplify the warming predicted using conventional estimates of climate sensitivity. This tipping point shifts our model into a post-threshold regime with increased climate sensitivity $\hat{s}$:

$$\hat{s} = 2s \; . \tag{7}$$

Climate sensitivity in the pre-threshold regime is 3°C as in DICE, and post-threshold climate sensitivity becomes 6°C (Hansen et al., 2008). Each unit of emissions causes more temperature change than previously expected, which quadruples damages from equilibrium temperature change.

The second tipping point occurs when a sudden, irreversible change such as the collapse of the West Antarctic or Greenland ice sheets increases impact assessments for higher temperatures (Oppenheimer, 1998; Vaughan, 2008; Notz, 2009). DICE uses a damage function that is quadratic in temperature, but many have raised concerns about the fit of a quadratic function at high levels of temperature change (e.g., Wright and Erickson, 2003; Ackerman et al., 2009; Newbold and Daigneault, 2009; Weitzman, 2009; Ackerman et al., 2010; Hanemann et al., 2010). This tipping point increases the convexity of the damage function by changing the exponent on $T_t$ to $\hat{b}_3$. We parameterize this as making damages a cubic function of temperature in the post-threshold regime:[14]

$$\hat{b}_3 = b_3 + 1 = 3 \; . \tag{8}$$

Each unit of temperature change reduces output by more than it would have prior to crossing the threshold. Damages are now multiplied by an additional factor equal to current temperature.

The third and fourth tipping points model degradation of carbon sinks and and activation of methane sources. The third tipping point reflects the possibility that carbon sinks weaken beyond the predictions of coupled climate-carbon cycle models (Raupach et al., 2008). The many processes through which the climate and carbon cycle affect each other are difficult to model and to calibrate (Luo, 2007), making it hard to rule out extreme outcomes (e.g., Sitch et al., 2008). Warming-induced changes in oceans (Le Quéré et al., 2007), soil carbon dynamics (Eglin et al., 2010), and standing biomass (Huntingford et al., 2008) could affect the uptake of $CO_2$ from the atmosphere. In order to represent an extreme form of weakened sinks, we parameterize this tipping point as causing a 75% reduction in the decay rate of $CO_2$:

$$\hat{\phi} = \phi/4 \; . \tag{9}$$

The change from $\phi$ to $\hat{\phi}$ increases the time for which a unit of emitted $CO_2$ affects atmospheric $CO_2$.

The fourth tipping point corresponds to a permanent increase in forcing from non-$CO_2$ greenhouse gases, as if from a large, sustained release of methane from melting permafrost or subsea

[14] This regime with more convex damages is similar to the case considered by Azar and Lindgren (2003), but their regime switch can happen only in 2035 and with a low probability that is exogenous (i.e., does not depend on emission decisions).

Table 3: The model runs used to assess the effects of climate tipping points and additional aversion to tipping point uncertainty.

| Model | Description |
|---|---|
| Base case | No thresholds can occur |
| Certain threshold | Policy precisely controls when and whether a threshold occurs |
| Uncertain threshold | Policy controls the probability that a threshold occurs |
| Ambiguity aversion | Additional aversion to threshold uncertainty |

hydrates (Hall and Behl, 2006; Archer, 2007; Schaefer et al., 2011). A large methane release is one of the hypothesized triggers for ancient periods of rapid warming (Zachos et al., 2008). We represent this tipping point as permanently increasing each period's non-$CO_2$ forcing by 1.5 W $m^{-2}$:

$$\widehat{EF}_t = EF_t + 1.5 \; . \tag{10}$$

The 1.5 W $m^{-2}$ additional forcing is at the low end of a range of plausible methane emission rates during ancient warming (Schmidt and Shindell, 2003), and it is equivalent to increasing the $CO_2$ concentration by 30%. The methane event corresponding to the permanent increase in forcing is a large initial release followed by an ongoing release of declining magnitude. The additional forcing from the initial release tends to decrease as methane decays to $CO_2$, but the ongoing release offsets this decay.

## 4 Results

We compare several sets of model runs to assess how the social cost of carbon and optimal $CO_2$ concentrations respond to different tipping points, to uncertainty about the temperature threshold for a tipping point, and to additional aversion to tipping point uncertainty (Table 3). The baseline version of the model is the standard DICE model plus temperature stochasticity. Period-to-period temperature uncertainty has a negligible effect on policy in these runs. Temperature variability affects residual output (i.e., investment), but the independently distributed fluctuations are not important enough to affect the consumption and abatement policies chosen before a period's temperature variability is resolved. A second set of runs has a tipping point occurring at a known threshold. The decision-maker knows that the world will change once reaching the temperature $T^*$. She can therefore adjust emissions to delay or avoid crossing the threshold and also adjusts optimally after crossing the threshold. A third set of runs makes the decision-maker uncertain about the temperature threshold that triggers the tipping point. Additional emissions increase the chance of crossing into the new regime, and the decision-maker updates beliefs about the threshold based on whether a tipping point has occurred. Finally, a fourth set of runs includes aversion to tipping point ambiguity, which makes the decision-maker more averse to tipping point uncertainty than to temperature change uncertainty. In each model run with tipping points, the decision-maker only faces one type of tipping point and knows which type that is.

The effect of each tipping point's occurrence on optimal policy and welfare determines how pre-threshold policy responds to awareness of the tipping point. Pre-threshold policy affects the chance of crossing the tipping point and also affects the capital and $CO_2$ stocks at the time the threshold is crossed. Figure 4 depicts the optimal time paths of the social cost of carbon (optimal carbon tax) and $CO_2$ concentrations conditional on not having crossed the threshold. The depicted paths have each year's temperature take its expected value. We show optimal policy conditional on not having crossed the threshold because we want to inform the choice of policy when tipping points still might occur. Each graph compares the baseline scenario without tipping point awareness to runs with tipping point awareness and with moderate and strong ambiguity aversion. Comparing time paths in graphs on different rows shows the effects of different types of tipping points. A possible tipping point can increase the year 2015 social cost of carbon by 50% (Table 4), and optimal policy may still incur a substantial (though reduced) chance of crossing the uncertain threshold (Table 5). The effects of tipping point considerations on the near-term social cost of carbon and on the peak $CO_2$ concentration are not highly sensitive to the upper bound used in the uniform distribution for the threshold's location (Figure 5). If the threshold is known with certainty to be at 2.5°C or higher, the policymaker will ensure that it is not crossed, but the abatement required to avoid lower thresholds may be overly costly (Figure 6).

In order to understand the effects of tipping points on optimal policy, we analyze the effect of an additional unit of emissions on a policy program's value. In the baseline scenario without tipping points, the loss from an additional unit of emissions is given by the derivative of the pre-threshold value function with respect to emissions: $d\,V_{\psi=0}(k_t^*, M_t^*, t)/d\,M_t$. However, when tipping points are possible, this derivative only captures the change in value conditional on staying in the pre-threshold regime. The social cost of carbon also depends on the change in value conditional on crossing into the post-threshold regime and on the change in the probability of crossing into the post-threshold regime. To more readily analyze the right-hand side of equation (3) at the optimum, we suppress all arguments that do not change for an additional unit of emissions $e_t$. Then the value of the program is:

$$u^* + \beta_t \int f_{amb}^{-1}\Bigg[\underbrace{\big[1 - h(e_t^*)\big] f_{amb}\big[V_{\psi=0}(M_{t+1}(e_t^*))\big] + h(e_t^*) f_{amb}\big[V_{\psi=1}(M_{t+1}(e_t^*))\big]}_{V_{eff}(e_t^*)}\Bigg] d\mathbb{P} \ ,$$

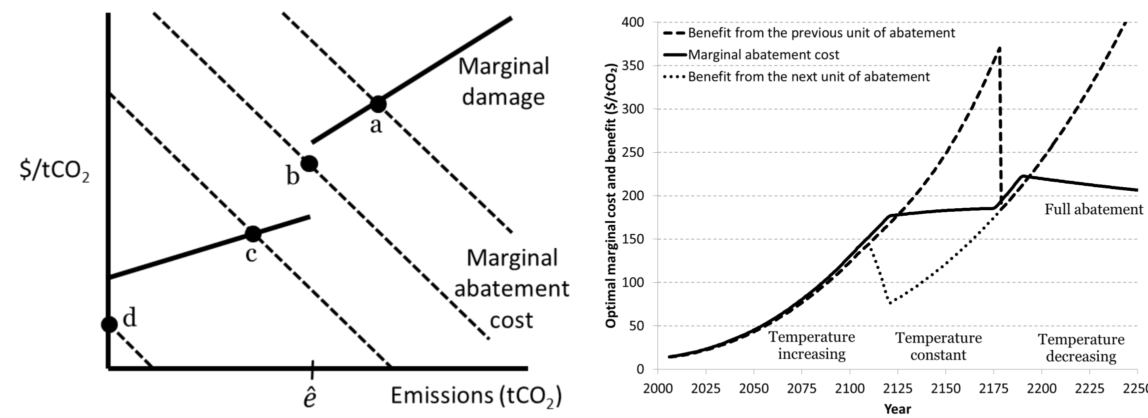
where $h(e_t^*) = h(E_t[T_{t+1}(e_t^*)])$. The change in value from an additional unit of emissions is then given by:

$$\begin{aligned}\beta_t \int \frac{d\,f_{amb}^{-1}\big[V_{eff}(e_t^*)\big]}{d\,V_{eff}} \Bigg\{ &\big[1 - h(e_t^*)\big] \frac{d\,f_{amb}\big[V_{\psi=0}(M_{t+1}(e_t^*))\big]}{d\,V_{\psi=0}} \frac{\partial V_{\psi=0}(M_{t+1}(e_t^*))}{\partial M_{t+1}} \\ &+ h(e_t^*) \frac{d\,f_{amb}\big[V_{\psi=1}(M_{t+1}(e_t^*))\big]}{d\,V_{\psi=1}} \frac{\partial V_{\psi=1}(M_{t+1}(e_t^*))}{\partial M_{t+1}} \\ &+ \underbrace{\frac{d\,h(E_t[T_{t+1}(e_t^*)])}{d\,E_t[T_{t+1}]}}_{\text{i}} \underbrace{\frac{d\,E_t[T_{t+1}(e_t^*)]}{d\,e_t}}_{\text{ii}} \underbrace{\Big(f_{amb}[V_{\psi=1}(M_{t+1}(e_t^*))] - f_{amb}[V_{\psi=0}(M_{t+1}(e_t^*))]\Big)}_{\text{iii}} \Bigg\} d\mathbb{P} \ .\end{aligned} \tag{11}$$

When temperature is decreasing, there is no chance of crossing a threshold in the next period ($h(e_t^*) = 0 = d\,h(e_t^*)/d\,e_t$). In this case, only the first line contributes to the damage from additional emissions, and as in the baseline scenario without tipping points, the social cost of carbon is determined by the effect of an increased carbon stock on the next period's pre-threshold value function. When there is positive probability of crossing a threshold, the decision-maker also considers the effect of an increased carbon stock on the next period's post-threshold value function (via the second line in the equation above). This additional consideration does not greatly affect the social cost of carbon in our simulations because the difference between the value function derivatives in the first and second lines is not large enough to matter after weighting by the small annual hazard rate. However, the social cost of carbon is not simply a convex combination of the effects of a higher $CO_2$ stock in the pre- and post-threshold regimes. Instead, additional emissions also affect the probability of crossing the threshold as in the third line above. Additional emissions increase the temperature expected in the next period (term ii), and if this expected temperature is above any historical level, then the additional emissions also raise the hazard of crossing into the post-threshold regime (term i). The cost of the regime switch is proportional to the difference between the pre- and post-threshold value functions (term iii). If, on the other hand, the expected temperature is not greater than historic levels, then the third line does not contribute to the social cost of carbon because the hazard rate will be zero with or without the additional emissions.

The contribution of the third line to the social cost of carbon (the "hazard effect") is crucial to understanding the influence of possible tipping points on optimal policy. The value lost from switching regimes (term iii) is often large. As long as optimal temperatures are increasing (so term i is nonzero), the hazard effect increases the social cost of carbon relative to a scenario without tipping points. As $CO_2$ concentrations increase, term i grows bigger while term ii grows smaller. The change in term i usually dominates in our model, which makes the hazard effect become more important as $CO_2$ increases. The effect of tipping points on the social cost of carbon therefore tends to increase over time. Eventually, abatement becomes cheap enough that it is no longer optimal for temperature to increase. Figure 3a describes the relation between the marginal benefit of abatement and the marginal cost of abatement. When marginal abatement cost is high, optimal emissions are greater than the level $\hat{e}$ that would stabilize temperature (point a). There is a discontinuity in the marginal benefit of abatement at $\hat{e}$, as the hazard effect is positive for greater emissions and zero for lower emissions. When marginal abatement cost becomes low enough to cross marginal benefit at this discontinuity (point b), it is optimal to stabilize temperature and thereby eliminate the contribution from the second and third lines in equation (11). The benefit from the next unit of abatement would be less than the cost of the additional abatement because there is no hazard effect for further emission reductions, but the benefit from the last unit of abatement was greater than the cost of that abatement because additional emissions would indeed produce a nonzero hazard rate. Abatement is set so as to keep temperature constant, making $CO_2$ concentrations nearly level out (see Figure 4).[15] When temperature is held constant, the marginal benefit of abatement differs between the last unit emitted—to which only the first line in equation (11) contributes—and the next unit emitted—to which all three lines contribute. Figure 3b shows the social cost of carbon

[15] Because the $CO_2$ decay rate, the ocean cooling adjustment, and the non-$CO_2$ forcing change over time, keeping temperature constant implies slightly decreasing the $CO_2$ stock over time.




(a) (b)

Figure 3: On the left, marginal damage (i.e., the marginal benefit of abatement) exhibits a discontinuity at the net emission level $\hat{e}$ that keeps temperature constant. Marginal abatement cost (dashed lines) falls over time. On the right, the benefit from the next unit of abatement (dotted line) is less than the benefit from the previous unit of abatement (dashed line) when optimal policy holds temperature constant. The plotted values are results for an unknown temperature threshold that would increase climate sensitivity.

for the last unit emitted as a dotted line and the social cost of carbon for the next unit emitted as a dashed line. The solid line depicts optimal marginal abatement cost. These three lines diverge once it is optimal to hold temperature constant, as the calculation of marginal benefit depends on whether emissions have a nonzero hazard rate and marginal abatement cost depends on the emissions needed to hold temperature constant. Once marginal abatement cost falls far enough, the decision-maker reduces optimal emissions below $\hat{e}$ (point c), making temperature fall. Finally, once the economy reaches full abatement, the marginal benefit of abatement and the marginal cost of abatement diverge once more (point d). Because the DICE model does not allow negative emissions, the decision-maker uses a boundary solution in abatement even though the benefit of the last unit of abatement exceeds its cost.

We see the effects of each tipping point possibility in Figure 4. All increase the social cost of carbon and decrease optimal $CO_2$. The climate sensitivity tipping point has the greatest effect on optimal pre-threshold policy, increasing the near-term cost of carbon by 50% and decreasing peak $CO_2$ from around 680 ppm to below 530 ppm. The tipping points with the next strongest effects on pre-threshold policy are the damage convexity and non-$CO_2$ forcing tipping points, which increase the near-term cost of carbon by about 40% and decrease peak $CO_2$ to around 540 ppm and 555 ppm, respectively. The decay rate tipping point has the least effect on optimal pre-threshold policy, increasing the near-term cost of carbon by only 12% and decreasing the pre-threshold peak $CO_2$ concentration to around 625 ppm. The effect of each tipping point on the social cost of carbon grows with time as the effect of emissions on the hazard rate increases with time (due to the increasing value of term i from equation (11) at higher $CO_2$ concentrations). Some tipping

Table 4: The social cost of carbon (\$/t$CO_2$, current value) in 2015 under expected draws. It is calculated from marginal abatement cost and is \$12/t$CO_2$ in the baseline case.

| | Tipping point | | | |
|---|---|---|---|---|
| | Climate sensitivity increased | Damage convexity increased | $CO_2$ sinks weakened | Non-$CO_2$ forcing increased |
| Uncertain threshold | 18 | 16 | 13 | 16 |
| Ambiguity aversion | 18 | 16 | 13 | 16 |
| Strong ambiguity aversion | 19 | 17 | 13 | 17 |

points affect policy more than others because the total loss $V_{\psi=0} - V_{\psi=1}$ from crossing a threshold (term iii) depends on the effect of the tipping point on system dynamics. The climate sensitivity tipping point quadruples damages for the equilibrium temperature, though its effect is smaller for the transient temperature modeled. The damage convexity tipping point multiplies pre-threshold damages by a factor equal to the temperature outcome, which means damages are usually not even tripled. The non-$CO_2$ forcing tipping point also has a smaller effect on damages than does the climate sensitivity tipping point, but its exact relation to the damage convexity tipping point depends on the current temperature and $CO_2$ stock. However, the effects of these two regime shifts on damages are often comparable. Finally, the decay rate tipping point has the smallest effect on optimal policy because the reduction in value that it causes is undercut both by the time it takes $CO_2$ to accumulate in the atmosphere and by the concavity of the relationship between temperature and $CO_2$.

Introducing tipping points requires considering the effect of emissions on the post-threshold continuation value and on the expected loss from crossing a threshold; in contrast, introducing ambiguity aversion (i.e., making $f_{amb}$ concave rather than linear) does not introduce a new term but changes the evaluation of each of these other terms. The most significant effect is on the loss from crossing a threshold (term iii in equation (11)), as increasing ambiguity aversion raises the perceived distance between the two continuation values. The importance for optimal policy of the total loss from crossing a threshold depends on the effect of additional emissions on the hazard rate. As discussed above, the effect of emissions on the hazard rate generally increases along the optimal path as the change in term i dominates the change in term ii for high $CO_2$ stocks. Ambiguity aversion increases the effect of tipping points on the cost of carbon through term iii, and ambiguity aversion itself becomes more important for policy over time as tipping points themselves become more important for policy through term i. When we implement ambiguity aversion in our recursive version of DICE, we find that the near-term effect on the social cost of carbon is usually minor even for strong ambiguity aversion (Table 4). Further, the optimal policy for a moderately ambiguity-averse decision-maker gives about the same chance of eventually crossing a threshold as does the optimal policy for an ambiguity-neutral decision-maker. Even a strongly ambiguity-averse decision-maker reduces the probability of a tipping point occurring by fewer than 6 percentage points relative to an ambiguity-neutral decision-maker (Table 5).

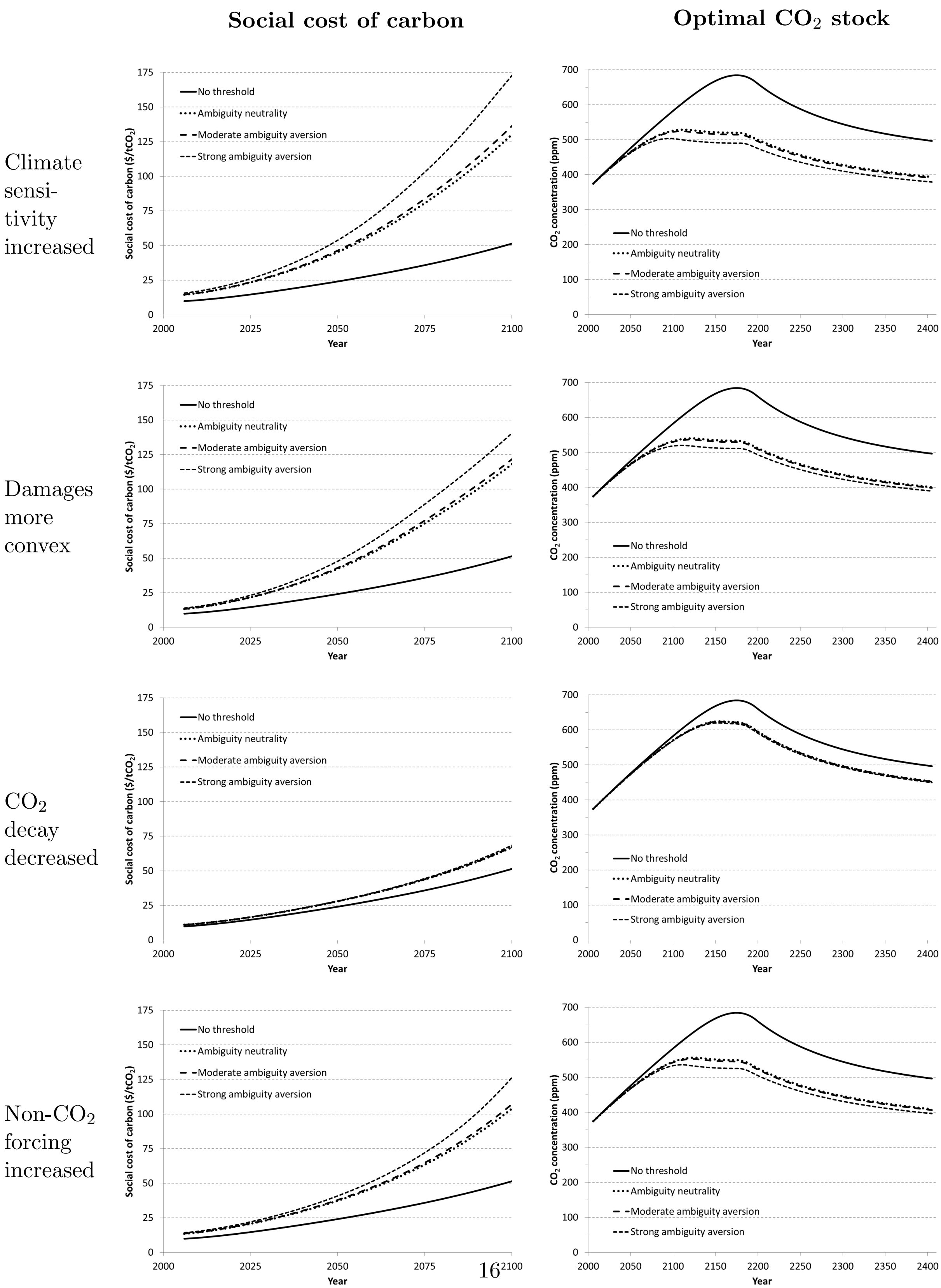


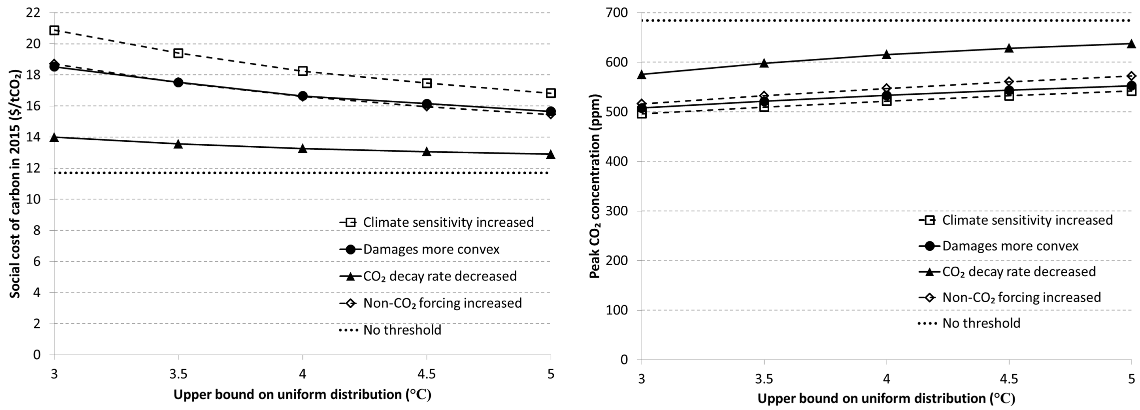



Figure 4: Time paths for the social cost of carbon (current value) and the $CO_2$ stock under each type of tipping point and using expected draws. With an uncertain threshold, we simulate a path that happens to never cross a threshold in order to see how the modeled policymaker adjusts to the possibility over time.

Table 5: The probability that the threshold will ever be crossed. This is determined by the peak temperature in a simulation with expected draws and a threshold that happens to never be crossed. In parentheses, the expected year of crossing, conditional on the threshold being crossed at some point and determined by repeated simulations. This expected crossing will tend to occur earlier when policy makes temperature peak earlier.

| | Tipping point | | | |
|---|---|---|---|---|
| | Climate sensitivity increased | Damage convexity increased | $CO_2$ sinks weakened | Non-$CO_2$ forcing increased |
| Uncertain threshold | 0.46<br>(2050) | 0.48<br>(2052) | 0.61<br>(2062) | 0.51<br>(2053) |
| Ambiguity aversion | 0.45<br>(2049) | 0.48<br>(2051) | 0.61<br>(2062) | 0.50<br>(2053) |
| Strong ambiguity aversion | 0.40<br>(2046) | 0.44<br>(2049) | 0.61<br>(2061) | 0.47<br>(2050) |

The baseline policy path has a 68% chance of ever crossing the unknown threshold with standard preferences. It crosses a 2.5°C threshold in 2098, and its expected year of crossing conditional on crossing at some point is 2068. The decision-maker follows the baseline path if unaware of tipping point possibilities.




(a) Social cost of carbon in 2015 (b) Peak $CO_2$ concentration

Figure 5: The sensitivity of the year 2015 social cost of carbon and of the peak $CO_2$ concentration to the upper bound $\bar{T}$ for the uniform distribution over the temperature threshold $\tilde{T}$. All of these runs use ambiguity-neutral preferences, and the simulations use expected draws for temperature while assuming the realized threshold is above the peak temperature reached.

Because the scientific literature does not support a particular distribution for the temperature threshold, it is important to assess the sensitivity of our conclusions to the bounds on the uniform distribution. We parameterized our base case distribution so that the year 2005 expected threshold was at 2.5°C, implying an upper bound of $\bar{T} = 4.33$°C. To assess sensitivity, we also vary the upper bound $\bar{T}$ between 3°C and 5°C.[16] The resulting optimal policies never lead temperature across $\bar{T}$. As expected, by reducing the threshold risk implied by a given temperature outcome, raising the threshold distribution's upper bound increases peak $CO_2$ and decreases the near-term social cost of carbon. Decreasing the upper bound to 3°C has a slightly greater effect on peak $CO_2$ and the near-term cost of carbon than does modeling strong ambiguity aversion. Our results regarding the relative importance of each type of tipping point and the influence of tipping points on optimal policy are therefore not highly sensitive to the precise form of the uniform distribution for the unknown threshold.

Finally, we consider the case where the temperature threshold's level is known with certainty. As described above, marginal damage depends on how the additional $CO_2$ stock affects each possible continuation value and on how additional emissions affect the probability of crossing the threshold before the next period. However, now the decision-maker knows—and controls—exactly which continuation value will hold in the next period. The decision to eventually cross the threshold or not depends on the cost of imposing a temperature constraint relative to the cost of switching regimes. If the threshold is sufficiently high, then the known tipping point is irrelevant because the temperature constraint would be slack. If the threshold is too close to the initial temperature, then the cost of avoiding it is extremely large and marginal damage is determined by the effect of additional $CO_2$ in the post-threshold regime. If the threshold is instead in an intermediate region,

[16]We are currently exploring sensitivity to greater upper bounds.

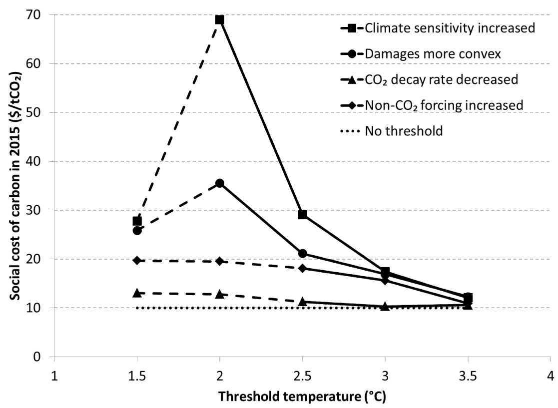



Figure 6: The social cost of carbon in 2015 (current value) when the temperature threshold is known to be at a certain level. Solid lines connect points for which the policymaker plans to avoid the threshold. All of these runs use ambiguity-neutral preferences. Note that these results are as yet only indicative and will change once we implement an improved solution method.

then optimal policy has temperature approach the threshold but never cross it. The social cost of carbon is then determined by the cost of adjusting future policy to still avoid the threshold. If future policy cannot be adjusted because additional emissions would trigger the tipping point, then these additional emissions bear the entire cost of the regime shift. This produces a discontinuity in marginal benefit as in Figure 3. For threshold levels that are worth avoiding, the cost of carbon decreases as the known threshold increases, but the cost of carbon can actually increase with a higher threshold if the higher threshold makes the decision-maker choose to avoid crossing it (Figure 6). The tipping points that impose greater damages are worth avoiding at lower thresholds. In these cases, optimal policy takes temperature up to the threshold and holds it constant until abatement becomes sufficiently cheap.

# 5 Conclusions

We have described an original extension of DICE to include the endogenous possibility of climatic tipping points, learning about the temperature threshold that triggers tipping points, and aversion to ambiguity about the threshold's location. Tipping points that increase the lifetime of $CO_2$ are less important than tipping points that increase the damage at a given time per unit of atmospheric $CO_2$. While previous work has shown that optimal policy is sensitive to the chosen value of climate sensitivity and to the convexity of the damage function, we have shown that it is also sensitive to the anticipation of endogenous, discontinuous changes in climate sensitivity or damage convexity. Ambiguity attitude does not become crucial to our evaluation of tipping point uncertainty until the probability of tipping points becomes much greater than in our model's early years. The type of tipping point faced is more important for near-term policy than is additional aversion to tipping point uncertainty or the precise form of the uniform distribution for the temperature threshold.

Our conclusions have implications for climate science, economic modeling, and climate policy.

First, it is important for IAMs that climate science improve knowledge about both the effects of tipping points on system dynamics and the types of temperature paths that trigger them (Alley et al., 2002, 2003). It is also important to translate tipping point results into the reduced climate models used by IAMs. Which variables might a tipping point affect? Is a tipping point triggered by medium-term average temperature, by short-term temperatures, by interannual variability, or by the rate of warming? What does the distribution for its occurrence look like? How might we expect to learn about tipping point risks? IAMs' conclusions might be sensitive to each of these answers and should be updated as the climate science literature progresses.

Second, economic modeling should gauge which simplifications are likely to be crucial for the results used in policy assessments. We have shown that optimal policy paths are sensitive to assumptions about damage convexity and climate sensitivity and to assumptions about the possibility of tipping points. Modeling exercises that do not vary key parameters or that assume smooth, predictable system dynamics need to be explicit about the omitted factors that tend to push their estimates in a given direction. This is especially important when all models tend to omit the same factors. In that case, the spread of models' estimates for the social cost of carbon should not directly give the distribution used in policy analysis. Past compilations of IAMs' estimates (e.g., Tol, 2008) described a set of models that omitted climatic features that we have shown could strongly affect the reported results. Further, by building uncertainty into the decision environment, we have shown that the information structure around tipping points has policy consequences. For instance, optimal policy is quite different if a threshold is known to be at 2.5°C versus if it is uncertain but currently expected to be at 2.5°C. As a result, it is important that other IAMs include uncertainty in a realistic way and vary the information structure. In some cases, certainty makes a variable less relevant, but in other cases, certainty means more effort will be expended to control a variable's effects. Because not much is known about tipping points or how to model them, they constitute an important form of model uncertainty that covers the effects of tipping points as well as knowledge about them.

Third, regarding climate policy, we find that including tipping points can increase DICE's estimate of the year 2015 social cost of carbon by 50% and can decrease DICE's year 2015 industrial $CO_2$ emissions by over 1 Gt $CO_2$. Using a recent bottom-up abatement cost curve, this increase in the social cost of carbon could increase the economical emission reductions in the U.S. by 0.25 Gt $CO_2$ per year (Creyts et al., 2007). Estimates of the social cost of carbon play an increasingly important role in the evaluation of government policies (e.g., Interagency Working Group on Social Cost of Carbon, 2010; Masur and Posner, 2010; Greenstone et al., 2011) and may well affect the carbon price eventually targeted by carbon taxes and cap-and-trade policies. The best estimate of the social cost of carbon probably does not treat tipping points as impossible, but it is not clear which type of model comes closest to representing the world we face. In any case, much work remains to make IAMs' representation of tipping points more realistic (Hall and Behl, 2006). The challenge when choosing values for the social cost of carbon is one also faced in, among others, the choice of climate projections for planning and policy evaluation (Knutti et al., 2010; Lemoine, 2010), of emission factors in transportation policy (Plevin et al., 2010), and of interest rates in monetary policy (Hansen and Sargent, 2001): policymakers must consider and combine the results not just of different models but also of different possible structures for a given model. The decision is further complicated when only a small set of model structures has been explored. It is therefore desirable that policy consider both how a given model's predictions change with its assumptions and how to adapt to the results of future modeling efforts. While we have shown how the possibility of a threshold affects the shadow value of $CO_2$ emissions in a standard IAM, climate policy constitutes buying insurance against not just the assumed possibility of a tipping point but also against the prospect that the available models do not adequately capture future changes in the climate and the economy.

## 6 Appendix 1: Model calibration

This appendix describes the feedback representation of temperature change, the decay of atmospheric $CO_2$ over time, and the model's calibration to DICE-2007. We simplify the carbon cycle and temperature change representations from DICE in order to include the tipping point possibilities that might produce a more realistic model.

DICE determines time $t$ surface temperature from the stock of $CO_2$, from temperature in the previous period, and from the difference in the previous period between temperature at the surface and in the deep ocean. We use a more parsimonious relationship to capture the influence of time $t$ $CO_2$ on time $t$ temperature, and we calibrate this representation so as to capture the marginal relationships important for economic evaluation. We model expected time $t$ temperature change $T_t$ relative to pre-industrial levels as determined by the $CO_2$ stock $M_t$ and by the net feedback $f_{atm} + f_t$ (Roe, 2009):

$$E_{t-1}[T_t] = \frac{\lambda[R(M_t) + EF_t]}{1 - (f_{atm} + f_t)} = \frac{\lambda[5.35 \ln(M_t/M_{pre}) + EF_t]}{1 - (f_{atm} + f_t)} \; . \tag{12}$$

The function $R(M_t) = 5.35 \ln(M_t/M_{pre})$ gives the additional radiative forcing in W m$^{-2}$ caused by changing $CO_2$ concentrations from the pre-industrial level $M_{pre}$ to the time $t$ level $M_t$ (Ramaswamy et al., 2001: Table 6.2), and $EF_t$ is the exogenous non-$CO_2$ forcing in W m$^{-2}$. The parameter $\lambda = 0.315$ °C (W m$^{-2}$)$^{-1}$ gives the reference system (black body) temperature change per unit of radiative forcing (Soden et al., 2008; Roe, 2009). The sum $f_{atm} + f_t$ is non-dimensional and must be less than 1. As described in the main text, realized temperature is given by $T_t = \epsilon_t E_{t-1}[T_t]$, with $\epsilon_t$ an independent, normally distributed shock with a mean of 1 and a variance calibrated to historical temperatures.

Feedbacks determine the change in temperature generated as the earth system responds to a unit of temperature change in the reference system. When feedbacks are positive, each non-dimensional feedback factor $f$ represents the portion of the total system's temperature change produced by its associated processes. When the $CO_2$ concentration increases, the atmosphere traps more outgoing radiation (given by $R(M_t)$) even as incoming radiation has not changed. The planet heats up to restore the balance between outgoing and incoming radiation at the top of the atmosphere. This effect is given by the constant $\lambda$. However, the increase in surface temperature causes changes in the earth system that in turn cause further changes in surface temperature. For instance, the warmer atmosphere now holds more water vapor, which traps additional outgoing radiation and causes the surface to warm faster. This amplifying response is captured by a positive feedback factor. We assume that feedbacks are linear functions of temperature and affect each other only through temperature. This common assumption allows us to aggregate the "atmospheric" feedbacks of sea ice, clouds, and water vapor-lapse rate in the constant $f_{atm}$ (Soden et al., 2008; Lemoine, 2010). Climate sensitivity $s$ is the equilibrium temperature change from doubled $CO_2$ concentrations:

$$s = \frac{5.35\lambda \ln 2}{1 - f_{atm}} , \tag{13}$$

where being in equilibrium means $f_t = 0$. DICE-2007 uses a climate sensitivity of 3°C, which implies $f_{atm} \approx 0.61$.

The time-varying feedback factor $f_t$ represents transient feedbacks. It adjusts equilibrium feedback strength to give time $t$ temperature. DICE-2007 has one state variable for surface temperature and another for deep ocean temperature. The interaction between the two allows the ocean's heat capacity to moderate each period's temperature change. Further, DICE-2007 also delays the effect of radiative forcing on temperature. We use $f_t$ as a reduced-form version of the difference between time $t$ temperature and equilibrium temperature for a given $CO_2$ concentration (Baker and Roe, 2009). When $CO_2$ concentrations are increasing, $f_t$ should always be a negative feedback because ocean heat uptake prevents all of the equilibrium surface warming from occurring immediately. When the $CO_2$ stock is constant, $f_t$ should weaken (i.e., move towards 0) with time as the ocean and atmosphere equilibrate. When $CO_2$ concentrations are decreasing, $f_t$ can be positive as the ocean transfers stored heat to the atmosphere.

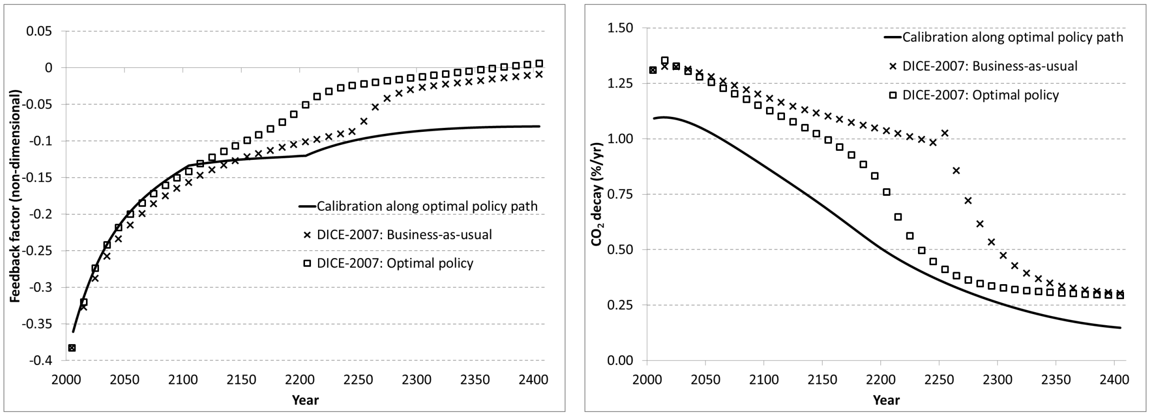
We also simplify the representation from DICE-2007 of the evolution of atmospheric $CO_2$. DICE-2007 has one state variable for atmospheric carbon, another for shallow ocean carbon, and a third for deep ocean carbon. These state variables and their associated transition matrix constitute a simple carbon cycle model determining how atmospheric $CO_2$ changes from period to period. We instead have an explicit decay rate $\phi(M, t)$ that is a function of time $t$ and of the atmospheric $CO_2$ stock $M_t$. As with the transient feedback factor $f_t$, including the atmospheric $CO_2$ stock as an argument proxies for the emission path up to time $t$. The decay rate $\phi$ implied by a change in DICE carbon stocks from $M_t$ to $M_{t+10}$ (where $t$ is in years) solves the following equation:

$$M_{t+10} = M_{pre} + (1 - \phi)^{10}(M_t - M_{pre}) + e_t \frac{1 - (1 - \phi)^{10}}{\delta} , \tag{14}$$

where $e_t$ gives annual $CO_2$ emissions over the decade and $M_{pre}$ is the pre-industrial $CO_2$ stock.

We calculate $f_t$ and $\phi$ as functions of time and the $CO_2$ stock. We use four runs from DICE to obtain temperature and carbon time series for calibrating these two functions: a business-as-usual run, a run with optimal policy, a lower-carbon run using an abatement path like that produced in our model with an uncertain climate sensitivity tipping point, and a still-lower-carbon run with an abatement path like that in our model when the climate sensitivity tipping point occurs in 2005. We infer the values of $f_t$ and $\phi$ at each 10-year timestep in these DICE runs, and we approximate each function over our state space using Chebychev polynomials. Figure 7 plots the resulting functions $\bar{f}_t$ and $\bar{\phi}$ along the business-as-usual and optimal policy paths from DICE, and it also plots each variable along our model's baseline path, after further adjusting them as described below. The inferred values of $f_t$ are at least broadly similar to the results of Baker and Roe (2009).

Inferring the functions for $f_t$ and $\phi$ enables us to replicate the $CO_2$-temperature relationship from DICE-2007 as well as the relation between $CO_2$ levels with decadal timesteps. However, these




(a) Transient feedback (b) $CO_2$ decay

Figure 7: The transient feedback $f_t$ and $CO_2$ decay rate $\phi$ along our model's baseline optimized path (after adjustments described in the text), along the optimized policy path in DICE-2007, and along the business-as-usual path in DICE-2007.

inferred functions do not reproduce the optimal $CO_2$ stock or social cost of carbon from DICE. Because we are primarily interested in how these policy-relevant values change under different specifications, we adjust these inferred functions to better match DICE-2007's optimized output (Figure 8). First, our use of an annual timestep instead of a decadal timestep tends to make the inferred $CO_2$ decay rate different from what we would need to replicate $CO_2$ dynamics with emissions varying each year. In particular, emissions in the second century often decline over a decade while DICE's stock transition equations treat annual emissions at the start of the decade as lasting for all ten years. We therefore adjust the inferred $CO_2$ decay rate $\bar{\phi}(M,t)$ as follows:

$$\phi(M,t) = 0.8\bar{\phi}(M,t) \ . \tag{15}$$

Second, while the transient feedback in the physical calibration reproduces the relationship between $CO_2$ stocks and temperature over representative time paths, it alters the marginal relationship between $CO_2$ and temperature by failing to delay the effect of radiative forcing on temperature. This would lead to a greater social cost of carbon and thus to greater abatement, which overly reduces both the $CO_2$ stock and temperature (Figure 8).[17] To better capture the social cost of carbon's time path, our calibration adjusts the inferred transient feedback $\bar{f}_t(M,t)$ as follows:

$$f_t(M,t) = \bar{f}_t(M,t) + \max\{-0.07, \ -0.0007\max(t-100,0)\} \ , \tag{16}$$

where $t = 0$ corresponds to 2005. This adjustment makes the transient feedback more negative after the year 2105. As Figure 8 shows, our calibration largely reproduces the time paths for the social cost of carbon, abatement, and the $CO_2$ stock, but this comes at the cost of making temperature 0.35°C too low near its its peak. Our calibration adjusts the inferred parameters to better represent

[17]We are currently shifting to a four-state model that will avoid this problem.

the marginal effect of $CO_2$ emissions in DICE, but in does so by reducing the effect of the total $CO_2$ stock on temperature and thereby increases available economic output relative to DICE.

# 7 Appendix 2: Model specification

This appendix provides the transition equations for state variables and exogenous variables (see Nordhaus, 2008; Crost and Traeger, 2010). It also describes the four modeled tipping points in terms of these equations and illustrates how post-threshold policy differs from policy in the absence of tipping points.

The transition equations for the state variables of effective capital ($k_t$) and atmospheric $CO_2$ ($M_t$) are:

$$k_{t+1} = e^{-(g_{L,t}+g_{A,t})}\left[(1-\delta_k)k_t + (1-\Psi_t\mu_t^{a_2})\frac{Y_{gross}}{1+D} - c_t\right] \qquad \text{(Capital)}$$

$$M_{t+1} = M_{pre} + (1-\phi(M,t))(M_t - M_{pre}) + \sigma_t(1-\mu_t)Y_{gross} + B_t\ . \qquad (CO_2)$$

In the transition equation for $CO_2$, $\sigma_t$ is the emission intensity of gross output and $B_t$ gives exogenous $CO_2$ emissions from non-industrial sources such as land use change. $M_{pre}$ is the pre-industrial $CO_2$ stock, and $\phi(M,t)$ is calculated as in Appendix 1. The first term in the capital transition equation has capital depreciating at constant rate $\delta_k$, and the last two terms define capital investment as any available output not allocated to the control variables of consumption $c_t$ and abatement. Here, $\Psi_t$ and $a_2$ determine the cost of abating the chosen fraction $\mu_t$ of emissions. The term outside the brackets adjusts for the growth of labor and technology to keep capital in effective terms. Gross output $Y_{gross}$ is a function of the capital stock:

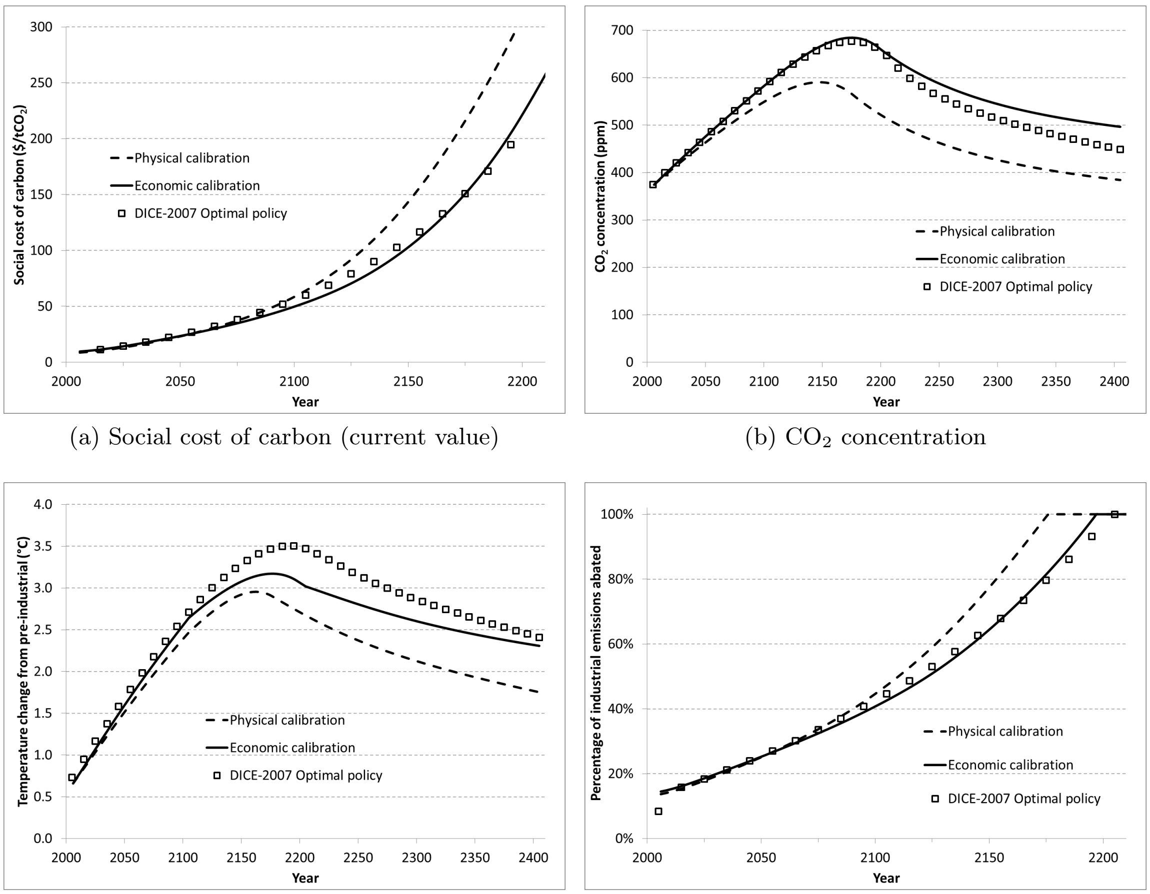
$$Y_{gross} = k_t^{\kappa}\ . \qquad \text{(Gross output)}$$

The parameter $\kappa$ gives the capital elasticity in a Cobb-Douglas production function. Climate damages $D$ reduce gross output more strongly as temperature increases:

$$D = b_2 T_t^{b_3}\ , \qquad \text{(Damages)}$$

where $b_3 = 2$ in DICE-2007 and temperature change $T_t$ relative to pre-industrial levels is as in Appendix 1.

The transition equations for the exogenous variables are as in DICE-2007, but adjusted for the annual timestep. We here list them and give the parameterization in Table 6. In each case, $t = 0$ corresponds to the year 2005. See Crost and Traeger (2010) for more details on variable definitions




(a) Social cost of carbon (current value)

(b) $CO_2$ concentration

(c) Temperature relative to the pre-industrial level

(d) Abatement

Figure 8: The optimized paths for our calibration and for DICE-2007. The "physical calibration" shows optimized policy with the transient feedback and decay rate functions ($\bar{f}_t(M,t)$ and $\bar{\phi}(M,t)$) inferred from DICE, and the "economic calibration" shows optimized policy with the adjusted functions $f_t(M,t)$ and $\phi(M,t)$.

and implementation.

$$A_t = A_0 \exp\left[\frac{g_{A,0}}{\delta_A}\left(1 - e^{-t\delta_A}\right)\right] \qquad \text{(Production technology)}$$
$$g_{A,t} = g_{A,0} e^{-t\delta_A} \qquad \text{(Growth rate of production technology)}$$
$$L_t = L_0 + (L_\infty - L_0)\left(1 - e^{-t\delta_L}\right) \qquad \text{(Labor)}$$
$$g_{L,t} = \delta_L \left[\frac{L_\infty}{L_\infty - L_0} e^{t\delta_L} - 1\right]^{-1} \qquad \text{(Growth rate of labor)}$$
$$\beta_t = \exp\left(-\rho + (1-\eta) g_{A,t} + g_{L,t}\right) \qquad \text{(Effective discount factor)}$$
$$\sigma_t = \sigma_0 \exp\left[\frac{g_{\sigma,0}}{\delta_\sigma}\left(1 - e^{-t\delta_\sigma}\right)\right] \qquad \text{(Uncontrolled emissions per output)}$$
$$\Psi_t = \frac{a_0 \sigma_t}{a_2}\left(1 - \frac{1 - e^{tg_\Psi}}{a_1}\right) \qquad \text{(Abatement cost factor)}$$
$$B_t = B_0 e^{tg_B} \qquad \text{(Non-industrial } CO_2 \text{ emissions)}$$
$$EF_t = EF_0 + 0.01(EF_{100} - EF_0)\min\{t, 100\} \qquad \text{(Non-} CO_2 \text{ forcing)}$$

The constraints prevent the decision-maker from using more than the output available after accounting for damages and from abating more than 100% of emissions in a period:

$$c_t + \Psi_t \mu_t^{a_2} \leq \frac{Y_{gross}}{1+D} \tag{17}$$
$$\mu_t \leq 1 \ . \tag{18}$$

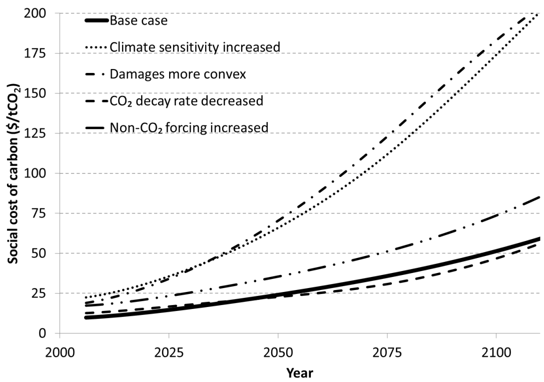
When the constraint in equation (17) is slack, we have positive capital investment, and when the constraint in equation (18) is slack, economic activity produces some $CO_2$ emissions that are not abated.

The challenge in solving the model lies not in finding the optimal actions for a given value function but in determining the value functions that satisfy the relations in equations (1) and (3) (see Kelly and Kolstad, 1999, 2001). We begin with a guess for the value function and a set of Chebychev nodes in the three-dimensional state space. We then use the initial guess for the continuation value to find each node's optimal controls $c_t^*$ and $\mu_t^*$ and optimal value. Knowing the optimal value at each Chebychev node, we approximate the value function across the rest of the state space using a set of Chebychev basis polynomials (Miranda and Fackler, 2002). We repeat the process using this approximated value function as the new initial guess, with iteration continuing until the coefficients of the value approximant's basis functions change by less than 0.0001.[18] When the temperature threshold is uncertain, the pre-threshold value function is smooth over the relevant state space because the hazard rate changes smoothly with time and the $CO_2$ level. However, when the threshold is known to be at $T^*$, each node in the state space is associated with either the pre-threshold or the post-threshold value function, and the combined value function

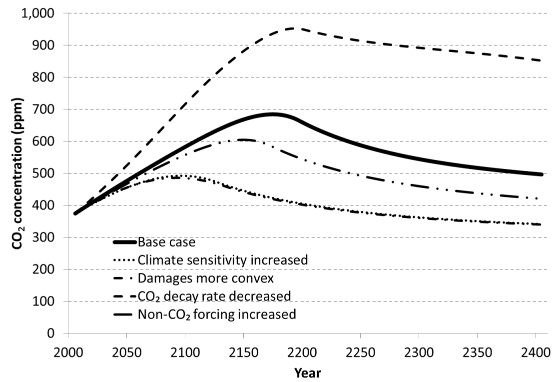
[18] To speed the solution procedure, we iterate over the value function calculation and approximation several times before re-optimizing the control variables. This is a type of modified policy iteration (Puterman and Shin, 1978).

Table 6: Parameterization of the transition equations. See Nordhaus (2008) and Crost and Traeger (2010) for more information. Also shows how these parameters change in the post-threshold regimes.

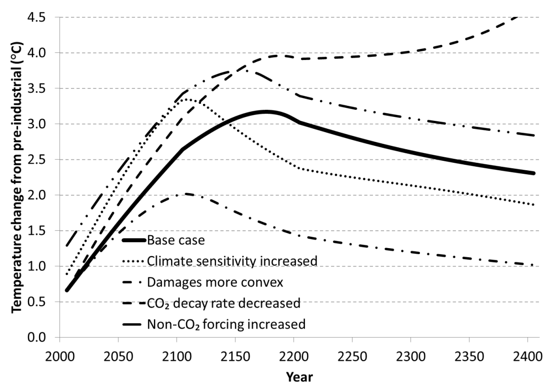
| Parameter | Value | Description |
|---|---|---|
| $A_0$ | 0.02722 | Initial production technology |
| $g_{A,0}$ | 0.009 | Initial annual growth rate of production technology |
| $\delta_A$ | 0.001 | Annual rate of decline in growth rate of production technology |
| $L_0$ | 6514 | Population in 2005 (millions) |
| $L_\infty$ | 8600 | Asymptotic population (millions) |
| $\delta_L$ | 0.035 | Annual rate of convergence of population to asymptotic value |
| $\sigma_0$ | 0.131418 | Initial emission intensity before emission reductions (GtC/output) |
| $g_{\sigma,0}$ | -0.00730 | Initial annual growth rate of emission intensity |
| $\delta_\sigma$ | 0.003 | Annual change in growth rate of emission intensity |
| $a_0$ | 1.17 | Cost of backstop technology in 2005 ($1000/tC) |
| $a_1$ | 2 | Ratio of initial backstop cost to final backstop cost |
| $a_2$ | 2.8 | Abatement cost exponent |
| $g_\Psi$ | -0.005 | Annual growth rate of backstop cost |
| $B_0$ | 1.1 | Initial non-industrial $CO_2$ emissions (GtC/y) |
| $g_B$ | -0.01 | Annual growth rate of non-industrial emissions |
| $EF_0$ | -0.06 | Initial exogenous forcing from non-$CO_2$ greenhouse gases (W $m^{-2}$) |
| $EF_{100}$ | 0.30 | Year 2105 exogenous forcing from non-$CO_2$ greenhouse gases (W $m^{-2}$) |
| $\kappa$ | 0.3 | Capital elasticity in Cobb-Douglas production function |
| $\delta_\kappa$ | 0.1 | Annual depreciation rate of capital |
| $b_2$ | 0.0028388 | Coefficient of temperature in the damage function |
| $b_3$ | 2 | Exponent on temperature in the damage function |
| $s$ | 3 | Climate sensitivity (°C) |
| $M_{pre}$ | 596.4 | Pre-industrial atmospheric $CO_2$ (GtC) |
| $\rho$ | 0.015 | Annual rate of pure time preference |
| $k_0$ | $137/(A_0 L_0)$ | Initial effective capital, with initial capital stock of 137 US$trillion |
| $M_0$ | 808.9 | Initial atmospheric $CO_2$ (GtC) |
| *Parameters for post-threshold regimes (i.e., for tipping points' effects)* | | |
| $\hat{s}$ | $2s$ | Climate sensitivity increased |
| $\hat{b}_3$ | $b_3 + 1$ | Damages more convex |
| $\hat{\phi}(M,t)$ | $\phi(M,t)/4$ | $CO_2$ sinks weakened |
| $\widehat{EF}_t$ | $EF_t + 1.5$ | Non-$CO_2$ forcing increased |



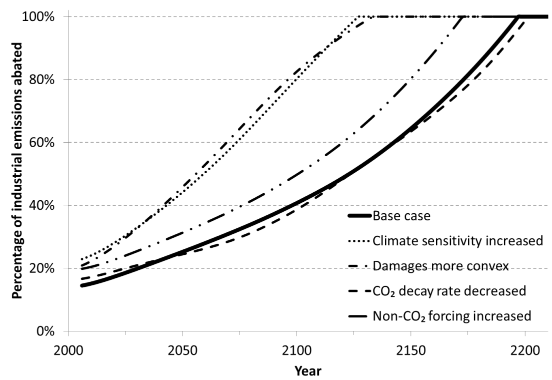

(a) Social cost of carbon (current value)




(b) Optimal $CO_2$ concentration




(c) Optimal temperature change from pre-industrial level




(d) Abatement

Figure 9: The evolution of the optimal social cost of carbon, $CO_2$ concentration, temperature, and abatement under expected temperature outcomes if each type of tipping point were crossed in 2005. These plots show how optimal policy and the climate respond to each tipping point's occurrence.

exhibits a discontinuity between regions in which optimal pre-threshold policy would and would not avoid crossing $T^*$. We undertake numerical approximation in this case by making the threshold stochastic in order to smooth out the value function (compare Brozović and Schlenker, 2011). This stochasticity comes from placing a low-variance normal distribution around $T^*$, which means that transitioning to a temperature above $T^*$ is virtually—but not totally—certain to cross the threshold. As the variance of this distribution approaches zero, we approach the case with a non-stochastic, certain threshold.[19]

We now describe how each post-threshold regime affects damages $D$, for simplicity ignoring

[19]The certain threshold model is currently solved with standard deviations between 0.01°C and 0.35°C, depending on the regime and threshold level under consideration. We are currently working on an improved solution method.

transient feedbacks. First, the climate sensitivity regime uses $\hat{s} = 2s = 6°$C, corresponding to $f_{atm} \approx 0.81$ (see Appendix 1). This change doubles equilibrium temperature, which increases equilibrium damages from $b_2T^2$ to $b_2(2T)^2$, where $T$ is calculated using a climate sensitivity of 3°C. Equilibrium damages therefore quadruple, though the presence of transient feedbacks reduces the effect on time $t$ damages. Second, increasing the convexity of damages means using $\hat{b}_3 = b_3 + 1 = 3$. This multiplies damages by $T$, which is always less than 4 in expectation in our baseline runs. Third, weakening $CO_2$ sinks means using $\hat{\phi}(M,t) = \phi(M,t)/4$, which increases the length of time for which a unit of $CO_2$ emissions affects the atmospheric $CO_2$ stock. The change from $\phi$ to $\hat{\phi}$ only has a significant effect on the $CO_2$ stock once enough time has passed for the additional accumulation to matter. Finally, increasing exogenous non-$CO_2$ forcing means using $\widehat{EF}_t = EF_t + 1.5$. This increases temperature by 1.2°C in equilibrium, which means damages increase in equilibrium to $b_2(T+1.2)^2$. When temperature is greater than 1.2°C, the effect of this tipping point on equilibrium damages is less than the quadrupling produced by the climate sensitivity tipping point, and when temperature is greater than 2.3°C, the effect of this tipping point on equilibrium damages is less than that of the damage convexity tipping point.

Figure 9 shows how optimal policy, temperature, and $CO_2$ concentrations would evolve if each tipping point occurred exogenously in 2005. Each regime affects policy and the climate differently, which in turn affects the degree to which the decision-maker tries to avoid that type of tipping point. Optimal policy responds in accord with how each tipping point affects damages. Increasing climate sensitivity or increasing damage convexity both raise the social cost of carbon and lower the $CO_2$ stock path. However, they have quite different effects on temperature because a given $CO_2$ stock produces higher temperatures when climate sensitivity is increased while a given temperature produces greater damages when the convexity of the damage function is increased. These results indicate how, aside from tipping points, IAMs' results are sensitive to assumptions about the uncertain parameters determining climate sensitivity and the convexity of the damage function. Decreasing the decay rate of $CO_2$ does not significantly affect abatement or the social cost of carbon, but the reduced stock decay does eventually produce higher $CO_2$ concentrations, temperatures, and damages. Emission decisions have impacts for a longer time than usual, but that change does not greatly affect the present value of the damage they produce. Finally, while increasing non-$CO_2$ forcing does increase temperature and damages, it does not affect the $CO_2$ stock path or the social cost of carbon as strongly as do the climate sensitivity or damage convexity tipping points. Entering this regime reduces economic output, but this effect does not interact as strongly with emission decisions.

## References

Ackerman, F., S. DeCanio, R. Howarth, and K. Sheeran (2009). "Limitations of integrated assessment models of climate change." Climatic Change 95(3-4): 297–315. doi:10.1007/s10584-009-9570-x.

Ackerman, F., E. A. Stanton, and R. Bueno (2010). "Fat tails, exponents, extreme uncertainty: Simulating catastrophe in DICE." Ecological Economics 69(8): 1657–1665. doi:10.1016/j.ecolecon.2010.03.013.

Alley, R. B., J. Marotzke, W. Nordaus, J. Overpeck, D. Peteet, R. Pielke Jr., R. Pierrehumbert, P. Rhines, T. Stocker, L. Talley, J. Wallace, A. Isern, J. Dandelski, C. Elfring, M. Gopnik, M. Kelly, J. Bachim, and A. Carlisle (2002). Abrupt Climate Change: Inevitable Surprises. Washington, DC: National Research Council, National Academy Press.

Alley, R. B., J. Marotzke, W. D. Nordhaus, J. T. Overpeck, D. M. Peteet, R. A. Pielke, R. T. Pierrehumbert, P. B. Rhines, T. F. Stocker, L. D. Talley, and J. M. Wallace (2003). "Abrupt climate change." Science 299(5615): 2005–2010. doi:10.1126/science.1081056.

Archer, D. (2007). "Methane hydrate stability and anthropogenic climate change." Biogeosciences 4(4): 521–544.

Ayong Le Kama, A., A. Pommeret, and F. Prieur (2011). "Optimal emission policy under the risk of irreversible pollution." FEEM Nota di Lavoro 14.2011.

Azar, C. and K. Lindgren (2003). "Catastrophic events and stochastic cost-benefit analysis of climate change." Climatic Change 56(3): 245–255. doi:10.1023/A:1021743622080.

Baker, M. B. and G. H. Roe (2009). "The shape of things to come: Why is climate change so predictable?" Journal of Climate 22(17): 4574–4589. doi:10.1175/2009JCLI2647.1.

Baranzini, A., M. Chesney, and J. Morisset (2003). "The impact of possible climate catastrophes on global warming policy." Energy Policy 31(8): 691–701. doi:10.1016/S0301-4215(02)00101-5.

Brozović, N. and W. Schlenker (2011). "Optimal management of an ecosystem with an unknown threshold." Ecological Economics 70(4): 627–640. doi:10.1016/j.ecolecon.2010.10.001.

Creyts, J., A. Derkach, S. Nyquist, K. Ostrowski, and J. Stephenson (2007). "Reducing U.S. greenhouse gas emissions: How much at what cost?" , McKinsey & Company.

Crost, B. and C. P. Traeger (2010). "Risk and aversion in the integrated assessment of climate change." CUDARE Working Paper 1104, University of California, Berkeley.

Eglin, T., P. Ciais, S. L. Piao, P. Barre, V. Bellassen, P. Cadule, C. Chenu, T. Gasser, C. Koven, M. Reichstein, and P. Smith (2010). "Historical and future perspectives of global soil carbon response to climate and land-use changes." Tellus B 62(5): 700–718. doi:10.1111/j.1600-0889.2010.00499.x.

Ellsberg, D. (1961). "Risk, ambiguity, and the Savage axioms." The Quarterly Journal of Economics 75(4): 643–669.

Fisher, A. C. and U. Narain (2003). "Global warming, endogenous risk, and irreversibility." Environmental and Resource Economics 25(4): 395–416. doi:10.1023/A:1025056530035.

Gerst, M. D., R. B. Howarth, and M. E. Borsuk (2010). "Accounting for the risk of extreme outcomes in an integrated assessment of climate change." Energy Policy 38(8): 4540–4548. doi:10.1016/j.enpol.2010.04.008.

Gjerde, J., S. Grepperud, and S. Kverndokk (1999). "Optimal climate policy under the possibility of a catastrophe." Resource and Energy Economics 21(3-4): 289–317. doi:10.1016/S0928-7655(99)00006-8.

Greenstone, M., E. Kopits, and A. Wolverton (2011). "Estimating the social cost of carbon for use in U.S. federal rulemakings: a summary and interpretation." National Bureau of Economic Research Working Paper Series No. 16913.

Guillerminet, M. and R. S. J. Tol (2008). "Decision making under catastrophic risk and learning: the case of the possible collapse of the West Antarctic Ice Sheet." Climatic Change 91(1): 193–209. doi:10.1007/s10584-008-9447-4.

Hall, D. C. and R. J. Behl (2006). "Integrating economic analysis and the science of climate instability." Ecological Economics 57(3): 442–465. doi:10.1016/j.ecolecon.2005.05.001.

Hanemann, M., S. H. Schneider, A. Rosencranz, M. Mastrandrea, and K. Kuntz-Duriseti (2010). "What is the economic cost of climate change?" In "Climate Change Science and Policy," pp. 185–193. Washington, D.C.: Island Press.

Hansen, J., M. Sato, P. Kharecha, D. Beerling, R. Berner, V. Masson-Delmotte, M. Pagani, M. Raymo, D. L. Royer, and J. C. Zachos (2008). "Target atmospheric CO2: Where should humanity aim?" The Open Atmospheric Science Journal 2: 217–231. doi:10.2174/1874282300802010217.

Hansen, L. P. and T. J. Sargent (2001). "Acknowledging misspecification in macroeconomic theory." Review of Economic Dynamics 4(3): 519–535. doi:10.1006/redy.2001.0132.

Hennlock, M. (2009). "Robust control in global warming management: An analytical dynamic integrated assessment." Discussion paper 09-19, Resources For the Future.

Huntingford, C., R. A. Fisher, L. Mercado, B. B. Booth, S. Sitch, P. P. Harris, P. M. Cox, C. D. Jones, R. A. Betts, Y. Malhi, G. R. Harris, M. Collins, and P. Moorcroft (2008). "Towards quantifying uncertainty in predictions of Amazon 'dieback'." Philosophical Transactions of the Royal Society B: Biological Sciences 363(1498): 1857 –1864. doi:10.1098/rstb.2007.0028.

Interagency Working Group on Social Cost of Carbon (2010). "Appendix 15a. Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866." United States Government.

Ju, N. and J. Miao (2009). "Ambiguity, learning, and asset returns." Working Paper wp2009-014, Boston University.

Karp, L. and Y. Tsur (2007). "Climate policy when the distant future matters: Catastrophic events with hyperbolic discounting." CUDARE working paper 1037, University of California, Berkeley.

Keller, K., B. M. Bolker, and D. F. Bradford (2004). "Uncertain climate thresholds and optimal economic growth." Journal of Environmental Economics and Management 48(1): 723–741. doi:10.1016/j.jeem.2003.10.003.

Kelly, D. L. and C. D. Kolstad (1999). "Bayesian learning, growth, and pollution." Journal of Economic Dynamics and Control 23(4): 491–518. doi:10.1016/S0165-1889(98)00034-7.

Kelly, D. L. and C. D. Kolstad (2001). "Solving infinite horizon growth models with an environmental sector." Computational Economics 18(2): 217–231. doi:10.1023/A:1021018417052.

Keynes, J. M. (1921). A Treatise on Probability. London: Macmillan and Co.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). "A smooth model of decision making under ambiguity." Econometrica 73(6): 1849–1892.

Klibanoff, P., M. Marinacci, and S. Mukerji (2009). "Recursive smooth ambiguity preferences." Journal of Economic Theory 144(3): 930–976. doi:10.1016/j.jet.2008.10.007.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl (2010). "Challenges in combining projections from multiple climate models." Journal of Climate 23(10): 2739–2758. doi:10.1175/2009JCLI3361.1.

Kriegler, E., J. W. Hall, H. Held, R. Dawson, and H. J. Schellnhuber (2009). "Imprecise probability assessment of tipping points in the climate system." Proceedings of the National Academy of Sciences 106(13): 5041–5046. doi:10.1073/pnas.0809117106.

Lange, A. and N. Treich (2008). "Uncertainty, learning and ambiguity in economic models on climate policy: some classical results and new directions." Climatic Change 89(1): 7–21. doi:10.1007/s10584-008-9401-5.

Le Quéré, C., C. Rodenbeck, E. T. Buitenhuis, T. J. Conway, R. Langenfelds, A. Gomez, C. Labuschagne, M. Ramonet, T. Nakazawa, N. Metzl, N. Gillett, and M. Heimann (2007). "Saturation of the Southern Ocean CO2 sink due to recent climate change." Science 316(5832): 1735–1738. doi:10.1126/science.1136188.

Leach, A. J. (2007). "The climate change learning curve." Journal of Economic Dynamics and Control 31(5): 1728–1752. doi:10.1016/j.jedc.2006.06.001.

Lemoine, D. M. (2010). "Climate sensitivity distributions depend on the possibility that models share biases." Journal of Climate 23(16): 4395–4415. doi:10.1175/2010JCLI3503.1.

Lempert, R. J., M. E. Schlesinger, and J. K. Hammitt (1994). "The impact of potential abrupt climate changes on near-term policy choices." Climatic Change 26(4): 351–376. doi:10.1007/BF01094402.

Lenton, T. M., H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber (2008). "Tipping elements in the Earth's climate system." Proceedings of the National Academy of Sciences 105(6): 1786–1793. doi:10.1073/pnas.0705414105.

Luo, Y. (2007). "Terrestrial carbon-cycle feedback to climate warming." Annual Review of Ecology, Evolution, and Systematics 38(1): 683–712. doi:10.1146/annurev.ecolsys.38.091206.095808.

Masur, J. S. and E. A. Posner (2010). "Climate regulation and the limits of cost-benefit analysis." John M. Olin Law & Economics Working Paper No. 525, The Law School, University of Chicago.

Millner, A., S. Dietz, and G. Heal (2010). "Ambiguity and climate policy." National Bureau of Economic Research Working Paper Series No. 16050.

Miranda, M. J. and P. L. Fackler (2002). Applied Computational Economics and Finance. Cambridge, Massachusetts: MIT Press.

Nævdal, E. (2006). "Dynamic optimisation in the presence of threshold effects when the location of the threshold is uncertain—with an application to a possible disintegration of the Western Antarctic Ice Sheet." Journal of Economic Dynamics and Control 30(7): 1131–1158. doi:10.1016/j.jedc.2005.04.004.

Nævdal, E. and M. Oppenheimer (2007). "The economics of the thermohaline circulation—A problem with multiple thresholds of unknown locations." Resource and Energy Economics 29(4): 262–283. doi:10.1016/j.reseneeco.2007.01.003.

Newbold, S. C. and A. Daigneault (2009). "Climate response uncertainty and the benefits of greenhouse gas emissions reductions." Environmental and Resource Economics 44(3): 351–377. doi:10.1007/s10640-009-9290-8.

Nordhaus, W. D. (1993). "Rolling the 'DICE': an optimal transition path for controlling greenhouse gases." Resource and Energy Economics 15(1): 27–50. doi:10.1016/0928-7655(93)90017-O.

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

Notz, D. (2009). "The future of ice sheets and sea ice: Between reversible retreat and unstoppable loss." Proceedings of the National Academy of Sciences 106(49): 20590 –20595. doi:10.1073/pnas.0902356106.

Oppenheimer, M. (1998). "Global warming and the stability of the West Antarctic Ice Sheet." Nature 393(6683): 325–332. doi:10.1038/30661.

Overpeck, J. T. and J. E. Cole (2006). "Abrupt change in Earth's climate system." Annual Review of Environment and Resources 31(1): 1–31. doi:10.1146/annurev.energy.30.050504.144308.

Plevin, R. J., M. OHare, A. D. Jones, M. S. Torn, and H. K. Gibbs (2010). "Greenhouse gas emissions from biofuels' indirect land use change are uncertain but may be much greater than previously estimated." Environmental Science & Technology 44(21): 8015–8021. doi:10.1021/es101946t.

Polasky, S., A. de Zeeuw, and F. Wagener (in press). "Optimal management with potential regime shifts." Journal of Environmental Economics and Management doi:10.1016/j.jeem.2010.09.004.

Puterman, M. L. and M. C. Shin (1978). "Modified policy iteration algorithms for discounted Markov decision problems." Management Science 24(11): 1127–1137.

Ramanathan, V. and Y. Feng (2008). "On avoiding dangerous anthropogenic interference with the climate system: Formidable challenges ahead." Proceedings of the National Academy of Sciences 105(38): 14245 –14250. doi:10.1073/pnas.0803838105.

Ramaswamy, V., O. Boucher, J. Haigh, D. Hauglustaine, J. Haywood, G. Myhre, T. Nakajima, G. Shi, and S. Solomon (2001). "Radiative forcing of climate change." In J. Houghton, Y. Ding, D. Griggs, M. Noguer, P. van der Linden, X. Dai, K. Maskell, and C. Johnson, eds., "Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change," p. 881. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Raupach, M. R., J. G. Canadell, and C. Le Qur (2008). "Anthropogenic and biophysical contributions to increasing atmospheric CO2 growth rate and airborne fraction." Biogeosciences 5: 1601–1613. doi:10.5194/bg-5-1601-2008.

Rockström, J., W. Steffen, K. Noone, Å. Persson, F. S. Chapin, E. F. Lambin, T. M. Lenton, M. Scheffer, C. Folke, H. J. Schellnhuber, B. Nykvist, C. A. de Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sörlin, P. K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R. W. Corell, V. J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, and J. A. Foley (2009). "A safe operating space for humanity." Nature 461(7263): 472–475. doi:10.1038/461472a.

Roe, G. H. (2009). "Feedbacks, timescales, and seeing red." Annual Review of Earth and Planetary Sciences 37(1): 93–115. doi:10.1146/annurev.earth.061008.134734.

Schaefer, K., T. Zhang, L. Bruhwiler, and A. P. Barrett (2011). "Amount and timing of permafrost carbon release in response to climate warming." Tellus B 63(2): 165–180. doi: 10.1111/j.1600-0889.2011.00527.x.

Schmidt, G. A. and D. T. Shindell (2003). "Atmospheric composition, radiative forcing, and climate change as a consequence of a massive methane release from gas hydrates." Paleoceanography 18(1): 1004. doi:200310.1029/2002PA000757.

Sitch, S., C. Huntingford, N. Gedney, P. E. Levy, M. Lomas, S. L. Piao, R. Betts, P. Ciais, P. Cox, P. Friedlingstein, C. D. Jones, I. C. Prentice, and F. I. Woodward (2008). "Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs)." Global Change Biology 14(9): 2015–2039. doi: 10.1111/j.1365-2486.2008.01626.x.

Smith, J. B., S. H. Schneider, M. Oppenheimer, G. W. Yohe, W. Hare, M. D. Mastrandrea, A. Patwardhan, I. Burton, J. Corfee-Morlot, C. H. D. Magadza, H. Fussel, A. B. Pittock, A. Rahman, A. Suarez, and J. van Ypersele (2009). "Assessing dangerous climate change through an update of the Intergovernmental Panel on Climate Change (IPCC) "reasons for concern"." Proceedings of the National Academy of Sciences 106(11): 4133–4137. doi:10.1073/pnas.0812355106.

Soden, B. J., I. M. Held, R. Colman, K. M. Shell, J. T. Kiehl, and C. A. Shields (2008). "Quantifying climate feedbacks using radiative kernels." Journal of Climate 21(14): 3504–3520. doi:10.1175/2007JCLI2110.1.

Tol, R. S. (2008). "The social cost of carbon: Trends, outliers and catastrophes." Economics: The Open-Access, Open Assessment E-Journal 2(2008-25).

Traeger, C. P. (2010). "Subjective risk, confidence, and ambiguity." CUDARE Working Paper 1103, University of California, Berkeley.

Tsur, Y. and A. Zemel (1996). "Accounting for global warming risks: Resource management under event uncertainty." Journal of Economic Dynamics and Control 20(6-7): 1289–1305. doi:10.1016/0165-1889(95)00900-0.

Vaughan, D. G. (2008). "West Antarctic Ice Sheet collapse—the fall and rise of a paradigm." Climatic Change 91(1-2): 65–79. doi:10.1007/s10584-008-9448-3.

von Neumann, J. and O. Morgenstern (1944). Theory of Games and Economic Behavior. Princeton: Princeton University Press.

Weitzman, M. L. (2009). "On modeling and interpreting the economics of catastrophic climate change." Review of Economics and Statistics 91(1): 1–19. doi:10.1162/rest.91.1.1.

Wright, E. L. and J. D. Erickson (2003). "Incorporating catastrophes into integrated assessment: Science, impacts, and adaptation." Climatic Change 57(3): 265–286. doi:10.1023/A:1022829706609.

Zachos, J. C., G. R. Dickens, and R. E. Zeebe (2008). "An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics." Nature 451(7176): 279–283. doi:10.1038/nature06588.

Zaliapin, I. and M. Ghil (2010). "Another look at climate sensitivity." Nonlinear Processes in Geophysics 17(2): 113–122. doi:10.5194/npg-17-113-2010.