NBER WORKING PAPER SERIES

CLIMATE TIPPING POINTS AND SOLAR GEOENGINEERING

Garth Heutel
Juan Moreno Cruz
Soheil Shayegh

Working Paper 21589
http://www.nber.org/papers/w21589

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2015

We thank Derek Lemoine and seminar participants at Tennessee, Stanford, Harvard, and the Southern Economic Association annual meeting for comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Climate Tipping Points and Solar Geoengineering
Garth Heutel, Juan Moreno Cruz, and Soheil Shayegh
NBER Working Paper No. 21589
September 2015
JEL No. C61,H23,Q54,Q58

ABSTRACT

We study optimal climate policy when climate tipping points and solar geoengineering are present. Solar geoengineering reduces temperatures without reducing greenhouse gas emissions. Climate tipping points are irreversible and uncertain events that cause large damages. We analyze three different rules related to the availability of solar geoengineering: a ban, using solar geoengineering as insurance against the risk of tipping points, or using solar geoengineering only as remediation in the aftermath of a tipping point. We model three distinct types of tipping points: two that alter the climate system and one that yields a direct economic cost. Using an analytic model, we find that an optimal policy, which minimizes expected losses from the tipping point, includes both emissions reductions and solar geoengineering from the onset. Using a numerical simulation model, we quantify optimal policy and various outcomes under the alternative scenarios. The presence of tipping points leads to more mitigation and more solar geoengineering use and lower temperatures.

Garth Heutel
436 Andrew Young School
Department of Economics
Georgia State University
PO Box 3992
Atlanta, GA 30302-3992
and NBER
gheutel@gsu.edu

Juan Moreno Cruz
School of Economics
Georgia Institute of Technology
221Bobby Dodd Way
Atlanta, Georgia 30332
juan.moreno-cruz@econ.gatech.edu

Soheil Shayegh
Carnegie Institution of Washington
260 Panama St.,
Stanford, CA 94305
sshayegh@carnegiescience.edu

1 Introduction

The accumulation of greenhouse gases in the atmosphere is associated with an increase in Earth's surface temperature, affecting economic performance and ecosystems as a whole. As temperature rises, the probability of crossing a climate tipping point (CTP) increases. CTPs are large disturbances that are rare, difficult to predict, and irreversible. The most common examples of such events are the collapse of the thermohaline circulation (THC) or the disintegration of the West Antarctic Ice Sheet (WAIS). Solar geoengineering (SGE), and more specifically solar radiation management (SRM), has been proposed as a way of limiting the probability of reaching a climate tipping point. By reducing the amount of radiation reaching Earth's surface, temperatures can be kept at a level below which catastrophes can occur even without reducing greenhouse gas concentrations. In this paper we analyze optimal climate policy in the presence of CTPs when both emissions reductions (mitigation) and SGE are available, using both an analytical theoretical model and numerical simulations.

We build a parsimonious analytic model of climate change economics with CTPs and SGE. We model a CTP as an irreversible event that changes the dynamics of the climate-carbon system, resulting in a welfare loss relative to the state of the world before the threshold is reached. The planner's problem is solved using stochastic dynamic programing techniques that allow us to accommodate the post-CTP transition in the system. In this model, the probability of reaching the tipping point is a stochastic function of the atmospheric temperature. An important characteristic of our model is that we build in a strong case of inertia, where temperatures *tomorrow* are a function of the stock of carbon in the atmosphere *today*, which is in turn a function of mitigation *yesterday*. On the other hand, SGE *today* affects temperatures *tomorrow*. Using the analytic model, we identify different roles for mitigation and SGE. While both instruments help reduce damages before and after reaching the CTP, SGE can reduce the risk of crossing the temperature threshold more quickly than can mitigation. We explore three different SGE rules currently discussed in the governance literature. The first rule is a *Ban*, in which society chooses not to engage in SGE under any circumstances. In the second rule, SGE is freely used in combination with mitigation. We call this the *Insurance* rule, since SGE can insure against the risk of reaching a CTP. Third, we consider a rule where SGE is allowed only when temperatures surpass the climate tipping point. This is called the *Remediation* rule, since SGE can be thought of as only a "last-resort" policy in the event that the tipping point is reached.

We then incorporate SGE into a quantitative integrated assessment model (IAM), the DICE (Dynamic Integrated Climate-Economy) model, following Heutel et al. (2015), to simulate a richer set of alternative scenarios allowing for both mitigation and SGE in the presence of CTPs. In the quantitative model, we relax many assumptions of the analytic model and confirm the results presented in the theory. The simulation model allows us to consider three distinct types of tipping points: a Climate Feedback CTP in which the climate sensitivity (the responsiveness of temperature to the carbon stock) is changed after the CTP, a Carbon Sink CTP in which the carbon dynamics are changed after the CTP, and an Economic Loss CTP in which there is a direct welfare loss from the CTP. We are able to quantify the effects of alternative SGE rules under various CTP specifications in terms of several outcome variables, including temperature, carbon stock, and the optimal carbon tax. The *Ban* rule yields a carbon tax that is twice as high as the other two rules. Under the *Insurance* rule, the risks associated with the tipping point are largely avoided. The *Remediation* rules, reduces damages and carbon taxes only when the threshold is crossed, but leaves current policy largely unaffected. Under all rules, and contrary to what has been expressed previously in the geoengineering literature, a substantial amount of mitigation is optimal to deal with the risks of climate change.

Our approach closely resembles that of Lemoine and Traeger (2014).[1] That paper uses a recursive version of DICE to consider CTPs where policymakers learn about the position of the tipping point, and where the costs associated with crossing the tipping point are a function of the state of the economy at the time the tipping point is crossed. Like our paper, Lemoine and Traeger (2014) model a Climate Feedback CTP and a Carbon Sink CTP. To that we add the Economic Loss CTP, as in Cai et al. (2013). Furthermore, we add SGE to the model. To best of our knowledge, our paper is the first to incorporate SGE in a model with CTPs.[2]

The use of SGE as part of the portfolio of options has been suggested in the literature under diverse scenarios. The use of SGE as an insurance against catastrophic climate change has been proposed early in the literature (Keith (2000), Victor (2008), Keith et al. (2010), Moreno-Cruz and Keith (2013)). The idea of SGE as a complement to mitigation is proposed in the literature as a way to achieve any given

[1] In the appendix we present an alternative model, which more closely follows Naevdal (2003), Naevdal (2006), and Naevdal and Oppenheimer (2007).

[2] By contrast, our earlier paper (Heutel et al. (2015)) and several others add SGE to an IAM but without CTPs.

temperature level at lower costs for society (Wigley (2006), Moreno-Cruz and Keith (2013), Heutel et al. (2015)). Finally, banning SGE has been proposed because of the large uncertainties surrounding the unintended consequences of SGE implementation and the asymmetry of impacts this intervention may have (Barrett (2008), Blackstock and Long (2010), Moreno-Cruz (2015), Victor (2008)).

A unique contribution of this paper in terms of methods is to model stochastic parameter values, rather than performing sensitivity analyses. Other studies have traditionally considered only sensitivity analyses but failed to develop a solution for the stochastic model. Among the papers that have actually modified DICE to include stochastic parameters are Baker and Solak (2011), Kolstad (1996) and Lemoine and Traeger (2014). However, none of these papers have included SGE as the source of uncertainty.

Other papers have added SGE to IAMs and examined the policy implications. Bickel and Lane (2009) and Goes et al. (2011) make several modifications to the DICE model, including allowing SGE and refining the climate dynamics. Their specification imposes an exogenous intermittency in SGE which makes it less effective. They present summaries of policies with an optimal mix of mitigation and SGE (subject to the intermittency). In contrast to Goes et al. (2011), Bickel and Agrawal (2013) find that under some scenarios a substitution of SGE for mitigation can pass a cost-benefit test. Gramstad and Tjøotta (2010) include SGE in DICE and conduct a cost-benefit analysis of SGE under various assumptions about the level undertaken and its costs. Emmerling and Tavoni (2013) use a different IAM, WITCH, to model SGE and mitigation policy. None of these papers consider the possibility of climate tipping points. See Heutel et al. (2015) for a more thorough comparison between our approach and previous papers introducing SGE on IAMs.

The rest of the paper is organized as follows. In section 2, we present our analytic model and its predictions. Section 3 describes and performs numerical simulations using the modified DICE model. A short conclusion section closes the paper. We leave for an appendix the description of our numerical approach as well as an alternative modeling framework.

2 Theoretical Model

We consider the case of a regulator who solves an infinite-horizon optimization problem with the goal of minimizing the total costs of climate change. In the model, the temperature threshold of CTPs is uncertain, there are different types of tipping points, and SGE and mitigation are imperfect substitutes.[3]

Optimal policy depends on the state of the world and the dynamics of the climate system. We use the following set of first order difference equations to represent the

[3]In the appendix, we present a simpler model in which the CTP threshold may be either uncertain or deterministic, where there is only one type of CTP, and where SGE and mitigation are perfect substitutes. Although these simplifications are unrealistic, the model presented in the appendix offers valuable intuitions relative to the more versatile model presented here and to the numerical simulations in the following section.

dynamics of the system:

$$S_{t+1} = e_t^{BAU} - m_t + (1 - \delta_t)S_t \tag{1}$$

$$T_{t+1} = \lambda_t(\beta \ln(S_t/S_0) - \theta_t g_t) + (1 - \gamma_t)T_t \tag{2}$$

$$S_0 > 0 \text{ and } T_0 > 0 \text{ given.}$$

Equation (1) captures the carbon dynamics. S_t is the stock of carbon in the atmosphere, e_t^{BAU} is business-as-usual emissions of greenhouse gases, m_t is mitigation, and δ_t is the absorption capacity of the planet. Equation (2) shows how temperature, T_t, responds to changes in radiative forcing at time t. The radiative forcing potential of carbon dioxide depends on the carbon stock S_t relative to its pre-industrial level S_0. β_t captures the relation between carbon concentrations and radiative forcing. g_t is the amount of SGE implemented at time t expressed in units of radiative forcing, and $\theta_t \in \{0, 1\}$ represents the rule regarding the availability of SGE: $\theta = 1$ when SGE is available and $\theta = 0$ when it is not. λ represents the climate sensitivity of the system that transforms radiative forcing into temperature levels. Finally, some fraction of the heat stored in the atmosphere escapes; this effect is captured by the term $\gamma_t T_t$ where γ_t is the heat transfer parameter (Naevdal and Oppenheimer (2007)). When γ_t approaches 1, temperature is only a function of atmospheric concentrations and independent of the temperature in previous periods.[4]

Equations (1)-(2) represent the inertia of the climate-carbon system in a simple way but highlights the main difference between mitigation and SGE: temperature in

[4] This timing of events also features predominately in Naevdal and Oppenheimer (2007) and Lemoine and Rudik (2014)

period $t+1$ is a function of the stock of carbon in period t which in turn is a function of mitigation in period $t-1$. That is, at time t society has already committed to an amount of warming in the next period. Mitigation efforts affect temperatures only two periods ahead. SGE in period t, on the other hand, affects temperatures in period $t+1$. Therefore, mitigation efforts in period t create benefits in future periods but can do little to reduce the warming we have committed for the next period, while SGE can alter temperatures more quickly, reducing the inertia of the climate-carbon cycle. These difference equations highlight the reason why SGE can serve as an insurance against catastrophic climate change.

We model a climate tipping point as an irreversible change in the climate-carbon system that occurs after a given temperature threshold is crossed. We define the vector $\nu_t = \nu = [\lambda_t, \beta_t, \delta_t, \gamma_t]$ that captures the state of the climate system at time t. When the threshold is crossed, and at least one of the parameters changes, the change in the dynamics of the system is represented by an vector $\nu_t = \tilde{\nu}$. If the CTP is reached at time $\bar{t}$, then we have:

$$\begin{aligned} \nu_{t+1} = \nu_t, \quad \nu_0 = \nu \\ \nu_{\bar{t}+1} - \nu_{\bar{t}} = \tilde{\nu} \end{aligned} \tag{3}$$

The regulator minimizes net costs, which are the sum of the costs of implementing mitigation, m_t, and SGE, g_t, plus the damages associated with climate change. The implementation costs are given by $c(m_t, g_t)$, where $c_m > 0$, $c_{mm} > 0$, $c_g > 0$, $c_{gg} > 0$ and $c_{mg} = 0$. Damages are given by $D(T_t, S_t, g_t)$ and are a function of the current state of the world. They are increasing and convex in temperature and atmospheric carbon concentrations, that is $D_T > 0$, $D_{TT} > 0$, $D_S > 0$, $D_{SS} > 0$. Solar geoengineering also create damages, $D_g > 0$ and $D_{gg} > 0$. To capture the current state of knowledge regarding the side-effects of SGE, we assume damages from SGE are stochastic. Damages per unit of SGE, ϕ_t, are known at time t, but not known at time $t+1$. We capture this stochastic process via a stochastic equation given:

$$\phi_{t+1} = f(\phi_t) \tag{4}$$

The exact location of the temperature threshold leading to a CTP is unknown to the regulator, but the probability of crossing the threshold is known to be an increasing function of the temperature at time t. In this specification of CTPs, the probability of crossing the threshold is captured by an endogenous hazard function given by $h(T_{t+1})$. This hazard function captures the idea that as temperature increases, the likelihood of crossing the threshold in the next period also increases.

We can solve the regulator's problem via backwards induction. We first analyze the situation after the CTP has be crossed, and then move backward to analyze the situation before the CTP has not been crossed. After the threshold is crossed, the value function is given by $V(S_t, T_t, \phi_t, \tilde{\nu})$, where $\tilde{\nu}$ captures the state of the dynamics of the climate-carbon system. We obtain the solution to $V(S_t, T_t, \phi_t, \tilde{\nu})$ by solving the following Bellman equation:

$$V(S_t, T_t, \phi_t, \tilde{\nu}) = \min_{m_t, g_t} \{c(m_t, g_t) + D(T_t, S_t, g_t|\phi_t) + \beta_t E_\phi[V(S_{t+1}, T_{T+1}, \phi_{t+1}, \tilde{\nu})]\} \tag{5}$$

subject to equations (1)-(2) and the stochastic equation (4). We take expectations over the future value function because SGE damages are stochastic.

Before crossing the CTP, the Bellman equation of this problem is as follows:

$$\begin{aligned} V(S_t, T_t, \phi_t, \nu) = &\min_{m_t, g_t} \{c(m_t, g_t) + D(T_t, S_t, g_t|\phi_t) \\ &+\beta \mathbb{E}_\phi \left[(1 - h(T_{t+1}))V(S_{t+1}, T_{t+1}, \phi_{t+1}, \nu) + h(T_{t+1})V(S_{t+1}, T_{t+1}, \phi_{t+1}, \tilde{\nu})\right]\} \end{aligned} \tag{6}$$

where $V(S, T, \nu)$ is the value function in period t given the state of the world. With probability $1 - h(T_t, T_{t+1})$ the system remains unchanged and with probability $h(T_t, T_{t+1})$ the CTP is crossed.

The first order conditions with respect to mitigation and SGE are given by the following equations:[5]

$$\begin{aligned} c_m(m_t, g_t) + \beta \mathbb{E}_\phi \Bigg(& V_S(S_{t+1}, T_{t+1}, \phi_{t+1}, \nu)\frac{\partial S_{t+1}}{\partial m_t} + \\ & \underbrace{h(T_{t+1}) \left[\tilde{V}_S(S_{t+1}, T_{t+1}, \phi_{t+1}, \tilde{\nu}) - V_S(S_{t+1}, T_{t+1}, \phi_{t+1}, \nu)\right] \frac{\partial S_{t+1}}{\partial m_t} \Bigg)}_{DWI_S} = 0 \end{aligned} \tag{7}$$

[5] To simplify notation we write $X_Y(Y) \equiv \partial X(Y)/\partial Y$.

and

$$c_g(m_t, g_t) + \beta \mathbb{E}_\phi \Bigg(V_T(S_{t+1}, T_{t+1}, \phi_{t+1}, \nu) \frac{\partial T_{t+1}}{\partial g_t} + \underbrace{h(T_{t+1}) \left[V_T(S_{t+1}, T_{t+1}, \phi_{t+1}, \tilde{\nu}) - V_T(S_{t+1}, T_{t+1}, \phi_{t+1}, \nu) \right] \frac{\partial T_{t+1}}{\partial g_t}}_{DWI_T} + \underbrace{h_T(T_{t+1}) \frac{\partial T_{t+1}}{\partial g_t} \left[V(S_{t+1}, T_{t+1}, \phi_{t+1}, \tilde{\nu}) - V(S_{t+1}, T_{t+1}, \phi_{t+1}, \nu) \right] \Bigg)}_{MHE} = 0 \tag{8}$$

The main difference between mitigation and SGE can be seen by comparing these two equations. The interpretation of equation (7) is straightforward. The marginal cost of mitigation equals the expected marginal climate damages from one extra unit of carbon in the atmosphere. The term V_S is the reduction in future climate costs achieved by reducing the stock of carbon in the atmosphere by one unit. The term $h[\tilde{V}_S - V_S]$ is called the "differential welfare impact", DWI_S and captures the difference in the marginal climate costs associated with changes in the carbon stock incurred if the system crosses a CTP (Lemoine and Traeger (2014)).

The interpretation of equation (8) yields a similar result, but includes one extra adjustment. Equation (8) states that the marginal costs of SGE equal the marginal benefits, given by the expected reduction in damages associated with a marginal reduction in temperature. The second term, $h[\tilde{V}_T - V_T]$, DWI_T, is the differential welfare impact associated with a change in temperature. The third term, $h_T[\tilde{V} - V]$, is the "marginal hazard effect", MHE, captures the marginal reduction in the hazard associated with an increase in SGE. The MHE does not appear in equation (7) because the threshold is a function of temperature in the next period, not carbon concentrations. Thus, the MHE captures the insurance properties associated with SGE: by increasing SGE we reduce the hazard rate and thus reduce the expected costs of climate change.

2.1 Comparing SGE rules

The regulator chooses the optimal levels of mitigation and SGE subject to one of three rules regarding SGE availability. These three rules encompass different options presented in the solar geoengineering debate, that we assume are exogneous to the regulator:

(a) Ban: SGE is never allowed; $\theta_t = 0$ for all t.

(b) Insurance: SGE is always allowed; $\theta_t = 1$ for t.

(c) Remediation: SGE is allowed only after the CTP has been reached; $\theta = 0$ for $t < \bar{t}$ and $\theta = 1$ for $t > \bar{t}$.

We expect the behavior of the system to satisfy the following hypotheses. Before the CTP is crossed:

i.) The optimal amount of mitigation under different rules is such that:

$$m^{ban} > m^{remediation} > m^{insurance}$$

ii.) The optimal amount of SGE under different rules is such that:

$$0 = g^{ban} < g^{remediation} < g^{insurance}$$

iii.) Temperature levels under different rules is such that

$$T^{ban} > T^{remediation} > T^{insurance}$$

iv.) Atmospheric carbon concentrations levels under different rules is such that:

$$S^{ban} < S^{remediation} < S^{insurance}$$

The intuition regarding the previous hypothesis is as follows. Consider first the *Ban* rule. In this case, SGE is zero at all times, and so equation (8) does not apply. Under this setup of enhanced inertia, the role of mitigation is to decrease damages, but little can be done to decrease the propensity to cross the tipping point. This is related to the notion of committed warming: temperatures will continue to increase for several decades into the future, even if we reduce emissions today. DWI_S in equation (7) implies that the benefits of mitigation occur in the future, and mitigation reduces damages before and after the CTP is crossed. While mitigation cannot do much about the propensity to cross the CTP in the immediate future, mitigation reduces the risk of crossing the threshold in the long-term. This effect implies that the presence of CTPs increases the optimal amount of mitigation, relative to the case without CTPs and all damages have to be dealt with mitigation alone.

Next, consider the *Insurance* rule, where SGE can be freely used at any period. This is the case captured by equations (7)-(8); both mitigation and SGE are used to tackle climate change. That is, by construction the *Insurance* rule represents the optimal policy, and the outcomes under the other two rules must be sub-optimal.[6]

[6] This of course follows from the assumption that all costs, damages, and risks of SGE are included in our model. Bans or limits on SGE use are generally recommended due to the fear of unforeseen damages excluded from models.

The differential welfare impacts DWI_S and DWI_T increase both mitigation and SGE, and the MHE increases SGE. We can also show that, given functional form assumptions on implementation costs and climate damages, the introduction of SGE reduces the amount of mitigation and increases atmospheric carbon concentrations, relative to the *Ban* rule.

Finally, consider the *Remediation* rule, where SGE can be used only after the CTP has been crossed. Under this rule, both mitigation and SGE levels account for the DWI, but now the MHE cannot be dealt with using SGE because once the threshold is crossed, the changes in the climate-carbon system cannot by reversed, even if we substantially reduce temperatures with SGE. It follows that, relative to the *Insurance* rule, the amount of SGE will be lower, and the amount of mitigation will be higher. This in turn results in higher temperatures and lower carbon concentrations before the threshold is crossed, relative to the *Insurance* rule. Relative to the *Ban* rule, mitigation is lower, carbon concentrations are higher, and temperature is lower.

To corroborate our intuition, in the next section we develop and implement and numerical simulation that allows us to explore the dynamics of the system in a more comprehensive framework.[7]

3 Numerical Simulations

The analysis presented in the preceding sections rely on a parsimonious model. In this section, we extend the generality of our analysis by modifying an integrated as-

[7]In the simpler model presented in the appendix, we are able to theoretically prove these hypotheses, while in our more complicated model they remain hypotheses. See the appendix for details on the proofs and the intuitions behind them.

sessment model, DICE, to incorporate CTPs and the possibility of SGE. We consider several types of CTPs and allow uncertainty, not only on the location of the tipping point, but also on the damages from SGE.

3.1 Summary of Modifications to DICE

The dynamic integrated climate-economy (DICE) model has been widely used to study climate change and optimal climate policy. A summary of the model's assumptions and equations is available in Nordhaus (2008) and replicated for completeness in the appendix. The model can be used to calculate optimal climate mitigation policy, including the optimal path for the carbon price.

Here we briefly summarize our modifications to DICE. These are based on the modifications in Heutel et al. (2015), and more detail is available there, as well as in this paper's appendix. There are six modifications made to DICE to incorporate SGE and CTPs.

3.1.1 SGE Intensity

We include a choice variable for the intensity of SGE, g, analogous to DICE's choice variable for the intensity of mitigation, m. Thus, in addition to choosing an optimal mitigation path, our model solves for an optimal SGE path. Both m and g are proportions; m is the proportion of emissions that are abated and is between 0 and 1. g is the proportion of radiative forcing that is reduced (see below), and it can take values greater than 1.

3.1.2 SGE's Effect on Radiative Forcing

SGE affects the radiative forcing of Earth's atmosphere, reducing the amount of sunlight entering and thereby reducing temperature. DICE has a dynamic model of temperature based on radiative forcing, and radiative forcing itself is determined by carbon concentrations. SGE reduces radiative forcing directly, therefore almost instantaneously reducing temperatures. Setting SGE to $g = 1$ corresponds to reducing radiative forcing to its pre-industrial levels. By considering $g > 1$, we effectively allow for SGE to reduce temperature even below preindustrial levels, which can be necessary to deal with the inertia of the climate system and the warming we have already committed to.

3.1.3 SGE Implementation Cost

SGE implementation is costly. Our specification of costs is analogous to DICE's specification of the cost of mitigation. It is a convex (quadratic) function of the intensity of SGE g. It is calibrated from back-of-the-envelope calculations based on Crutzen (2006), Rasch et al. (2008), and a recent study on the costs of deployment (McClellan et al. (2012)). As in DICE, costs are expressed as a fraction of gross output. Implementation costs are small; in our calibration the costs of SGE at intensity $g = 0.1$ is 0.06% of gross output. Instead, the larger costs of SGE come from its potential damages.

3.1.4 SGE Damages

Solar SGE may directly cause damages, for instance, by reducing the upper ozone layer (Heckendorn et al. (2009)). We model these damages analogously to DICE's specification of damages from climate change. They are modeled as a fractional lost of potential output. We know of no study that attempts to quantify these damages, and thus this parameterization is inherently uncertain. We attempt to be conservative (i.e., biased against SGE) in our parameterization and assume that full SGE ($g = 1$) causes damages equal to 3% of gross output. This is of the order of climate change damages in DICE from a 6 degrees Celsius temperature increase.

We also assume that SGE damages are stochastic and are drawn from a log-normal probability distribution before each period starts.

3.1.5 Climate Change Damages Directly from Carbon

In DICE, climate change damages are a function of global temperature only. Since SGE will reduce temperatures but not reduce atmospheric or ocean carbon concentrations, in our model damages from climate change are separated out between damages from temperature, from atmospheric carbon concentrations, and from ocean carbon concentrations. High ocean carbon concentrations result in ocean acidification, which can lead to damages (Brander et al. (2012)). High atmospheric carbon concentrations may yield benefits (Pongratz et al. (2012)) or damages (Bony et al. (2013)). Just like with damages from SGE, these damages are mostly unknown. We keep the total level of climate change damages identical to the calibrated level in DICE. We assume that the majority (80%) of climate change damages come directly from temperature, but a small amount of damages may come from ocean concentrations (10%) and from atmospheric concentrations (10%). As shown in Heutel et al. (2015), this implies that SGE is not a perfect substitute for mitigation.[8]

3.1.6 Climate Tipping Points

The incorporation of climate tipping points into DICE along with SGE is unique to this paper and not found in Heutel et al. (2015)[9]. CTPs are modeled as irreversible events. In dynamic programming language, these are absorbing states, meaning that once we hit a tipping point we enter a new state in terms of climate or economic systems, where there is no chance of returning to the old state. We consider three types of CTPs: two affecting climate dynamics and one affecting economic costs. The first two CTPs are analogous to the two CTPs modeled in Lemoine and Traeger (2014).

(i) Climate feedback: Crossing this CTP strengthens the temperature feedback loop by increasing the marginal effect of carbon on temperature.[10] In the IAM, after this CTP is crossed the climate sensitivity variable increases from 3°C to 5°C.[11]

(ii) Carbon sink: Crossing this CTP reduces the natural capacity of the planet

[8] In Heutel et al. (2015), we explore this calibration using sensitivity analysis. The qualitative behavior of the system remains the same so long as the temperature damages dominate the outcomes.

[9] CTPs are incorporated into DICE in Lemoine and Traeger (2014), but without SGE.

[10] In our analytical model, this amounts to an increase in λ_t in equation (2).

[11] Climate sensitivity measures the steady-state temperature increase due to doubling atmospheric carbon levels; see the appendix for details.

to absorb carbon.[12] In the IAM, after crossing this CTP, carbon sinks are weakened by 50%.[13]

(iii) Economic loss: Crossing this CTP causes a loss of economic welfare equivalent to a 10% proportional increase in the damages from climate change.

As in Lemoine and Traeger (2014), the probability of reaching a CTP in the next state is a function of the atmospheric temperature in the current state. A CTP is reached once we cross an unknown threshold temperature. The CTP threshold temperature takes a uniform distribution. The minimum value is the current temperature (since once the current temperature has been reached, we know the CTP threshold cannot be below it), and the maximum temperature is calibrated so that the expected value of the threshold temperature is $2.5°C$ in 2005. Therefore, in each period, the probability of reaching the threshold temperature in the next period is uniformly distributed between T_t, the current temperature, and $\bar{T}$, the upper limit temperature:

$$p = \max\left\{0,\ \frac{min(T_{t+1}, \bar{T}) - T_t}{\bar{T} - T_t}\right\} \tag{9}$$

Uncertainty in the timing of tipping points is introduced in the DICE model as a binary variable with its probability associated with observed atmospheric temperature. The value of this variable is set to zero for the states before crossing the tipping point. Once a tipping point is crossed the variable changes into one and stays at one for the rest of the simulation. Depending on the type of CTP, subsequent state variables (including temperature, carbon concentration, and economic output) are

[12] In our analytical model, this is a decrease in δ_t in equation (1).

[13] See the appendix for details.

calculated.

3.2 Solution Algorithm

The evolution of the climate-economy system under uncertain tipping points is modeled as a Markov decision process. We define S_t as the state variable with multiple dimensions. For this problem, the state variable has eight dimensions: capital, atmospheric temperature, lower ocean temperature, atmospheric carbon concentration, upper ocean carbon concentration, lower ocean carbon concentration, radiative forcing, and a binary state variable capturing whether or not the CTP has been crossed. Given the values of the state variable parameters at each time step, the mitigation action, the SGE action, and the realization of uncertainty (crossing the tipping point), we can calculate the state variable parameters for the next time step.

Since the state variable parameters and action space are continuous, finding an exact solution for this problem through conventional backward induction methods is infeasible. Therefore, to solve the DICE model with stochastic tipping points, we use the two-step-ahead approximation method described in Shayegh and Thomas (2015). The approximation technique was tested and tuned in the deterministic case and then applied to the stochastic model.[14] In this technique at each time step t, a value function $\bar{V}_t$ is defined and used to capture the future utility from taking a candidate action a_t:

[14] To test the accuracy of this solution algorithm, we use it to replicate the results in Lemoine and Traeger (2014). This exercise is described in the appendix. It is worth emphasising that our approach does not require a reduction in the dimension of the state space, and we are able to solve the problem using the full set of transition equations used in the original DICE model.

$$\hat{V}_t(S_t) = \max_{a_t}\{U_t(S_t, a_t) + \bar{V}_t(S_t)\} \tag{10}$$

where $\hat{V}_t(S_t)$ is the optimal value of state S_t based on the value approximation. The advantage of this technique is in using endogenous parameters to calculate the value function approximation by assuming a deterministic trajectory for the two steps into the future at any given time. The deterministic trajectory allows us to calculate the utilities of these two future steps and bring them back to the present time using an artificial and tunable discount rate. The adjusted value is then used as a proxy for the uncertain value of all future states. These values reflect the social utility under the deterministic assumption and are used to construct the value function of the current state. The optimal action (mitigation and SGE) is found by maximizing this value function. The algorithm starts at time $t = 1$ and progresses until the last time step. After calculating all value functions, these values are used to update the coefficients of the approximate value function in previous states. The algorithm then iterates until the error (the difference between approximate values of $\bar{V}_t$ and optimal values of $\hat{V}_{t+1}$) converges to approximately zero.

The algorithm is developed in MATLAB and is available upon request. The full description of the model and approximation algorithm is presented in the appendix.

3.3 Results and Discussion

In this section we discuss the results from applying our numerical simulation to the three types of tipping points under the three different rules regarding the availability of SGE. We analyze the optimal climate policy portfolio of mitigation and SGE, and then the resulting carbon price, temperature, and carbon concentrations.

3.3.1 Optimal policy intervention

In Figure 1 we present the amount of mitigation and SGE implemented under different CTP types and SGE rules. The three CTP types are organized by rows, and the three SGE rules are organized by columns. We begin by comparing a single row across columns and then move on to compare across rows.

In each panel, the horizontal axis shows the year of the simulation. The vertical axis shows the amount of climate intervention for mitigation and SGE. Because mitigation is expressed as a fraction of total emissions and SGE is a fraction of total radiative forcing, we can plot them both in the same axis. Mitigation is constrained to be a number between 0 and 1, because we do not allow for negative emissions. SGE, on the other hand, can exceed 100% reduction in radiative forcing. The optimal amount of mitigation is shown in blue and the optimal amount of SGE is shown in orange. As a baseline case, we plot two bold continuous lines to capture the amount of mitigation and SGE in the absence of any form of CTPs. The average value of mitigation and SGE for the uncertain cases are shown in dashed lines, and the shaded areas represent the 95% confidence intervals, given the uncertainty of the CTP threshold.

Thus, within any panel, to compare policy with and without CTPs, compare the solid line (without CTPs) to the dotted line (the mean value with CTPs) or the shaded area (the 95% confidence interval with CTPs). To compare policy with and without SGE, compare results under the *Ban* rule (no SGE) to those under the *Insurance* rule (free and optimal use of SGE). The *Remediation* rule simulations represent the case where SGE can only be used after the CTP is reached.

The top-left panel in Figure 1 shows the *Ban* rule, where SGE is not allowed. By construction, SGE is zero for all periods. Comparing the continuous blue line to the dashed blue line shows that the presence of a CTP increases the average amount of mitigation in each period. This is consistent with recent findings in the literature (Lemoine and Traeger, 2015). The blue area comprises the 95% confidence interval for optimal mitigation. Under this scenario, once the CTP is crossed, the marginal impact of one extra unit of emissions increases and therefore the benefits of mitigation also increase. The top of the blue zone captures the realization where the tipping point is not crossed and the bottom line captures the realization where the tipping point is crossed earliest. The differences between the continuous line and the bottom line of the blue zone captures the extra benefits of mitigation after the threshold is crossed. The difference with the top of the blue zone also captures the risk reduction effects of early mitigation intervention. Because once the CTP is crossed the risk reduction motive disappears, the amount of mitigation falls.

Next, consider the case of the *Insurance* rule in the top-middle panel in Figure 1. Here, both mitigation and SGE are used from the beginning of the simulation. Comparing this panel and the first panel, we can see that the amount of mitigation is substantially reduced relative to the *Ban* rule, while SGE comprises the bulk of climate policy. The attractiveness of SGE stems from factors other than its costs, mainly the quickness of response that allows SGE to reduce future damages not only by reducing the DWI but also by reducing the MHE. The increase in SGE due to the MHE is the difference between the horizontal axis and the bottom line in the orange area. To see this, compare the orange areas under the *Insurance* rule and the *Remediation* rule. As we move to the *Remediation* rule in the right panel, the effect of MHE on SGE is eliminated. Also, under the *Remediation* rule the amount of mitigation can be just as high as under the *Ban* rule, when the CTP is never crossed (and no SGE is allowed), and as low as under the *Insurance* rule, when the CTP is crossed early.

In the last two panels in the top row of Figure 1, the shaded ares are larger, relative to the *Ban* rule. This reflects the introduction of uncertainty in the damages of SGE, but also the quick response associated with SGE, which makes it more responsive to risk and its resolution.

Next we discuss the other CTP types. The second row presents the results for the Carbon Sink tipping point. Under the *Ban* rule, there is almost no difference in the mitigation level whether or not CTPs are possible. For the other rules, there is only a small difference between the optimal mitigation levels. SGE behaves about the same under the *Insurance* rule as it does for the other CTP type, but the amount of SGE is slightly larger under the *Remediation* rule for the Carbon Sink CTP. The reduction in the natural carbon decay rate caused by the Carbon Sink CTP increases temperature, thereby increasing the marginal benefits of SGE.

The third row presents simulations under an Economic Loss tipping point. Policy under this CTP is very different than the other two CTPs because the economic loss, once the threshold is crossed, cannot be attenuated by either mitigation or SGE. Thus, before the threshold is reached, both mitigation and SGE will be implemented at a higher intensity relative to the other CTP types, because the risk of crossing the CTP is too high. But once the threshold is reached, the incentives for mitigation and SGE are drastically diminished. Damages are proportional to output, so a reduction in output also reduces damages and therefore the incentives to do either mitigation or SGE.[15]

Figure 1 verifies the hypothesis derived from our theoretical model. It shows that $m^{ban} > m^{remediation} > m^{insurance}$ and that $0 = g^{ban} < g^{remediation} < g^{insurance}$ (although the difference between $m^{remediation}$ and $m^{insurance}$ is insubstantial). There is no qualitative difference across CTP types in policy response. However, SGE behavior is qualitatively different between the *Insurance* and *Remediation* rules. Under the *Insurance* rule, average SGE intensity with CTPs (the dotted line) is higher than SGE intensity without CTPs (the solid line). Under the *Remediation* rule, the opposite is true.

3.3.2 Optimal carbon price

In Figure 2, we present the optimal carbon price (in \$/tC) under the different CTP types and SGE rules. The panels are organized as in Figure 1. The green shaded area shows the 95% confidence interval for the possible CTP outcome realizations. The top line reflects cases where the CTP is not reached inside the planning horizon, and the bottom line reflects cases where the CTP is reached very early. The red line shows the outcome when there are no CTPs, and the dashed line shows the mean value of the stochastic cases.

[15]This result is also found in Cai et al. (2013).

We consider first the case of the *Ban* rule under the Climate Feedback tipping point. The presence of a CTP increases the price of carbon, reflecting the risk associated with a tipping point. Crossing the CTP early reduces the carbon price as those risks are eliminated. Moving to the next two panels, we can see that the carbon price is the lowest under the *Insurance* rule, followed by the *Remediation* rule. When SGE is introduced, the marginal damages associated with each unit of carbon in the atmosphere are reduced, less mitigation is used and the carbon price falls.[16] Under the *Insurance* rule, the uncertainty band is narrow, but under the *Remediation* rule, the uncertainty band is very large, reflecting the use of SGE only after the CTP has been reached. If the CTP is never reached, since SGE cannot be used, mitigation levels remain high and so the carbon price stays also high. If the CTP is reached early, SGE is used and so mitigation and the carbon price fall.

Comparing these results to the two other CTP types, we see generally the same outcomes. The carbon price is the lowest and exhibits the narrowest uncertainty band under the *Insurance* rule. The uncertainty in the carbon price is highest for the Economic Loss tipping point. As soon as the tipping point is reached, irreversible losses are incurred and there is little that can be done to deal with those losses. When SGE use is restricted, either under the *Ban* rule or the *Remediation* rule, the mean carbon price and the uncertainty in the carbon price are substantially higher than when SGE use is unrestricted (*Insurance* rule).

[16] This is consistent with the findings in Heutel et al. (2015)

3.3.3 Temperature and Atmospheric Carbon Concentrations

Lastly, we show how temperature and carbon concentrations behave under the different scenarios. Figure 3 presents the temperature (in degrees ° C deviation from preindustrial temperatures). For all CTP types, the lowest temperature occurs under the *Insurance* rule. SGE is used most intensively when it is used to avoid crossing a CTP that is difficult to control with SGE after being crossed; this is true of the Carbon Sink and Economic Loss tipping points. Under the *Insurance* and *Remediation* rules we even observe negative temperature changes at the end of the planning horizon. The uncertainty regarding the future temperatures is largest under the *Remediation* rule. Since using SGE after crossing the tipping point cannot eliminate all damages from the tipping point, the mean value of temperature under the *Remediation* rule (dotted line) is higher than in the case without any tipping points (solid line).

In Figure 4 we show the resulting atmospheric carbon concentrations, measured in GtC, under the optimal amount of climate intervention in each scenario. Compared to the *Ban* rule, allowing SGE yields the highest concentrations of carbon, although temperatures are lower. In the time horizon presented (200 years), carbon eventually decreases under the *Ban* rule but continues to increase under the other rules. The Carbon Sink CTP exhibits the highest carbon concentrations across all SGE rules and also the highest amounts of uncertainty. The *Insurance* rule, while keeping temperature to very low levels, allows for high concentration levels.

Figures 3 and 4 are consistent with our initial hypotheses. Temperatures are such that $T^{ban} > T^{remediation} > T^{insurance}$ and carbon concentration levels are such that $S^{ban} < S^{remediation} < S^{insurance}$. Carbon concentrations are highest when SGE is allowed (*Insurance* and *Remediation*), reflecting the fact that SGE and mitigation are (imperfect) substitutes.

3.3.4 Summary

These simulations demonstrate that SGE can be used as a substitute, albeit an imperfect substitute, for mitigation in managing the risks of CTPs. Without the availability of SGE (the *Ban* rule), the presence of CTPs causes more mitigation to be used and consequently a higher carbon price. Depending on the type of CTP, temperatures and carbon stocks may be higher or lower with the CTP than without it.

Under the optimal policy portfolio (the *Insurance* rule), CTPs increase the use of SGE but do not substantially affect mitigation or the optimal carbon price. Thus, nearly all of the risk of CTPs is managed by SGE rather than by mitigation.

When SGE is restricted to only be allowed after the CTP is reached (the *Remediation* rule), mitigation is used much more intensively before the CTP is crossed, since it is the only policy option that can manage that risk. Once the CTP is crossed and SGE is allowed, SGE is used less intensively than under the *Insurance* rule, since there is no benefit in terms of reduced probability of CTP risk (no marginal hazard effect). The *Remediation* rule does not increase welfare and in fact substantially increases policy uncertainty, especially in regards to optimal mitigation levels and the carbon price.

4 Conclusion

We consider optimal climate policy when solar geoengineering is included as a policy option and tipping points are potential threats. Solar geoengineering is part of the optimal policy portfolio for two reasons. First, it provides a means to control temperature at (potentially) a lower cost than mitigation. Second, it can be used as insurance against the risk of reaching a climate tipping point. Thus, refraining from using SGE only until a tipping point has been reached (our *Remediation* rule) is not a welfare-maximizing policy.

Our analytic results were reached using a simple model; we have done so to concentrate on the importance of SGE in dealing with mega-disasters caused by CTPs. Our numerical approach modifies the DICE model to incorporate SGE, three rules governing its use, and three types of tipping points. The simulation results confirm our predictions from the analytical model, but they also provide us with a quantitative characterization of alternative policy scenarios. As with any integrated assessment model, results depend on the parametrization and calibration of the model, much of which could be highly speculative.[17]

We find that tipping points call for more action, but this action can take the form of a combination of mitigation and SGE, rather than mitigation alone. This allows for a climate policy with lower carbon taxes and overall lower risk, relative to a world without SGE. SGE will not eliminate the risk from CTPs altogether, but it may substantially reduce it.

[17] See Pindyck (2013) for a critique of IAMs.

Figure 1: Optimal Climate Policy

Figure 2: Carbon Price [$\$/tC$]

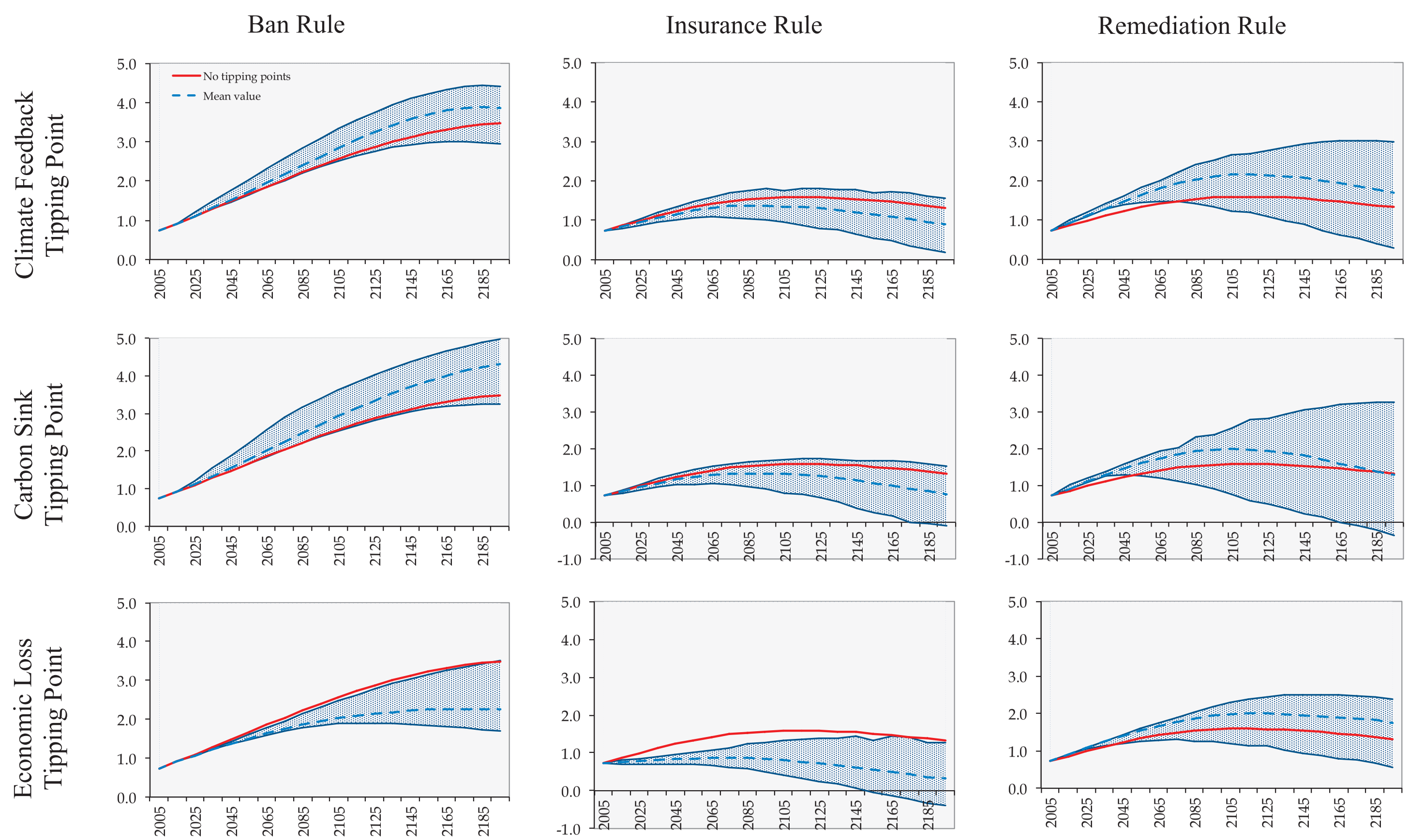


Figure 3: Temperature change [$°C$]

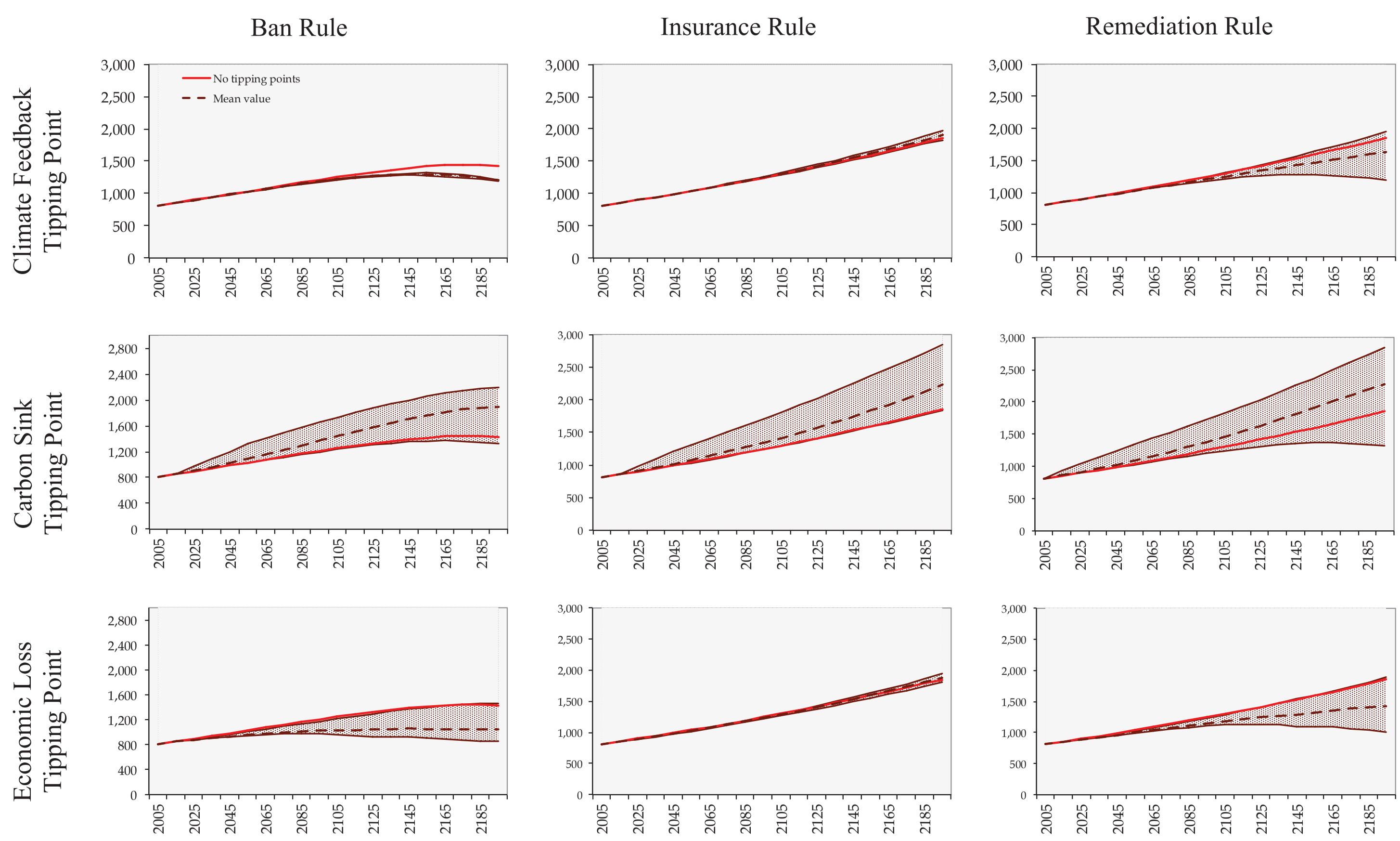


Figure 4: Atmospheric Carbon Concentrations [GtC]

References

Baker, E. and S. Solak

2011. Climate change and optimal energy technology r&d policy. *European Journal of Operational Research*, 213(2).

Barrett, S.

2008. The incredible economics of geoengineering. *Environmental and Resource Economics*, 39:45–54.

Bickel, J. and S. Agrawal

2013. Reexamining the economics of aerosol geoengineering. *Climatic Change*, 119(3-4).

Bickel, J. and L. Lane

2009. An analysis of climate engineering as a response to climate change. Copenhagen Consensus Center Report, 40.

Blackstock, J. J. and J. C. S. Long

2010. The Politics of Geoengineering. *Science*, 327(5965):527–.

Bony, S., G. Bellon, D. Klocke, S. Sherwood, S. Fermepin, and S. Denvil

2013. Robust direct effect of carbon dioxide on tropical circulation and regional precipitation. *Nature Geoscience*, 6(6):447–451.

Brander, L. M., K. Rehdanz, R. S. Tol, and P. J. Van Beukering

2012. The economic impact of ocean acidification on coral reefs. *Climate Change Economics*, 3(01).

Cai, Y., K. Judd, and T. Lontzek

2013. The social cost of stochastic and irreversible climate change. Working Paper 18704, National Bureau of Economic Research.

Crutzen, P. J.

2006. Albedo enhancement by stratospheric sulfur injections: A contribution to resolve a policy dilemma? *Climatic Change*, 77:211–219.

Emmerling, J. and M. Tavoni

2013. Geoengineering and abatement: a 'flat' relationship under uncertainty. Fondazione Eni Enrico Matei, Nota Di Lavoro 2013.031.

Goes, M., N. Tuana, and K. Keller

2011. The economics (or lack thereof) of aerosol geoengineering. *Climatic Change*, 109(3-4).

Gramstad, K. and S. Tjøotta

2010. Climate engineering: Cost benefit and beyond. University of Bergen, Department of Economics, No 05/10.

Heckendorn, P., D. Weisenstein, S. Fueglistaler, B. P. Luo, E. Rozanov, M. Schraner, L. W. Thomason, and T. Peter

2009. The impact of geoengineering aerosols on stratospheric temperature and ozone. *Environmental Research Letters*, 4(4):045108.

Heutel, G., J. Moreno-Cruz, and S. Shayegh

2015. Solar geoengineering, uncertainty, and the price of carbon. *NBER Working Paper 21355*.

Keith, D. W.

2000. Geoengineering the climate: History and prospect. *Annual Review of Energy and Environment*, 25:245–284.

Keith, D. W., E. Parson, and M. G. Morgan

2010. Research on global sun block needed now. *Nature*, 463:426–427.

Kolstad, C. D.

1996. Learning and stock effects in environmental regulation: the case of greenhouse gas emissions. *Journal of environmental economics and management*, 31(1).

Lemoine, D. and I. Rudik

2014. Steering the climate system: Using inertia to lower the cost of policy. Technical report, University of Arizona Department of Economics.

Lemoine, D. and C. Traeger

2014. Watch your step: Optimal policy in a tipping climate. *American Economic Journal: Economic Policy*, 6(1).

Lemoine, D. and C. Traeger

2015. Playing the climate dominoes: Tipping points and the cost of delaying policy. Working paper.

McClellan, J., D. W. Keith, and J. A. r

2012. Cost analysis of stratospheric albedo modification delivery systems. *Environmental Research Letters*, 7:034019.

Moreno-Cruz, J.

2015. Mitigation and the geoengineering threat. *Resource and Energy Economics*, 41:248–263.

Moreno-Cruz, J. and D. Keith

2013. Climate policy under uncertainty: a case for solar geoengineering. *Climatic Change*, 121:431–444.

Naevdal, E.

2003. Optimal regulation of natural resources in the presence of irreversible threshold effects. *Natural Resource Modeling*, 16(3):305–333.

Naevdal, 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.

Naevdal, 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.

Nordhaus, W. D.

1993. Optimal greenhouse-gas reductions and tax policy in the" dice" model. *The American Economic Review*, Pp. 313–317.

Nordhaus, W. D.

2008. *A question of balance: Weighing the options on global warming policies*. Yale University Press.

Pindyck, R. S.

2013. Climate change policy: What do the models tell us? *Journal of Economic Literature*, 51(3):860–872.

Pongratz, J., D. B. Lobell, L. Cao, and K. Caldeira

2012. Crop yields in a geoengineered climate. *Nature Climate Change*, 2(2):101–105.

Rasch, P., S. Tilmes, and R. Turco

2008. An overview of geoengineering of claimte using stratospheric sulphate aerosols. *Philosophical Transactions of the Royal Society A—Mathematical Physical and Engineering Sciences*, 366:4007–4037.

Shayegh, S. and V. M. Thomas

2015. Adaptive stochastic integrated assessment modeling of optimal greenhouse gas emission reductions. *Climatic Change*, 128(1-2):1–15.

Victor, D. G.

2008. On the regulation of geoengineering. *Oxford Review of Economic Policy*, 24:322–336.

Wigley, T.

2006. A combined mitigation/geoengineering approach to climate stabilization. *Science*, 314:452–454.

A The Stochastic DICE model with Tipping Points

In this appendix, we describe in more detail our modifications to the DICE model to incorporate SGE and CTP. We modify the DICE model that has been first introduced by Nordhaus (Nordhaus, 1993) and the model parameters and equations have been represented in (Nordhaus, 2008). We have modified the DICE 2007 version of the model in order to include a probability of the tipping points and the SGE action. We model the stochastic DICE as a Markov decision process with a state space, an action space, an information space, a transition function, and a reward function.

- **State Space**

  The global climate-economy system can be defined as a state with seven continuous variables: T_t^{at} is atmospheric temperature (degrees Celsius above preindustrial), T_t^{lo} is lower ocean temperature (degrees Celsius above preindustrial), M_t^{at} is the atmospheric concentration of carbon (Giga Tons of Carbon, GTC), M_t^{up} is the concentration in the biosphere and upper oceans (GTC), M_t^{lo} is the concentration in deep oceans (GTC), K_t is capital (\$trill), and F_t is radiative forcing (W/m^2). We define the state space as $S_t = \{T_t^{at}, T_t^{lo}, M_t^{at}, M_t^{up}, M_t^{lo}, K_t, F_t\}$.[18]

[18]And, as described in the text, when CTPs are added to the model, there is an eighth state variable representing whether or not the CTP has been reached.

- **Action Space**

  At each time step, a mitigation action (control rate) a_t and a SGE action g_t are taken which indicate the percentage reduction of GHG emissions and the percentage of radiative forcing reduction respectively. Both actions are costly and impose immediate costs to the economy but prevent the future damage costs of higher temperature due to the abated emissions or lowered radiative forcing. Taking actions a_t and g_t at any given state will determine the next state deterministically. Therefore the action space is defined as $a_t \in [0, 1]$ and $g_t > 0$

- **Information Space**

  We can introduce uncertainty into this system by modeling two types of uncertainty, one from the atmospheric temperature shocks and the other from the SGE damage dynamic. We define a Normal probability distribution for the temperature shocks and a truncated Lognormal distribution for the SGE damage coefficient. For the analysis here we assume $W_t = Normal(1, 0.0068)$ for the weather shocks and $\nu_t^g = Lognormal(\ln(0.03), 1)$ for the SGE damage cost function.

- **Transition Functions**

  The gross economic output, Y_t, is calculated from the given level of technology, capital, and labor in the current state:

$$Y_t = \Gamma_t \times K_t^{\beta} \times L_t^{1-\beta} \tag{A.1}$$

where Γ_t is technology and L_t is labor at time t. β is the output elasticity of capital. The net output, Q_t, is calculated after subtracting climate change damages and mitigation and SGE costs from the gross output:

$$Q_t = Y_t - \big(\Delta(T_t^{at}) + A(a_t) + G(g_t)\big) \times Y_t \tag{A.2}$$

$$\Delta(T_t^{at}) = \frac{(1-\mathbf{u_3})(1+\nu_t^g\ g_t^2)}{1+\xi_1(W_t \times (T_t^{at})^2) + \xi_2(M_t^{at} - M_0^{at})^2 + \xi_3(M_t^{up} - M_0^{up})^2} \tag{A.3}$$

$$A(a_t) = \theta_1 \times a_t^{\theta_2} \tag{A.4}$$

$$G(g_t) = \theta_1^g \times g_t^{\theta_2^g} \tag{A.5}$$

where Δ is the damage function that depends on the atmospheric temperature, atmospheric carbon concentration and upper ocean carbon concentration. The state of the world determines the value of $\mathbf{u_3}$: if the economic tipping point has not passed yet, $\mathbf{u_3} = 0$ and if the tipping point is passed $u_3 = 10\%$. The parameters ξ_1, ξ_2, and ξ_3 are the damage cost coefficients and are adjusted to replicate the damage cost of the original DICE model for the year 2005. The parameters θ_1 and θ_2 are the coefficients of the mitigation cost function $A(a_t)$ and θ_1^g and θ_2^g are the coefficients of SGE cost function.

Part of the net output at each time step is saved and invested back in the form of the capital and the rest is consumed:

$$K_{t+1} = (1-\delta) \times\ K_t + \theta_3 \times\ Q_t \tag{A.6}$$

where δ is the capital depreciation rate and θ_3 is the saving rate. The industrial emissions E_t are found from the carbon intensity of the output σ_t, taking into account the abatement decision:

$$E_t = \sigma_t \times (1 - a_t) \times \ Y_t \tag{A.7}$$

Other state variables in the next time epoch are found as:

$$M^{at}_{t+1} = E_t + (1 - \mathbf{u_2}) \times \ M^{at}_t + \phi_{21} \times \ M^{up}_t \tag{A.8}$$

$$M^{up}_{t+1} = u_2 \times \ M^{at}_t + \phi_{22} \times \ M^{up}_t + \phi_{32} \times \ M^{lo}_t \tag{A.9}$$

$$M^{lo}_{t+1} = \phi_{23} \times \ M^{up}_t + \phi_{33} \times \ M^{lo}_t \tag{A.10}$$

where $\phi_{21}, ..., \phi_{33}$ are carbon cycle transition coefficients. The parameter $\mathbf{u_2}$ indicates the carbon sink tipping point. When the tipping point is crossed it drops to half of its initial value.

The temperature equations for the next state are:

$$T^{at}_{t+1} = T^{at}_t + \eta_1 \times \left\{F_{t+1} - \eta_2 T^{at}_t - \eta_3 \times \ \{T^{at}_t - T^{lo}_t\}\right\} \tag{A.11}$$

$$T^{lo}_{t+1} = T^{lo}_t + \eta_4 \times \ \{T^{at}_t - T^{lo}_t\} \tag{A.12}$$

$$F_{t+1} = \eta_2 \mathbf{u_1} (\log_2^{(M^{at}_{t+1} - M^{at}_0)}) \times (1 - g_t) \tag{A.13}$$

where $\eta_1, ..., \eta_4$ are temperature coefficients and $\mathbf{u_1}$ is the tipping point indicator for the climate sensitivity. If the tipping point is crossed, $\mathbf{u_1}$ will go up from $3°C$ to $5°C$.

- **Reward Function**

  The reward is calculated as the social utility of consumption at each time epoch:

$$U_t = \frac{\left\{(1-\theta_3) \times \ Q_t\right\}^{1-\alpha}}{1-\alpha} \tag{A.14}$$

  where α is the elasticity of marginal utility of consumption. The objective is to maximize the sum of discounted expected social utilities over the modeling horizon given uncertainty in climate sensitivity:

$$\max_{a_t \in A(S_t)} \mathbb{E}\left\{ \sum_{t=0}^{T} \gamma^t \ U_t\big(S_t, a_t, W_t\big) \right\} \tag{A.15}$$

- **Look-ahead approximation heuristic**

  To demonstrate the algorithm, consider a simple transition between two states S_t to S_{t+1} as shown in Figure A.1. The uncertainty from the weather shocks and SGE damage cost is shown as ΔT. After observing this information and taking the action a_t at state S_t, we will be able to calculate the next state S_{t+1} . In order to find the optimal action a^* we deploy our two-step-ahead algorithm. First, under the deterministic assumption, the value of the current state S_t will be calculated by taking any candidate action a_t and two consecutive null actions to obtain two post-decision states S_t^a and S_{t+1}^0 and with immediate rewards of $U_t^{a_t}$, U_{t+1}^0, and U_{t+2}^0. The post-decision state variable S_t^a is a transient state between the current state S_t and the next state S_{t+1}. This state is generated by implementing the chosen action a_t on the current state S_t but before realization of the random parameter ΔT_{t+1}. The optimal action is the one that maximizes the value of the current sate:

$$a_t^*(S_t) = argmax_{a_t}\big(U_t^{a_t} + \bar{V}_t(S_t^a)\big) \tag{A.16}$$

The value function $\bar{V}_t$ is the approximation of the post-decision state S_t^a. For this problem we consider a very simple function approximation with only one parameter $\bar{V}_t(S_t^a) = \theta_1 \times U_{t+2}^0$, where θ_1 is the tunable parameter of the value function approximation and defines the "policy". The initial value of this parameter is assumed to be one and it is updated at the end of each iteration. The value of state is calculated from $\hat{V}_t(S_t) = \max_{a_t}\big(U_t^{a_t} + \bar{V}_t(S_t^a)\big)$. Once the optimal action a^* is found, a realization of the uncertain parameter is drawn from the sample path and the values of state variables of the next state S_{t+1} is calculated accordingly. The approximate value is used to update the approximation function that was used to estimate the value of the post-decision state S_t^a using the following stochastic gradient algorithm: $\theta_1^{new} = \theta_1^{old} - \alpha \times (\bar{V}_t - \hat{V}_{t+1}) \times U_{t+2}^0$ The step size α is chosen as $[U_{t+2}^0]^{-2}$ to simplify the updating equation and guarantees the convergence. Therefore the new coefficient for the next iteration is calculated as

$$\theta_1^{new} = \theta_1^{old} - \frac{\bar{V}_1 - \hat{V}_2}{U_1(S_2^0, 0)} \tag{A.17}$$

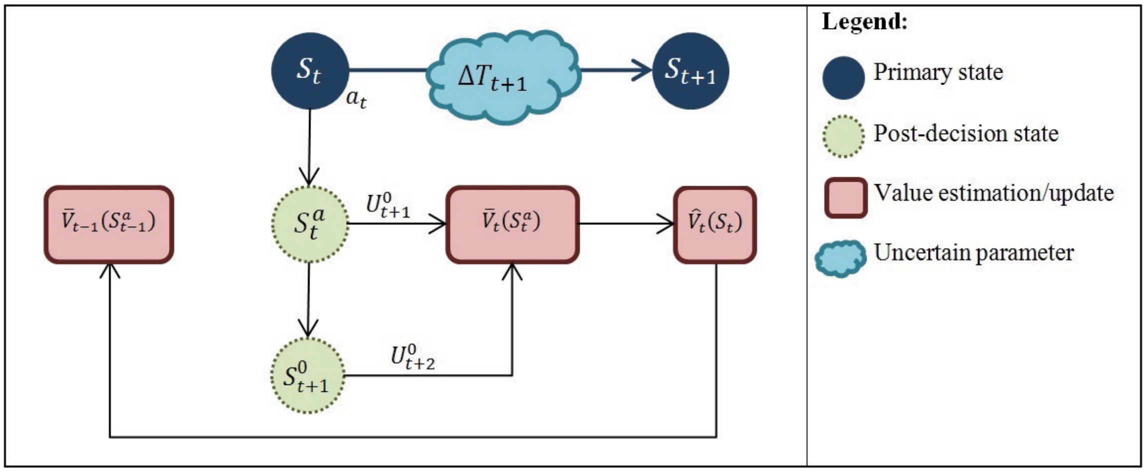


Figure A.1: An example of the two-step-ahead algorithm for for the DICE model.

B Comparing our numerical solution to Lemoine and Traeger (2014)

In this appendix, we verify the two-step-ahead solution algorithm from Shayegh and Thomas (2015) by using it to solve the model in Lemoine and Traeger (2014). We demonstrate that our results are identical to theirs, which was solved by them using an alternative solution algorithm. In that study, for numeric efficiency, they reformulated the DICE-2007 model to use effective labor units for capital and combined biosphere and shallow ocean stock for carbon dynamics. Moreover, they downscale the original decadal time steps in the DICE model to an annual step size. Their solution is based on approximating the value function using a 10^4 basis of Chebychev polynomials.

Compared to Lemoine and Traeger (2014)'s solution method, our method is significantly simpler, is faster to converge, and uses only one tunable parameter for each approximation. We keep the original structure of the DICE model and use the three carbon circulation layers (atmosphere, upper ocean, and lower ocean). Furthermore, we keep the decadal structure of the DICE model and show that the results follow Lemoine and Traeger (2014)'s very closely.

We consider the two types of CTPs modeled by Lemoine and Traeger (2014) for this comparison. The first, the Climate Sensitivity tipping point, increases the climate sensitivity parameter and therefore amplifies global warming. We consider three levels of increased climate sensitivity and model them separately. The second tipping point, the Carbon Sink, increases the lifetime of CO_2 in the atmosphere by reducing the fraction of atmospheric emissions that is transferred to the upper ocean layer at each time step. We reduce this fraction by either 25%, 50%, and 75%. The Climate Sensitivity tipping point changes the effect of emissions on temperature, and the Carbon Sink tipping point changes the timing of such an effect. The results are shown in Figures B.1 and B.2, for the case in which the CTP happens to never be crossed, in order to see how the modeled policymaker adjusts to the possibility over time.

Figure B.1 shows the carbon concentrations, and Figure B.2 shows temperature, both comparing the results of our model that has decadal time steps and a two-layer ocean with the results from Lemoine and Traeger (2014).[19] The results are nearly identical.[20] This verifies that the solution method from Shayegh and Thomas (2015)

[19]The figures from Lemoine and Traeger (2014) are cut and pasted directly from their paper. Note that all the results in this section are without SGE and without the Economic Loss CTP, since we are solely concerned in replicating the model in Lemoine and Traeger (2014).

[20]Our model generates slightly higher temperature that can be attributed to the longer time steps.

is appropriate to use in this context.

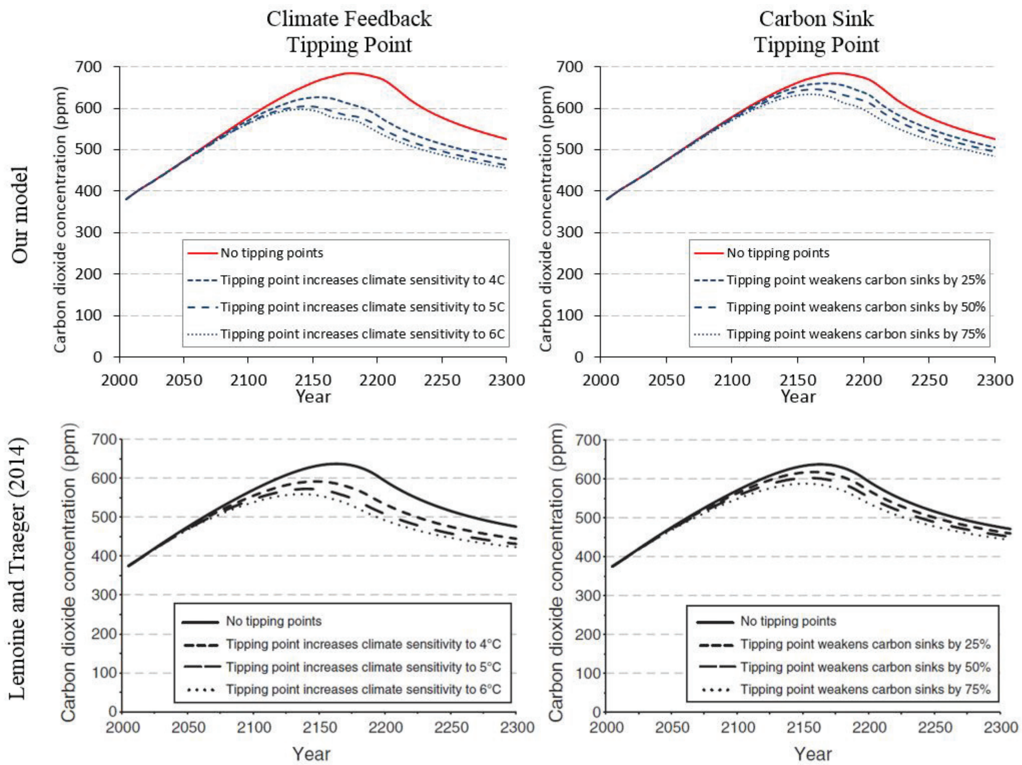


Figure B.1: Comparison of carbon concentration results in our model with the Lemoine and Traeger's model. We replicated the tipping points definition from their model but used the original DICE model's time step and carbon circulation structure.

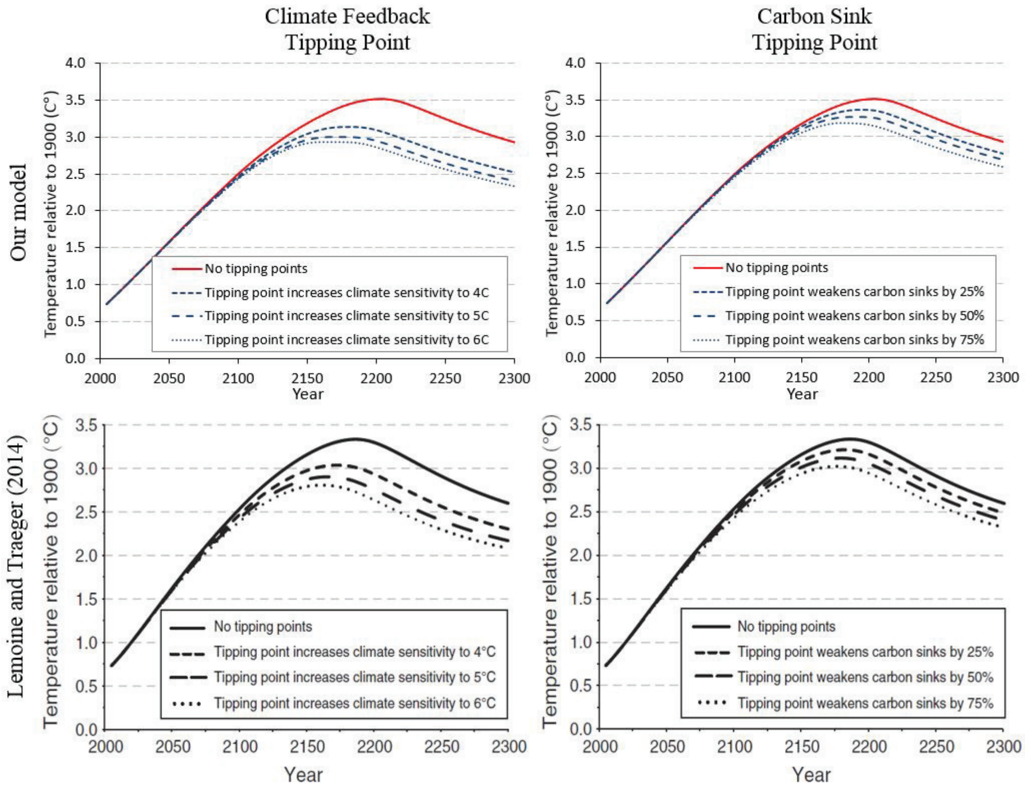


Figure B.2: Comparison of temperature results in our model with the Lemoine and Traeger's model. We replicated the tipping points definition from their model but used the original DICE model's time step and carbon circulation structure.

C Alternative Theoretical Model

In this appendix we present an alternative theoretical model that employs more simplifying assumptions than the model presented in the paper. In particular, this alternative model considers only one type of CTP (a fixed drop in utility), includes a case where the threshold of the CTP is known with certainty (though also includes a case in which the threshold location is uncertain), and models SGE and mitigation as perfect substitutes. The benefit of these simplifying assumptions is that the model provides some additional intuition and closed-form solutions, which are relevant to our numerical simulations. In particular, with this simpler model we can formally prove the predictions about how outcomes differ under the different SGE rules, while in the model in the main text these predictions are just hypotheses. The structure of this model is substantially different than that of the model in the paper; in particular, the model in the paper uses a dynamic programming approach. Thus, the paper's model is not merely an extension of the simpler model, so we present this model in an appendix.

First, we present the case where the threshold location is known with certainty. Then, we present the case of an uncertain threshold location.

C.1 Known Threshold Location

The dynamics governing the change in temperature are represented by the following first-order differential equation:

$$\dot{T} \equiv \frac{dT}{dt} = S^0 - m - \theta g - \delta T \tag{C.1}$$

where T represents temperature, m and g are mitigation and SGE, respectively; both measured in terms of their potential to reduce temperature. We assume society chooses whether or not to make SGE available to the regulator. And the regulator takes this choice as given. The parameter $\theta \in \{0, 1\}$ represents the rule regarding the availability of SGE: $\theta = 1$ when SGE is available and $\theta = 0$ when it is not. S^0 is the radiative forcing caused by the unabated concentration of greenhouse gases. Some fraction of the heat stored in the atmosphere escapes; this effect is captured by the term δT where δ is the heat transfer parameter (Naevdal and Oppenheimer (2007)). All variables are a function of time, but we omit reference to the time variable to avoid notation clutter.

The costs of mitigation are a strictly increasing and strictly convex function of m denoted by $C_m(m)$ with $C_m(0) = 0$. SGE costs are a strictly increasing and strictly convex function denoted by $C_g(g)$ with $C_g(0) = 0$. For simplicity, and in order to explore the model further, we assume quadratic costs for both mitigation and SGE; that is

$$C_m(m) = \frac{1}{2}\beta m^2 \quad \text{and} \quad C_g(g) = \frac{1}{2}\gamma g^2. \tag{C.2}$$

We model a climate tipping point (CTP) as an irreversible loss in welfare that occurs after a given temperature threshold $\bar{T}$ is crossed. In particular, before the threshold is crossed society receives a constant stream of utility $\alpha = A$. When the threshold is crossed, α jumps to zero. A regulator chooses the level of mitigation and SGE that maximizes welfare. This problem can be stated as

$$W(t) = \max_{\{m,g\}} \int_0^\infty \left[\alpha - C_m(m) - C_g(g)\right] e^{-rt} dt \tag{C.3}$$

subject to:

$$\dot{T} = S^0 - m - \theta g - \delta T,\ T(0) = 0 \tag{C.4}$$

$$\dot{\alpha} = 0,\ \alpha(0) = A \tag{C.5}$$

$$\alpha(\bar{t}^+) - \alpha(\bar{t}^-) = -\alpha(\bar{t}^-) \tag{C.6}$$

Equation (C.6) describes how, at the time when the CTP is reached ($\bar{t}$), the utility α drops to zero. At all other periods it is constant (equation C.5).

In principle, if the costs of avoiding crossing the threshold are too high, the regulator could allow it to be crossed. We assume that the utility loss from crossing the CTP is high enough so that it is never optimal to cross, which we ensure by setting $A > C_m(S^0)$. Moreover, in order for the problem to be economically interesting, we assume that the threshold is below the steady-state temperature level in the absence of regulation, that is $\bar{T} < \frac{S^0}{\delta}$. Under these two assumptions, the threshold will never be crossed but will be reached at some finite time $t = \bar{t}$. When the location of the threshold is known, the regulator can optimally choose $\bar{t}$. That is, without uncertainty, the optimal policy is to just reach the threshold temperature $\bar{T}$ but not exceed it. The utility loss from the CTP never occurs, and after reaching the threshold temperature, the temperature will be maintained at that level.

The problem is solved backwards in two stages. In the second stage, we find the solution to the problem after the threshold is reached. In the first stage, we analyze the problem starting at $t = 0$, using the welfare of the second stage as a scrap value function that depends only on the reaching time $\bar{t}$.

C.1.1 After Reaching the Threshold

After the threshold $\bar{T}$ is reached, the temperature must be kept at that constant level to avoid triggering the CTP: $T = \bar{T}$ and $\dot{T} = 0$. Thus, for all $t > \bar{t}$, the problem for the regulator is:

$$W(t|t > \bar{t}) = \max_{\{m,g\}} \int_{\bar{t}}^{\infty} [A - C_m(m) - C_g(g)]\, e^{-rt} dt \tag{C.7}$$

subject to:

$$0 = S^0 - m - \theta g - \delta \bar{T} \tag{C.8}$$

The optimality condition is given by:

$$\theta C'_m(m) = C'_g(g) \tag{C.9}$$

After reaching the threshold, mitigation and SGE will be kept at constant levels, which we denote $\bar{m}$ and $\bar{g}$. Equation (C.9) shows that once the threshold is reached, the optimal policy is such that the marginal costs of mitigation and SGE are equalized. Combining equations (C.8) and (C.9), the optimal values of $\bar{m}$ and $\bar{g}$ can be calculated. Using the functional forms defined above, the optimal policy after the threshold is crossed is given by:

$$\bar{m} = \frac{\gamma}{\gamma + \theta^2\beta}[S^0 - \delta\bar{T}] \text{ and } \bar{g} = \frac{\theta\beta}{\gamma + \theta^2\beta}[S^0 - \delta\bar{T}] \tag{C.10}$$

Replacing equation (C.10) back into equation (C.7) yields:

$$W(t|t > \bar{t}) = \frac{1}{r}\left[A - \frac{1}{2}\frac{\beta\gamma}{\gamma + \theta^2\beta}\left[S^0 - \delta\bar{T}\right]^2\right]e^{-r\bar{t}} = \bar{W}e^{-r\bar{t}} \tag{C.11}$$

which is constant over time and is a function of $\bar{t}$ and parameters only.

C.1.2 Before Reaching the Threshold

Before reaching the threshold, the Hamiltonian of the problem is given by:

$$\mathscr{H} = A - C_m(m) - C_g(g) + p[S^0 - m - \theta g - \delta T] \tag{C.12}$$

where p is the co-state variable associated with the increase in temperature. Applying the maximum principle, the optimality conditions are given by:

$$C'_m(m) + p = 0 \tag{C.13}$$

$$C'_g(g) + \theta p = 0 \tag{C.14}$$

$$\dot{p} = (r + \delta)p \tag{C.15}$$

As usual, p has the interpretation of the social cost of a marginal increase in temperature, and it is equated to the marginal cost of mitigation and the marginal cost of SGE, as in equations (C.13) and (C.14). The optimal policy ensures that the marginal costs of all climate intervention technologies are equalized. Equation (C.15) can be interpreted as an arbitrage rule for investing in reducing temperature. By investing today, and paying a price p, society saves in costs of intervention in the future; thus reducing $\dot{p}$. Directly from equation (C.15) the solution for p is given by:

$$p(t) = \kappa e^{(r+\delta)t} \tag{C.16}$$

where p must be negative because the increase in temperature decreases welfare; hence $\kappa < 0$. Using the functional forms defined above the optimal levels of mitigation and SGE are given by:

$$m(t) = -\frac{1}{\beta}\kappa e^{(r+\delta)t}, \text{ and } g(t) = -\frac{\theta}{\gamma}\kappa e^{(r+\delta)t} \tag{C.17}$$

where $\kappa = \frac{\beta\gamma}{\gamma+\theta^2\beta} f(\bar{t})$, and $f(\bar{t}) = -\frac{r+2\delta}{\delta}\frac{S^0-(S^0-\delta\bar{T})e^{\delta\bar{t}}}{e^{(r+2\delta)\bar{t}}-1}$

As we mentioned above, when the location of the threshold is known, the regulator can choose the optimal time to reach the threshold. Specifically, the optimal reaching time satisfies the following condition:

$$\mathscr{H}(\bar{t}) = -\frac{\partial W(t|t>\bar{t})}{\partial \bar{t}} e^{r\bar{t}}. \tag{C.18}$$

Replacing the values from equations (C.16) and (C.17) back into equation (C.12), and evaluating at $t = \bar{t}$ yields:

$$\mathscr{H}(\bar{t}) = A + \frac{1}{2}\left[\frac{\gamma + \theta^2\beta}{\beta\gamma}\right]\kappa^2 e^{2(r+\delta)\bar{t}} + \kappa e^{(r+\delta)\bar{t}}[S^0 - \delta\bar{T}] \tag{C.19}$$

and using equation (C.11), the righthand side of equation (C.18) is given by:

$$-\frac{\partial W(t|t > \bar{t})}{\partial \bar{t}} e^{r\bar{t}} = A - \frac{1}{2}\left[\frac{\beta\gamma}{\gamma + \theta^2\beta}\right][S^0 - \delta\bar{T}]^2 \tag{C.20}$$

The optimal policy solution, that is, the values of κ (which gives $m(t)$ and $g(t)$) and $\bar{t}$, depends on the rule regarding the availability of SGE. We turn to study the different rules next.

C.1.3 Comparing SGE Rules

The regulator chooses the optimal levels of mitigation and SGE subject to one of the three rules regarding SGE availability:

(a) Ban: SGE is never allowed, that is, $\theta = 0$ for all $t > 0$.

(b) Insurance: SGE is always allowed , that is, $\theta = 1$ for $t > 0$.

(c) Remediation: SGE is allowed only after the threshold has been reached, that is, $\theta = 0$ for $t < \bar{t}$ and $\theta = 1$ for $t \geq \bar{t}$.

Each rule can be compared under different criteria. Here, we compare them in terms of optimal reaching time, temperature level, and welfare. In the analysis that follows, we denote the different optimal policies using the superscripts *ban*, *ins*, and *rem*. We start by comparing optimal reaching times for the different policies, then we analyze them in terms of temperature changes and welfare levels.

Proposition 1: *The optimal reaching times under the Ban and the Insurance rules are the same. Under the Remediation rule, the threshold is reached sooner, compared to the other two policies:* $\bar{t}^{rem} < \bar{t}^{ban} = \bar{t}^{ins}$.

Proof: We use Figure C.1 to help with the proof. The optimal reaching time can be shown to occur when:

$$S^0 - \delta\bar{T} = -Lf(\bar{t})e^{(r+\delta)\bar{t}} \tag{C.21}$$

where $L = 1$ for *ban* and *ins*, and $L = \frac{\gamma+\beta}{\gamma}\left[1 + \sqrt{\frac{\beta}{\beta+\gamma}}\right] > 1$ for *rem*. When $L > 1$ the slope of the right hand side of the equality defined in equation (C.21) is steeper, which implies a lower reaching time.

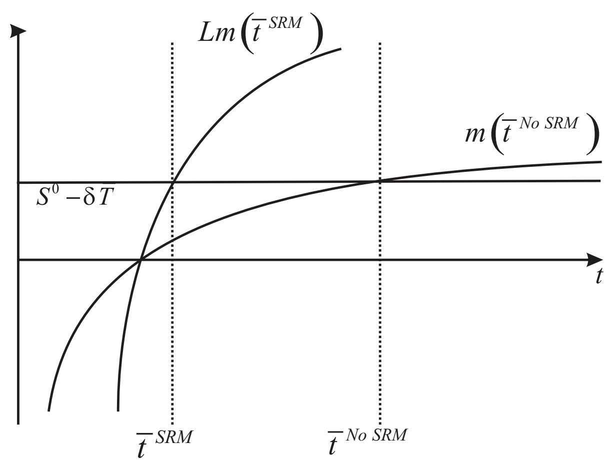


Figure C.1: Optimal reaching time.

Proposition 1 establishes that the threshold is reached at the same time in the *ban* and *ins* rules, that is when SGE is not available or when SGE is always available. However, when SGE is only available after the threshold has been reached (*rem*), there is an incentive to reach the threshold faster. The reason behind this result is that it is too costly to deal with the problem with only one instrument; hence, the regulator finds it optimal to allow a faster approach to the threshold in order to be able to use SGE. This result, however, does not imply higher levels of climate intervention once the threshold is reached.

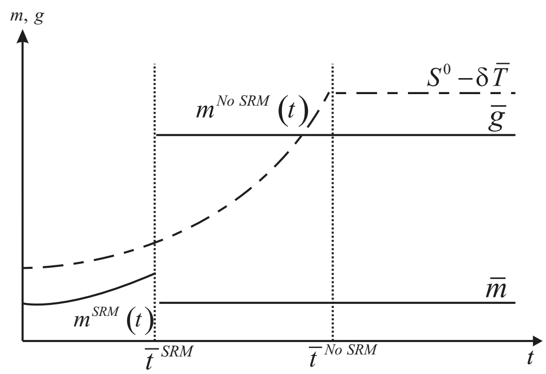
Proposition 2:

i.) $m^{ban}(\bar{t}^{ban-}) = \bar{m}^{ban}$, $m^{ins}(\bar{t}^{ins-}) = \bar{m}^{ins}$, $g^{ins}(\bar{t}^{ins-}) = \bar{g}^{ins}$, *and* $m^{rem}(\bar{t}^{rem-}) < \bar{m}^{rem}$. *For the rem rule there is a jump at time* $t = \bar{t}$.

ii.) For $t \leq \bar{t}^{rem}$*:* $m^{ban}(t) = m^{ins}(t) + g^{ins}(t) > m^{rem}(t)$*. For* $\bar{t}^{rem} < t < \bar{t}^{ban}$*:* $m^{rem}(t) > m^{ins}(t) + g^{ins}(t) = m^{ban}(t)$*. For* $t > \bar{t}^{ban}$*:* $m^{rem}(t) = m^{ins}(t) + g^{ins}(t) = m^{ban}(t)$

Proposition 2 shows how the intervention level varies under the different SGE rules. Figure C.2 presents this result graphically. From Proposition 1.i, if the reaching time is the same for *ban* and *ins*, it must be true that mitigation in *ban* is equal to the total amount of intervention in *ins* (see Figure C.2). Also, directly from Proposition 1.ii, if the system reaches the threshold fastest under *rem*, then a lower level of mitigation must be implemented under this rule. Once the threshold is reached, however, the amount of intervention increases to $\bar{m} + \bar{g}$ (see Figure C.2).

The change in the reaching time $\bar{t}$ for the different rules has implications in terms of temperature levels and welfare over time. From Propositions 1 and 2, it is difficult



(A)

Figure C.2: Comparing optimal policies. Climate intervention outcomes under the Ban, Insurance and Remediation rules.

to predict what are temperature and welfare under the different rules. In particular, we should expect temperature to be the same under rules *ban* and *ins*, but welfare can differ. Also, under rule *rem*, temperature can be lower given the earlier jump to a higher level of climate intervention. The next proposition shows results regarding temperature and welfare.

Proposition 3:

i.) For $t < \bar{t}^{ban}$, $T^{ban} = T^{ins} < T^{rem}$. *For* $t > \bar{t}^{ban}$, $T^{ban} = T^{ins} = T^{rem} = \bar{T}$.

ii.) For $t < \bar{t}^{rem}$, $W^{ban} < W^{ins} < W^{rem}$. *For* $\bar{t}^{rem} < t < \bar{t}^{ban}$, $W^{ban} < W^{rem} < W^{ins}$. *For* $t > \bar{t}^{ban}$, $W^{ban} < W^{ins} = W^{rem}$

The results in Proposition 3 are illustrated in Figure C.3. Proposition 3.i follows directly from Proposition 2. (see Panel A in Figure C.3). The lower level of mitigation under *rem* for $t < \bar{t}^{rem}$ implies higher temperature relative to *ban* and *ins*. Proposition 3.ii says that welfare under *ban* is lower than *ins*. Given the assumption of increasing and convex costs of mitigation, and from Proposition 2.i, it follows that it is cheaper to deal with climate change using two instruments. Thus, while the Insurance and the Ban rules are no different in terms of reaching time and temperature levels, the Insurance rule is better in terms of welfare. The Remediation rule also shows higher welfare levels, compared to the Ban. However, it is not clear whether the welfare is higher under Insurance or Remediation. This result follows from the lower level of mitigation as well as the fact that there are not direct costs associated to temperature. That is, temperature is a bad only in terms of crossing the threshold. However, once the threshold is reached, society must deal with higher costs of intervention to keep temperature constant. These costs are larger relative to the Insurance rule. Thus, whether the discounted welfare is higher (lower) under the Insurance rule relative to the Remediation depends on whether area X is smaller (larger) than area Y. We will explore this aspect with our numerical simulation.

Under the assumption of a known CTP temperature, allowing SGE technologies is always better than not allowing for them. Contrary to what has been proposed in the literature before Victor (2008) Keith et al. (2010), the Remediation rule yields higher temperature, lower welfare, and a faster approach to the threshold relative to the other two rules.

As we said above, it is unlikely policy makers will know the tipping temperature. In the next section we introduce uncertainty and analyze the optimal mix of mitigation and SGE and the ranking of the different SGE rules.

C.2 Unknown Threshold Location

The exact temperature threshold leading to a CTP is likely to be unknown to the regulator. When the location of the threshold is unknown, the procedure to derive the optimal policy must be modified to include this new risk. The specific characteristics of this threshold problem makes it suitable to the use of piecewise deterministic control techniques.

In the deterministic case the threshold ends up being reached but never crossed. In the unknown location case, however, the threshold may end up being crossed. When the threshold is crossed, we assume that SGE is fast enough to maintain temperatures at the threshold level. After the threshold has been reached or crossed, the problem becomes deterministic, since the only source of uncertainty is the location of the threshold. The residual value after the threshold is reached is calculated using standard deterministic optimal control techniques.

Under the Ban rule, once the threshold is reached, there will be no effective climate intervention policy available because of the inertia associated with the climate system and the ineffectiveness of mitigation to quickly reduce temperatures. The CTP utility loss will occur, mitigation will be zero, and the temperature will reach its maximum level. Under the Insurance and Remediation rules, once the threshold is reached, SGE and mitigation will be employed to maintain temperature at $\bar{T}$ to avoid the utility loss.

Following Naevdal (2003, 2006), we assume the location of the threshold $\bar{T}$ is distributed according to the function $h(\bar{T})$ with support $[0, \infty]$. If a value $T > 0$ is attained without reaching the threshold, then it must be true that $\bar{T} > T$. Thus, as temperature increases, the regulator updates her beliefs about the threshold temperature. The updated probability distribution is given by $\phi(T(t)) = h(\bar{T}) / \int_T^\infty h(\bar{T}) d\bar{T}$ over the range $[T, \infty]$. We can then transform the distribution in state-space to a distribution over time using the following function:

$$\psi(t) = \begin{cases} \dot{T} & \text{if } \dot{T} \geq 0 \\ 0 & \text{if } \dot{T} < 0. \end{cases} \tag{C.22}$$

The previous function captures the idea that only new values of T remain risky and provide new information. Whenever a value of T has already taken place there is no longer risk. Introducing this function, the hazard rate of the occurrence of the CTP is given by:

$$\lambda(T(t), t) = \phi(T(t))\psi(t). \tag{C.23}$$

Without further loss of generality, we assume that $h(\bar{T})$ is exponential so the function $\phi(T(t)) = \phi_0$. This implies a Poisson arrival rate for $\bar{t}$. This process in time is different from a traditional Poisson process because the regulator can reduce the probability of an event happening by simply reducing $\dot{T}$. By implementing mitigation and SGE, the regulator reduces the risk of crossing the threshold. In particular, whenever $\psi(t) = 0$ the probability of crossing the threshold is zero.

The problem for the regulator is given by:

$$W(t) = \max_{\{m,g\}} E\left[\int_0^\infty \left[\alpha - C_m(m) - C_g(g)\right] e^{-rt} dt\right] \tag{C.24}$$

subject to:

$$\dot{T} = S^0 - m - \theta G - \delta T,\ T(0) = T_0 \tag{C.25}$$

$$\dot{\alpha} = 0,\ \alpha(0) = A \tag{C.26}$$

The jump in the state variable that occurs at time $\bar{t}$ depends on whether or not SGE is available at the moment the threshold is reached. If SGE is not available at $t = \bar{t}$ (under the Ban rule), then

$$\alpha(\bar{t}^{+}) - \alpha(\bar{t}^{-}) = -\alpha(\bar{t}^{-}) \tag{C.27}$$

If SGE is available (under the Insurance or Remediation rule), the loss of utility associated to crossing the tipping point can be avoided, but at a cost. In particular,

$$\alpha(\bar{t}^{+}) - \alpha(\bar{t}^{-}) = -\bar{A} \tag{C.28}$$

where $\bar{A} = \frac{1}{2}\frac{\beta\gamma}{\gamma+\beta}[S^0 - \Delta T]^2$ are the costs associated with $\bar{m}$ and $barg$ that keep temperature at the threshold level. Finally, $\bar{t}$ follows a Poisson process with hazard rate given by equation (C.23).

Following Naevdal (2003, 2006) and Naevdal and Openheimer (2007), the risk-augmented Hamiltonian is given by:

$$\mathscr{H} = \alpha - C_m(m) - C_g(g) + p(e^0 - m - \theta g - \delta T) + \lambda(T, m, g)\,[W(t|t > \bar{t}) - z] \tag{C.29}$$

$W(t|t > \bar{t})$ is the value of the objective function after the threshold is crossed.

Unlike in the deterministic case, here where the threshold is uncertain it may end up being crossed. Under the Ban rule, since SGE is not available, once the threshold is crossed temperature will be allowed to reach its maximum level by setting $m = 0$. Hence, temperature will stabilize at $T^0 = S^0/\delta$, α jumps to zero, and $W(t|t > \bar{t}) = 0$. Under the other two rules, SGE is available after reaching the CTP. We maintain the assumption that damages from crossing the CTP are sufficiently high, so SGE will be used after reaching the CTP to keep temperature fixed at $\bar{T}$ and avoid triggering the CTP utility loss. Thus, α jumps from A to $A - \bar{A}$ and $W(t|t > \bar{t}) = \int_{\bar{t}}^{\infty}(A - \bar{A})e^{-r(t-\bar{t})} = \bar{W}$ where $\bar{W}$ was defined in equation (C.11).

z is the value of the remaining welfare starting from any time t, given by

$$z(t) = e^{rt}E\left[\int_t^{\infty} \left[\alpha - C_m(m) - C_g(g)\right] e^{-rs}ds\right] \tag{C.30}$$

The differential equation governing the evolution of z is:

$$\dot{z} = rz - \left[\alpha - C_m(m) - C_g(g)\right] - \lambda(T, m, g)\left[\theta\bar{W} - z\right] \tag{C.31}$$

Therefore, the term $\lambda(T, m, g)\left[\theta\bar{W} - z\right]$ represents the expected loss of welfare that society would suffer if the threshold is crossed at any given time t.

Having defined z and $W(t|t > \bar{t})$, the maximum principle can be applied to equation (C.29) to obtain the following optimality conditions:

$$-C_m'(m) - p + \lambda_m'(T, m, g)\left[\theta\bar{W} - z\right] = 0 \tag{C.32}$$

$$-C_g'(g) - p + \lambda_g'(T, m, g)\left[\theta\bar{W} - z\right] = 0 \tag{C.33}$$

$$\dot{p} = (r + \delta + \lambda(T, m, g))p - \lambda_T'(T, m, g)\left[\theta\bar{W} - z\right] \tag{C.34}$$

The interpretation of equations (C.32) and (C.33) are very similar to the deterministic case. The marginal costs of mitigation and SGE are equalized to the marginal benefits of a reduction in temperature. However, in this case, the benefits have two components. First, a marginal increase in the level of mitigation (or SGE) directly reduces the temperature level, this effect is captured by p. Second, a marginal increase in mitigation (or SGE) also reduces the probability of the temperature threshold being crossed. This second component is captured by $\lambda'(T, m, g)$ which is a decreasing function of m and g.

Equation (C.34) captures the benefits from implementing SGE sooner. As in the deterministic case, the first term captures the direct effect, which is equivalent to the discounted reduction in future mitigation and SGE implementations. In this case, however, the hazard rate is part of the discounting term. Once the threshold is crossed, the benefits from mitigation are zero, thus the present value of climate intervention is discounted harder given the possibility of them not being useful after the threshold has been crossed. The second term shows the direct benefits in terms of a reduction in the probability of crossing the threshold. Increasing mitigation and SGE reduce temperature and with it the risks of reaching a CTP.

Using the functional forms defined above, the optimal policy is the solution to the following set of equations:

$$m = -\frac{p}{\beta} - \frac{\phi^0}{\beta}\left[\theta\bar{W} - z\right] \tag{C.35}$$

$$g = -\frac{p}{\gamma} - \frac{\theta\phi^0}{\gamma}\left[\theta\bar{W} - z\right] \tag{C.36}$$

$$\dot{p} = (r+\delta)p + \delta\phi^0\left[\theta\bar{W} - z\right] + \phi^0(S^0 - m - \theta g - \delta T)p \tag{C.37}$$

$$\dot{z} = rz - \alpha + \frac{1}{2}\beta m^2 + \frac{1}{2}\gamma g^2 - \phi^0(S^0 - m - \theta g - \delta T)\left[\theta\bar{W} - z\right] \tag{C.38}$$

$$\dot{T} = S^0 - m - \theta g - \delta T \tag{C.39}$$

C.2.1 Comparing SGE Rules

From equations (C.35)-(C.39) and evaluating at $\dot{T} = 0$, $\dot{z} = 0$, and $\dot{p} = 0$ the steady state equilibrium levels of emissions, SGE, and temperature are given by:

$$m^* = \sqrt{\left[\frac{\gamma[r+\delta]}{\phi^0[\gamma+\theta^2\beta]}\right]^2 + \frac{2\gamma\theta\bar{W}}{\beta[\gamma+\theta^2\beta]}} - \frac{\gamma[r+\delta]}{\phi^0[\gamma+\theta^2\beta]} > 0 \tag{C.40}$$

$$g^* = \theta\left[\sqrt{\left[\frac{\beta[r+\delta]}{\phi^0[\gamma+\theta^2\beta]}\right]^2 + \frac{2\beta\theta\bar{W}}{\gamma[\gamma+\theta^2\beta]}} - \frac{\beta[r+\delta]}{\phi^0[\gamma+\theta^2\beta]}\right] \geq 0 \tag{C.41}$$

$$T^* = \frac{S^0}{\delta} + \frac{1}{\delta}\left[\frac{[r+\delta]}{\phi^0} - \sqrt{\left[\frac{[r+\delta]}{\phi^0}\right]^2 + \frac{2[\gamma+\theta^2\beta]\theta\bar{W}}{\gamma\beta}}\right] \leq \frac{S^0}{\delta} \tag{C.42}$$

where the equality in (C.42) hold when SGE is banned, $\theta = 0$. The steady state behavior of the system is captured in the following proposition:

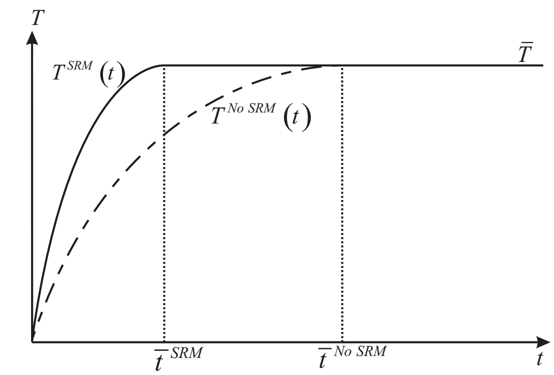
Proposition 4:

i.) The steady state levels of climate intervention under the three different SGE rules are such that $m^{*ban} > m^{*ins} + g^{*ins} > m^{*rem}$.

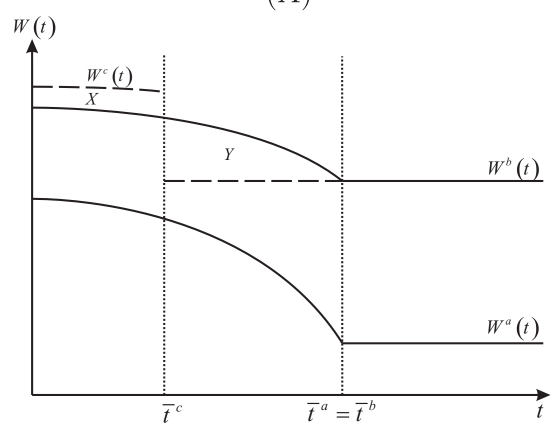
ii.) The steady state levels of temperature under the three different SGE rules are such that $T^{*ban} < T^{*ins} < T^{*rem}$.

Proposition 4 establishes that under the Ban rule, there will be more mitigation than the combined levels of intervention under the Insurance rule, which in turn exhibits higher intervention levels than the Remediation rule. Compared to the deterministic case, the main difference is that here the Ban and Insurance rules show different intervention levels. The reason is clear: without SGE the risk of crossing the threshold makes the regulator more cautious, and in order to avoid the loss of utility A, the regulator decides to pay a "risk premium" in terms of higher mitigation costs. The risk premium payment disappears in the deterministic case because the regulator can decide when to reach the threshold, which allows her to optimally choose mitigation levels to ensure the threshold is not crossed. The ranking of climate intervention levels transfer directly in terms of temperature. The Ban rule exhibits the lowest temperature in steady state reflecting the precautionary behavior induced by the inability of mitigation to quickly reduce temperatures. This result,however, follows from our assumption that climate damages are only represent by the loos of utility associated with the CTP. When we allow for damages to be a continuous function of temperature and carbon, we should observe more climate intervention overall, implying the Insurance and Remediation rules will exhibit lower temperature levels in steady state.

Although the risk-augmented Hamiltonian allows for a simpler derivation of the results and a cleaner interpretation of the optimality conditions, it comes at the cost of not being able to find an analytic solution for the system of equations.



(A)



(B)

Figure C.3: Comparing temperature and welfare. Panel A shows temperature levels. Panel B shows welfare levels.