

Managing Termination Shock Risk of Solar Geoengineering in a Warming World

Guo-Liang Zheng ^{a, b}, Hua Liao ^{a, b}, Yi-Ming Wei ^{a, b}, Michael Freiberger ^c, Stefan Wrzaczek ^c, Michael Kuhn ^c

a Center for Energy and Environmental Policy Research, Beijing Institute of Technology, Beijing 100081, China

b School of Management, Beijing Institute of Technology, Beijing 100081, China

c International Institute for Applied Systems Analysis, Laxenburg A-2361, Austria

Emails

glzheng99@foxmail.com (Guo-Liang Zheng); hliao@bit.edu.cn (Hua Liao);
wei@bit.edu.cn (Yi-Ming Wei); freiberger@iiasa.ac.at (Michael Freiberger)
wrzaczek@iiasa.ac.at (Stefan Wrzaczek); kuhn@iiasa.ac.at (Michael Kuhn)

Note: This manuscript is a preliminary version presented at the "2026 RFF and Harvard SRM Social Science Research Workshop". The analysis and results are subject to further revision. Please contact authors for the latest version of the manuscript

<https://ceep.bit.edu.cn/english/>

Abstract

Solar Geoengineering (SRM) temporarily offsets global warming but introduces a critical concern: the termination shock risk triggered by the sudden failure of SRM deployment. This study investigates the value of proactive adjustments of carbon abatement and SRM deployment in managing termination risk. We introduce termination shock as a stochastic regime-switching process into integrated assessment models that combine geophysical and economic dynamics. Then we propose a transformation technique that reformulates the stochastic model into a deterministic structure amenable to optimal control theory, overcoming the numerical challenges of dimensionality and non-stationarity. We find that, *before the shock*, a “myopic” decision-maker (DM) who ignores the shock, relies heavily on SRM and allows carbon emissions to rise steadily from 40 GtCO₂/yr in 2020 to 78 GtCO₂/yr by 2100. By contrast, internalizing termination risk raises the Social Cost of Carbon (SCC) but slightly lowers the Social Cost of Radiative Forcing (SCF). A forward-looking “smart” DM deploys less SRM and strengthens abatement significantly. *After the shock*, both DMs abandon SRM and sharply cut emissions below the “abatement-only” pathway. Although “smart” strategy incurs higher pre-shock abatement costs and warming, it substantially reduces the post-shock temperature surge and secures GDP recovery to the “abatement-only” level—an unattainable recovery for “myopic” strategy. By 2100, it narrows the GDP uncertainty range from 665–807 trillion USD (“myopic”) to 718–803 trillion USD. Sensitivity analyses show that the dynamic structure of strategy remains robust, while revealing contrasting effects of termination probability, short-term damages, and long-term damages on SRM policy design.

Keywords: Solar geoengineering; Stochastic Dynamic Integrated Assessment Model; Social Cost of Carbon; Social Cost of Radiative Forcing; Termination Shock Risk

JEL Codes: Q54; C61; Q58

1. Introduction

Rapid global warming has intensified interest in Solar Geoengineering, also known as Solar Radiation Modification (SRM) (MacMartin et al., 2018; UNEP, 2023). Unlike indirect cooling strategies that target CO₂, such as carbon emission reductions or carbon dioxide removal (Victor & Nichols, 2024), SRM bypasses the substantial inertia of the carbon cycle, offering a potentially fast and relatively low-cost cooling effect. For instance, stratospheric aerosol injection—the most intensively studied SRM approach—appears extremely inexpensive compared to other climate interventions, with estimated deployment costs of about \$18 billion yr⁻¹ per degree Celsius of warming avoided (in 2020 USD), which is less than 0.02% of global GDP (Smith, 2020).

However, SRM does not address the root cause of climate change (greenhouse gas accumulation) and may introduce novel risks to human societies and ecosystems (Kravitz & MacMartin, 2020; Schneider et al., 2020; Zarnetske et al., 2021). One particularly concerning consequence of failing to address the underlying drivers of climate change is the risk of a termination shock, i.e., the potentially catastrophic consequences of a sudden cessation of SRM deployment (Bellamy & Lezaun, 2017), which could lead to warming rates reaching 20 times the present-day level (Matthews & Caldeira, 2007). Hence, integrating the risks of its potential termination into decision-making is crucial, when SRM is included in the optimal climate strategy.

This study aims to explore the impacts of termination shock risk and to identify optimal response strategies before and after the termination shock. To what extent would a termination shock damage the climate–economy system? How would incorporating the risk of termination into policy decisions in advance influence the optimal climate policy (including both carbon abatement and SRM deployment)? And to what degree could such proactive planning reduce the adverse impacts of a termination shock? To answer these questions, we build on the Dynamic Integrated Climate-Economy (DICE) model (Barrage & Nordhaus, 2024) and develop a stochastic dynamic integrated assessment model (IAM) that includes the termination shock risk of SRM, capturing

both its dynamic evolution and associated damages. This IAM includes the uncertainty of the cooling effect of SRM, which may lead to the occurrence of termination shock. It also distinguishes between long-term and short-term climate damages, thereby capturing the consequences of rapid warming induced by a termination shock. While we use DICE2023 as a convenient framework to explore the shock risk, our methodology and conclusions are general and can be applied to other IAMs as well. The costs and benefits of SRM are calibrated based on previous studies that have conducted extensive sensitivity analyses. Rather than advocating for a specific deployment level of SRM, this work underscores the critical role of factoring termination risks into early-stage policy planning.

A Termination shock refers to the significant risk of a rapid rebound in global temperatures following an abrupt suspension of SRM (Kemp et al., 2022; Parker & Irvine, 2018). The causes of SRM termination can be broadly divided into two categories: (1) forced termination driven by exogenous shocks, such as infrastructure failure, terrorism, natural disasters, war, or economic collapse; and (2) elective termination driven by endogenous political choices, such as growing opposition to SRM (Parker & Irvine, 2018). This study focuses on exogenous termination shocks because, compared with elective termination, exogenous termination is harder to anticipate and may leave little time for adjustment, potentially making its consequences more severe. Because the true probability of an SRM termination shock cannot be empirically identified, we treat 0.5%/yr as a plausible baseline calibration, with nuclear risk serving as a heuristic benchmark.

Although widely recognized as a critical concern, most studies to date have focused on ethical or governance debates from a qualitative perspective (Baum et al., 2013; Malm, 2023), while quantitative analyses have predominantly examined post-shock outcomes, such as climatic, ecological, or economic consequences (Jones et al., 2013; McCusker et al., 2014; Trisos et al., 2018). Ex-ante policy responses remain largely unexplored. In particular, there is a notable lack of economic analyses on how the

anticipation of termination risk should shape decision-making regarding carbon abatement and the scale of SRM deployment.

Within the cost-benefit analyses of SRM, one stream of research has examined its optimal deployment, developing foundational modeling approaches to incorporate the costs, potential damages, and cooling effects of SRM into cost-benefit analysis frameworks (Acemoglu & Rafey, 2023; Belaia et al., 2021; Bickel & Lane, 2009; Gramstad & Tjøtta, 2010; Heutel et al., 2016). Heutel et al. (2018) synthesized this body of work and introduced additional modifications to the DICE model, placing particular emphasis on the uncertainties associated with SRM damages (such as ozone depletion) and climate risks. They find that uncertainty over climate sensitivity increases both abatement and SRM deployment, whereas uncertainty over SRM damages reduces its optimal use. Distinct from earlier studies (Bickel & Agrawal, 2013; Goes et al., 2011), Heutel et al. (2018) explicitly incorporated the process of dynamic decision-making under stochastic conditions.

A similar stochastic perspective is evident in other contributions. Moreno-Cruz and Keith (2013) introduce SRM into a simple stochastic model to analyze the value of SRM in managing climate risk, emphasizing its speed, low cost, and the role of reducing inertia despite significant risks. Emmerling and Tavoni (2018) employ the probability of success of SRM into the WITCH model and show that while abatement decreases with higher probabilities of SRM success, the effect is relatively flat and robust across model specifications. Kelly et al. (2023) incorporate endogenous learning over climate sensitivity and SRM effectiveness into DICE, showing that while SRM deployment may reduce the value of learning by masking temperature signals with noise, some SRM use remains optimal, and its deployment is highly sensitive to prior beliefs about these uncertainties. Similarly, Traeger and Meier (2023) extend ACE (Analytic Climate Economy) model to analyze how uncertainties and the anticipation of learning shape the case for SRM in a world of uncertain climate responses (Traeger, 2023), highlighting that results are highly sensitive to assumptions about SRM damages and cooling effectiveness. Zheng et al. (2025) and G.-L. Zheng et al. (2026)

extend DICE model to analyze the optimal mix of abatement and SRM that address the tipping point risk of the stratocumulus clouds breakup and coral reef ecosystem collapse. Beyond a stochastic-risk perspective, G. L. Zheng et al. (2026) develops a robust SRM framework that combine cost–benefit IAM and MMR decision rule, with the aim of reducing the risk of temperature target failure under deep uncertainty. Collectively, these studies demonstrate the importance of embedding stochastic dynamics into the economic analysis of SRM, yet none explicitly examine how anticipation of termination shock risk affects optimal mitigation and SRM strategies.

Our main contribution to the existing literature is to focus on the risk of an SRM termination shock, an underexplored source of uncertainty in the assessment of SRM's cooling effectiveness. Research by Emmerling and Tavoni (2018), Kelly et al. (2023) and Traeger and Meier (2023) make important contributions to the literature on uncertainties in SRM's cooling performance but neglects the termination shock as a key risk. Some studies explicitly included termination shock in their findings. Specifically, Kosugi (2011) and Kosugi (2013) consider an irreversible termination shock event, implying that SRM deployment, once discontinued, cannot resume. They fall short in modeling the damages caused by the termination shock, particularly with respect to the impacts associated with the rate of temperature change. Goes et al. (2011) focus on intermittent SRM deployment, in which SRM is suspended for a single year and then resumed immediately. However, this does not account for intermittent SRM within a mixed policy scenario that combines both SRM and emissions mitigation strategies. Furthermore, Kosugi (2011) and Goes et al. (2011) focus only on the post-shock climate–economic impacts and on how myopic decision-makers might respond after the shock occurs. In summary, existing studies neither capture decision-making under stochastic dynamics nor consider how termination shock might be mitigated through forward-looking optimization of abatement and SRM strategies. Parker and Irvine (2018) argue that possible triggers of unexpected termination of SRM deployment include political conflicts, governance failures, technical malfunctions, natural disasters, or even deliberate attacks. Therefore, in our framework, the exact

timing of a shock is uncertain and modeled as a stochastic event, while its probability is specified exogenously. This approach enables us to analyze how the anticipation of a termination risk shapes optimal abatement and SRM strategies, a dimension that has been largely overlooked in previous studies.

Beyond the termination shock, our research provides three additional unique contributions to the literature. **First**, we calculated the Social Cost of Carbon (SCC) under termination shock risk and introduced the Social Cost of Radiative forcing (SCF). These metrics correspond to the optimal levels of emission reduction (carbon tax) and SRM deployment respectively, thereby offering intuitive policy implications (Iverson & Karp, 2021; Mattauch et al., 2020). The *SCC* is defined as the marginal rate of substitution between consumption and carbon emission, consistent with previous literature (Nordhaus, 2019; van der Ploeg, 2014; Zhao et al., 2023). Correspondingly, we introduced the *SCF* for the first time, defined as the marginal rate of substitution between consumption and radiative forcing (after deploying SRM). **Second**, we propose a method to model the damages from termination shock, resulting from the abrupt temperature increase following the suspension of SRM. Previous studies have either ignored these damages or represented them as an exogenous fixed loss. In fact, the severity of termination shock damages depends on the level of abatement and the extent of SRM deployment prior to the shock (Parker & Irvine, 2018). Therefore, we account for how optimal pre-shock policies (abatement and SRM) depend on climate damage uncertainties by distinguishing between short- and long-term climate damages. **Third**, we propose a transformation technique that enhances the transparency of analyzing models with stochastic tipping process, supporting theoretical insights while delivering stable and accurate numerical solutions. Inspired by the literature on age-structured models (Feichtinger & Wrzaczek, 2024; Freiburger et al., 2023; Wrzaczek et al., 2020), this approach reformulates a stochastic framework with a single tipping event into a structure amenable to optimal control theory (Feichtinger et al., 2003). Stochastic analysis of climate and economy poses severe computational challenges for conventional numerical approaches, particularly the

stochastic tipping/switch process (Barnett, 2023; Collin-Dufresne et al., 2016; Golosov et al., 2014; Van den Bremer & Van der Ploeg, 2021). One significant strength of our approach is to overcome computational difficulties commonly encountered in solving stochastic cost–benefit IAM, such as the curse of dimensionality and issues of non-stationarity (Cai, 2019; Cai & Lontzek, 2019; Lemoine & Traeger, 2014).

We first present the case of a “myopic” decision-maker (DM), who does not take into account the risk of a termination shock. Although the myopic scenario has been discussed in previous literature, such as Kosugi (2011), those studies typically do not include the endogenous relationship between the pre- shock strategy and post-shock impacts. Besides, the myopic case provides a benchmark for alternative scenarios to be compared to. In this study, we find that relative to the “abatement-only” scenario (corresponding to the standard solution of the DICE model), the myopic DM substantially increases both carbon emissions and SRM deployment. Once termination occurs, however, SRM has to be abandoned abruptly and emissions in turn are drastically reduced, falling below abatement-only levels. This sharp adjustment also causes savings rates to drop temporarily in order to smooth consumption before gradually recovering. The termination shock further triggers rapid warming—about 0.6 °C within five years if occurring mid-century, and 1.2 °C if occurring late-century. Economic output rises slightly before termination but drops steeply afterward, remaining below the abatement-only trajectory, though it eventually recovers. Due to the possible arrival (or not) of the termination shock, by the end of the century, the myopic path yields a GDP uncertainty range of roughly 665–807 trillion USD.

We then turn to a “smart” DM who anticipates that termination risk will occur with a certain probability, yet remains uncertain of the timing. Compared to the myopic DM, we find that the forward-looking DM deploys less SRM and emits less carbon, although still above the abatement-only baseline. After termination, similar sharp cuts are required, yet the preemptive reduction in SRM lowers exposure to abrupt warming and results in smaller economic losses. By 2100, the GDP range narrows to about 718 to 803 trillion USD, indicating that anticipation of termination risk mitigates post-shock

damages. From the perspective of expected pathways, the myopic DM emits 6–40 Gt CO₂ more annually than abatement-only, whereas the smart DM emits only 2–20 Gt CO₂ more.

Finally, since neither the probability of occurrence nor the scale of post-shock damages have been empirically observed so far, we conduct sensitivity analyses on the probability of a termination shock, as well as on the severity of shock damages. In terms of damages, we distinguish between two components of shock damages, the short-term damages that capture the catastrophic but transitory impacts, as well as the long-term damages that capture the persistent warming losses. These exercises demonstrate that the Smart DM's strategy is robust in its overall structure, consistently favoring moderate SRM deployment and stronger abatement **before the shock** (compared to the myopic case), and scaling up abatement rapidly to compensate for the loss of SRM **after the shock**. However, the SRM and abatement efforts adjust depending on the specifications of shock probability and climate damages. The Smart DM decides to adopt even more conservative SRM use and to strengthen abatement earlier when facing higher probability or higher short-term damage. In contrast, higher long-term damages reinforce the use of both SRM and abatement simultaneously. This results from higher long-term damages not only amplifying the effects of post-impact warming but also exacerbating climate damages occurring prior to the termination shock.

The remainder of the paper is organized as follows. Section 2 introduces the modeling framework, the continuous-time DICE-2023 model with stochastic termination, and its age-structured transformation. Section 3 provides the theoretical analysis, presenting the measures of SCC and SCF. Then, the following section reports the numerical results, sensitivity analyses and summary.

2. Methodology

The stochastic IAM modifies the Dynamic Integration of Climate and Economy (DICE) 2023 model, a work-horse model in climate-economy analysis, which integrates a stylized climate system (DFAIR models) with an economy system (Ramsey–Cass–Koopmans economic model) (Barrage & Nordhaus, 2024). In the DICE model, a social planner determines savings and emission reductions in each period to achieve economic growth while mitigating climate change and its associated damages.

Various approaches have been proposed for modeling SRM—for instance, reducing radiative forcing in either an additive or a multiplicative form. We follow the specification of Heutel et al. (2018), who provides a comprehensive summary of the existing modeling approaches and evaluates their implications. Subsequently, we extend the model by incorporating stochastic dynamics and associated damages of the termination shock risk, which turns the model into a stochastic decision-making framework.

The theoretical analysis reformulates the stochastic IAM into a two-stage optimal control model with stochastic switching (2SOCS), providing clear economic intuition. In parallel, the proposed numerical algorithm for 2SOCS addresses key challenges in solving stochastic IAMs, such as the “curse of dimensionality” and “non-stationarity”(Cai, 2019). The numerical results quantify the climate–economic impacts of termination shock and identify the optimal response strategy, combining abatement and SRM policies. For parameter calibration, most parameters adhere to the DICE 2023 standard. For parameters related to implemented solar radiation management (SRM) technologies, such as deployment costs and potential destructive power, the findings from Heutel et al. (2018) were referenced. Parameters regarding SRM termination shock are provided in this section.

The methodology unfolds in three steps: we (1) introduce the core modeling approach for incorporating termination shock risk into the DICE framework.; (2) present the full continuous-time stochastic integrated assessment model, detailed in the Appendix; and (3) reformulate the problem into a 2SOCS structure to enable tractable analysis and numerical implementation.

2.1 Modeling Termination Shock Risk of SRM

Dynamic process

The DICE-2023 climate model (D-FAIR) is based on the FaIR model. An increase in atmospheric carbon concentration (which is affected by abatement efforts μ_t) in period t M_t^{at} affects F_t (radiative forcing) as follows:

$$F_t = \left(\eta * \log_2 \left(\frac{M_t^{at}}{M_*^{at}} \right) + F_t^{abate} + F_t^{ex} \right) * (1 - I_t * g_t) \quad [1]$$

where η is a climate parameter, measuring the level of radiative forcing induced by doubling carbon concentrations. F_t^{ex} is the exogenous radiative forcing from non-abatable GHG and other sources, and F_t^{abate} is the radiative forcing from abatable non-CO2 greenhouse gases (GHG), depending on the carbon abatement rate μ_t :

$$F_{t+1}^{abate} = \lambda^f * F_t^{abate} + \lambda^c * CE_t^{abate} * (1 - \mu_t), \quad [2]$$

where λ^f and λ^c are parameters, CE_t^{abate} are exogenous baseline GHG emissions from abatable non-CO2. Therefore, $\left(\eta * \log_2 \left(\frac{M_t^{at}}{M_*^{at}} \right) + F_t^{abate} + F_t^{ex} \right)$ measures the total radiative forcing of GHGs before deploying SRM.

SRM directly reduces the radiative forcing, thus leading to a cooling effect. SRM efforts are typically short-lived and there is no (significant) long-run SRM capital. Therefore, a new decision variable, g_t , is introduced in (eq. 1). g_t is the radiative control rate, representing the proportion of radiative forcing offset by SRM, which follows the specification of Heutel et al. (2018). When g_t is equal to one, SRM may completely offset the radiation caused by the rise of the GHG stock.

Further, I_t is a random binary variable, representing the state of the shock event. $I_t = 1$ and $I_t = 0$ represents the pre-shock state (state 1) and post-shock state (state 2), respectively. I_t is a stochastic Markov process, and the corresponding probability transition matrix at t is:

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

$$p_t = 1 - e^{-\eta t} \quad [3]$$

where the (i, j) element of the matrix is the probability of I_t going from state i to j in period t . According to Parker and Irvine (2018), partial failure or temporary suspension of SRM may not qualify as a termination shock, as they are typically not associated with significant abrupt risks. Our modeling assumes that the occurrence of a termination shock leads to the complete and permanent failure of SRM. This assumption represents the scenario that captures the most severe form of termination risk. Alternative cases—such as partial failure or temporary suspension of SRM—can be interpreted as lower-impact variants of this worst-case scenario. Therefore, in our model, the probability to transit from state 2 to state 1 is 0, which represents the irreversibility of the termination shock under worst-case. Parker and Irvine (2018) emphasize that termination is likely to result from external shocks (Martin & Pindyck, 2015), thus, we model the shock probability of p_t exogenously and treat the timing as stochastic. Potential termination triggers are not limited to engineering failures of the SRM system itself, but may also arise from broader systemic and geopolitical shocks (Parker & Irvine, 2018). Since the true likelihood of such an event is neither empirically observable nor well established in the literature, the hazard rate used here should be interpreted as a modeling calibration rather than an empirical estimate.

Nevertheless, we consider a baseline termination hazard of 0.5%/yr (the hazard rate is 0.5%) to be plausible. Nuclear risk provides a useful heuristic comparison because both systems depend on complex infrastructure, state-level decisions, and geopolitical stability. Hellman’s order-of-magnitude estimate places the annual probability of full-scale nuclear war at about 1%/yr, with a rough range of 0.3%–3%/yr (Hellman & Cerf, 2021); our baseline hazard of 0.5%/yr lies within this range. A 0.5% annual hazard also implies an expected waiting time of about 200 years. By comparison, Chernobyl (1986) and Fukushima (2011) occurred only 25 years apart, and statistical analyses estimate a 50% probability of a Fukushima-scale or larger accident within roughly 60–150 years (Wheatley et al., 2017). More broadly, even highly reliable critical infrastructures, such as the Internet and GPS, have operational histories far shorter than 200 years.

In this framework, the radiative forcing F_t affects the temperature box T_t^j , which contributes to the atmospheric temperature increase due to deep and upper ocean processes:

$$T_{t+1}^j = T_t^j * \exp\left(-\frac{1}{\varsigma^j}\right) + T_*^j * F_t * \left(1 - \exp\left(-\frac{1}{\varsigma^j}\right)\right), \quad [4]$$

$$T_t^{at} = \sum_1^2 T_t^j, \quad [5]$$

for $j=1, 2$, where T_*^j represents the thermal equilibration parameter (diffusion rate) for box j , and ς^j is the thermal response timescale (time lag) for box j . The T_t^{at} denotes the atmospheric temperature anomaly. Thus, the model captures the dynamic mechanism of a termination shock, in which the abrupt loss of SRM (I_t from 1 to 0) results in additional radiative forcing and induces rapid warming.

Associated damage

The termination shock induces a sudden temperature increase, accompanied by significant damages. Both the magnitude and velocity of temperature change contribute to the severity of climate and economic impacts. However, previous studies on termination shock risk have largely ignored the damage caused by the velocity of temperature change (Kosugi, 2011, 2013), although it has been found that some climate damages depend on the velocity of climate change (Differbaugh & Field, 2013). For example, Texas, which has a relatively warm winter, experienced a rare temperature plunge in February 2021 (i.e., the "Texas Freeze"), which led to catastrophic power outages and triggered a chain of events such as fuel shortages, burst pipelines and potable water shortages, icy roadways, and traffic accidents. These events ultimately resulted in mass casualties and vast economic damages estimated to be as high as \$130 billion (Ghosh et al., 2021).

To model the damages from a termination shock, we propose a functional form that allows us to incorporate the effects of ex-ante decisions (mixed SRM and abatement policy) on mitigate termination shock risk. Our modeling approach is inspired by the methods of Goes et al. (2011) and Lempert et al. (2000), where the total damage from climate change D_t (measured as a percentage of output) is decomposed into

short-term and long-term components:

$$D_t = D_t^L + D_t^S \quad [6]$$

where D_t^L and D_t^S represent the climate damage from long-term warming and sudden temperature change, respectively.

The long-term damages D_t^L follow the DICE-2023 formulation:

$$D_t^L = \beta_1 (T_t^{at})^2 \quad [7]$$

where T_t^{at} represents the atmospheric temperature anomaly (i.e., the increase in atmospheric temperature compared to the pre-industrial level) in period t .

The short-term damages, which are driven by abrupt climate disruption, are assumed to depend on the difference between the current warming level and the average warming of the past s years. More specifically,

$$D_t^S = \beta_2 \left(\frac{T_t^{at} - \overline{T_{s,t}^{at}}}{\omega \cdot s} \right)^2, \quad [8]$$

where $\overline{T_{s,t}^{at}}$ is the moving average value of T_t^{at} over the last s years (in our baseline setting, we assume $s=5$ years). ω is the maximum acceptable annual warming rate. Parker and Irvine (2018) suggested a maximum acceptable warming rate of 0.2°C per decade based on RCP2.6 scenario. The $T_t^{at} - \overline{T_{s,t}^{at}}$ then reflect short-term changes in temperature. Our choice of the five-year average in the baseline setting implies the assumption that the effects of a sudden temperature change mainly manifest within five years, which means that it is likely to take five years to adapt to the sudden temperature change and attenuate its effects. Therefore, the cumulative acceptable warming rate over the past s years is 0.1°C , i.e., $\omega \cdot s = 0.1$. In addition, we assume that $\beta_2 = 1\%$, implying that it causes 1% damage if current warming exceeds the moving average of the past s years by 0.1°C . Since the temperature deviation is squared, the inclusion of D_t^S has only a marginal effect in the standard DICE setting as temperature increases gradually. Hence, we abstain from adjusting the parameter β_1 of the long-term damage and choose the same value as in the DICE 2023. As there is no consent in the literature on the extent of climate damages for specific levels of short- and long-term temperature increases (Okullo, 2020), we conduct an extensive sensitivity analysis on the damage parameters β_1 and β_2 . These two parameters

comprehensively evaluate the critical role of climate damage from short-term and long-term in shaping optimal climate policy decisions under termination shock risk.

2.2 Continuous-Time DICE model with Termination shock risk

Objective function

In our extended version of the DICE-2023 model, the DM adjusts economic policy (the savings rate s), abatement policy (the abatement rate μ) and SRM policy (the radiative forcing control rate g) to maximize total social welfare over time, which is constituted by the flow of household consumption over time. The social welfare is the discounted integral of the population-weighted utility of per-capita consumption.

$$\max_{s_t, \mu_t, g_t \in [0,1]} \int_0^T e^{-\int_0^t \rho(\tau) d\tau} L(t) \cdot U(c(t)) dt \quad [9]$$

where $\rho(t)$ is the time-varying rate of time preference (a discount applied to future utility or wellbeing at time t), $L(t)$ is the exogenous size of the population/labor, $U(c(t))$ is the utility of consumption, and $c(t)$ is the per-capita consumption $C(t)/L(t)$. The total consumption $C(t)$ is the portion of net output $Q(t)$ not used for saving, i.e., $(1 - s(t))Q(t)$.

Our model incorporates a stochastic shock process represented by the binary indicator $I(t, \omega) \in \{0,1\}$, which denotes whether an irreversible shock event has occurred by time t along a given realization $\omega \in \Omega$ of the stochastic environment. The termination shock risk ultimately affects social welfare by influencing consumption flows $c(t; I(t, \omega))$. The social planner maximizes expected total welfare over all possible realizations of the shock process:

$$\max_{s_t, \mu_t, g_t \in [0,1]} \mathbb{E}_{\omega \in \Omega} \left\{ \int_0^T e^{-\int_0^t \rho(t') dt'} L(t) \cdot U(c(t; I(t, \omega))) dt \right\} \quad [10]$$

For notational simplicity, we denote the stochastic process as $I(t)$, omitting explicit dependence on the random realization $\omega \in \Omega$. The process $I(t)$ follows a Poisson jump process with a constant hazard rate η , implying that the probability of a shock

occurring in a small interval $[t, t + \Delta t]$ is $\eta \cdot \Delta t + o(\Delta t)$. The waiting time until the shock, denoted τ , follows an exponential distribution with parameter η , which means that the shock event occurs at a random time τ with probability distribution $P(\tau > t) = e^{-\eta t}$. Once the shock occurs, $I(t)$ switches from 1 to 0 permanently and modifies the system dynamics to feature a loss of the SRM cooling effect.

Controls and State Variables

A social planner is assumed to set the values of the following control variables at each point in time: the *savings rate* $s \in [0, 1]$ denotes the fraction of net economic output allocated to capital reinvestment; the *abatement rate* $\mu \in [0, 1]$ represents the proportion of industrial carbon emissions that is reduced through mitigation efforts; the *radiative forcing control rate* $g \in [0, 1]$ indicates the fraction of radiative forcing from all kinds of greenhouse gases that is offset by SRM.

At every timepoint, the state of the system is described by the following state variables: the *capital stock*, K ; the *carbon reservoirs*, $R = (R_1, R_2, R_3, R_4)$, representing carbon concentrations in four different reservoirs (j), measured from the pre-industrial baseline (starting from the year 1765); the *accumulated carbon emissions*, C_{acc} ; the *radiative forcings from abatable non-CO2 greenhouse gases*, F_{GHG} ; the *temperature box model*, $Tat = (Tat_1, Tat_2)$, where Tat_i is the increase in atmospheric temperature in box i , relative to pre-industrial levels, and the *moving average of global atmospheric warming*, $\overline{Tat}$. In this framework, the sum of carbon reservoirs R_t at time t plus the initial atmospheric carbon stock M_*^{at} , represents the current atmospheric carbon concentration M_t^{at} . The sum of the temperature boxes $\sum_i Tat_{i,t}$ at time t represents the current level of global atmospheric warming relative to the pre-industrial level; the *shock indicator*, I , a binary state variable captures the occurrence of an irreversible shock. It equals 1 prior to the shock event and switches to 0 upon occurrence, thereafter remaining constant.

Controls (3):

$s \in [0, 1]$ savings
 $\mu \in [0, 1]$ abatement
 $g \in [0, 1]$ SRM

States (11):

K capital stock
 R_j carbon concentration in reservoir j ($j=1,2,3,4$)
 C_{acc} accumulated carbon emissions
 F_{GHG} radiative forcings from abatable non-CO2 GHG
 Tat_i warming from box i ; ($i=1,2$)
 $\overline{Tat}$ moving average of warming
 I shock indicator

The controlled dynamics governing the evolution of the eleven state variables form a complex system, the detailed formulation of which will be presented in the Appendix.

Full model

Putting it all together results in a stochastic dynamic optimization problem in which a social planner maximizes expected intertemporal welfare under the termination shock risk (detailed in Appendix). This framework combines an integrated assessment model with stochastic elements of the shock risk. The planner chooses three control variables: the savings rate $s(t)$, the abatement rate $\mu(t)$, and the SRM deployment rate $g(t)$.

$$\max_{s_t, \mu_t, g_t \in [0,1]} \mathbb{E}_{\omega \in \Omega} \left\{ \int_0^T e^{-\int_0^t \rho(t') dt'} L(t) \cdot U(c(t; I(t, \omega))) dt \right\} \quad [11]$$

subject to

$$\dot{K} = s \cdot Q(\mu, g, K, Tat_i, \overline{Tat}) - \delta_k K \quad K(0) = K_0 \quad [12]$$

$$\dot{R}_j = \xi_j \frac{1}{co2c} E_{co2}(\mu, K) - \frac{1}{\tau_j \psi(C_{acc}, \sum_j R_j, \sum_i Tat_i)} R_j, \quad j = 1, 2, 3, 4 \quad R_j(0) = R_{j_0} \quad [13]$$

$$\dot{C}_{acc} = \frac{1}{co2c} E_{co2}(\mu, K) \quad C_{acc}(0) = C_{acc_0} \quad [14]$$

$$\dot{F}_{GHG} = (1 - \mu) \psi_1 E_{GHG} - \psi_2 F_{GHG} \quad F_{GHG}(0) = F_{GHG_0} \quad [15]$$

$$\dot{Tat}_i = \frac{1}{\kappa_i} [\chi_i (1 - I \cdot g) TF(\sum_j R_j, F_{GHG}) - Tat_i], \quad i = 1, 2 \quad Tat_i(0) = Tat_{i_0} \quad [16]$$

$$\dot{\overline{Tat}} = \bar{\chi} \cdot (\sum_i Tat_i - \overline{Tat}) \quad \overline{Tat}(0) = \overline{Tat}_0 \quad [17]$$

$$\frac{d\mathbb{P}(\tau > t)}{dt} = -\eta \mathbb{P}(\tau > t) \quad \text{with } I(t) = \mathbf{1}_{\{t \leq \tau\}} \quad \mathbb{P}(\tau > 0) = 1 \quad [18]$$

where

$$c = (1 - s) \cdot Q(\mu, g, K, Tat_i, \overline{Tat}) / L \quad [19]$$

2.3 Two-stage optimal control models with stochastic switch

We transform stochastic problems to the two-stage optimal control models with stochastic switch (2SOCS) (Wrzaczek et al., 2020). Specifically, we divide the time horizon T into two stages, with the shock time τ serving as the dividing point. **Stage 1** refers to the pre-shock period. In this stage, the DM does not know the exact realization of τ , but is aware that it may occur at any time and understands its potential effects on the system. Thus, when formulating a Stage 1 strategy ex ante, the planner must design a policy that spans the entire horizon, i.e., for all $t \in [0, T]$. We denote the decision variables in the first stage by

$$s_1(t), \mu_1(t), g_1(t)$$

Stage 2 corresponds to the post-shock period. Here, the DM knows both that the shock has occurred and the exact time τ . The strategy can therefore be adjusted conditional on the occurrence of the event. Accordingly, a Stage 2 strategy must be defined for all possible realizations of the switching time, i.e., for all $\tau \in [0, T]$ and for all $t \in [\tau, T]$. we denote by

$$s_2(t, \tau), \mu_2(t, \tau), g_2(t, \tau)$$

the Stage 2 savings, abatement and SRM rate respectively, at time t , given that the shock occurred at time τ . The same holds for the state trajectories, e.g. $K_1(t)$ in Stage 1 and $K_2(t, \tau)$ in Stage 2.

Therefore, the objective function can be rewritten as

$$\max_{s_1, \mu_1, g_1, s_2, \mu_2, g_2 \in [0, 1]} \mathbb{E}_\tau \left\{ \int_0^\tau e^{-\int_0^t \rho(t') dt'} L(t) U(c_1(t)) dt + \int_\tau^T e^{-\int_0^t \rho(t') dt'} L(t) U(c_2(t, \tau)) dt \right\} \quad [20]$$

Next, we replace the indicator variable I with two auxiliary state variables, z_1, z_2 in Stage 1 and 2 respectively, that are linked to the distribution of the random shock time τ .

The auxiliary state variables are defined as,

$$z_1(t) = \mathbb{P}(\tau > t), \quad z_2(t, \tau) = \eta(\tau) z_1(\tau) \quad [21]$$

The interpretation of $z_2(t, \tau)$ can be found in Wrzaczek et al. (2020). They obey the following dynamics:

$$\dot{z}_1(t) = -\eta z_1 \text{ with } z_1(0) = 1 \quad [22]$$

$$\dot{z}_2(t, \tau) = 0 \text{ with } z_2(\tau, \tau) = \eta(\tau) z_1(\tau) \quad [23]$$

Then reformulating the objective value using a vintage structure. Note that the warming box (Tat_i) dynamics differ between the two Stages, with SRM no longer affecting radiation in Stage 2. Finally, the 2SOCS formulation of the stochastic IAM is,

$$\max_{s_1, \mu_1, g_1, s_2, \mu_2, g_2 \in [0,1]} \int_0^T e^{-\int_0^t \rho(t') dt'} [z_1(t) L(t) U(c_1(t)) + H(t)] dt \quad [24]$$

subject to

Stage 1

$$\dot{z}_1 = -\eta z_1 \quad z_1(0) = 1 \quad [25]$$

$$\dot{K}_1 = s_1 Q(\mu_1, g_1, K_1, Tat_{i,1}, \overline{Tat}_1) - \delta_k K_1 \quad K_1(0) = K_0 \quad [26]$$

$$\dot{R}_{j,1} = \xi_j \frac{E_{co2}(\mu_1, K_1)}{co2c} - \frac{R_{j,1}}{\tau_j \psi(C_{acc1}, \sum_j R_{j,1}, \sum_i Tat_{i,1})}, j = 1, 2, 3, 4 \quad R_{j,1}(0) = R_{j,0} \quad [27]$$

$$\dot{C}_{acc1} = \frac{E_{co2}(\mu_1, K_1)}{co2c} \quad C_{acc1}(0) = C_{acc0} \quad [28]$$

$$\dot{F}_{GHG1} = \psi_1(1 - \mu_1) E_{GHG} - \psi_2 F_{GHG1} \quad F_{GHG1}(0) = F_{GHG0} \quad [29]$$

$$\dot{Tat}_{i,1} = \frac{1}{\kappa_i} [\chi_i \cdot TF(R_{j,1}, F_{GHG1})(1 - g_1) - Tat_{i,1}], i = 1, 2 \quad Tat_{i,1}(0) = Tat_{i,0} \quad [30]$$

$$\dot{\overline{Tat}}_1 = \bar{\chi} \cdot (\sum_i Tat_{i,1} - \overline{Tat}_1) \quad \overline{Tat}_1(0) = \overline{Tat}_0 \quad [31]$$

Stage 2

$$H = \int_0^t z_2 LU(c_2) d\tau \quad [32]$$

$$\dot{z}_2 = 0 \quad z_2(\tau, \tau) = \eta(\tau) z_1(\tau) \quad [33]$$

$$\dot{K}_2 = s_2 Q(\mu_2, g_2, K_2, Tat_{i,2}, \overline{Tat}_2) - \delta_k K_2 \quad K_2(\tau, \tau) = K_1(\tau) \quad [34]$$

$$\dot{R}_{j,2} = \xi_j \frac{E(\mu_2, K_2)}{co2c} - \frac{R_{j,2}}{\tau_j \psi(C_{acc2}, \sum_j R_{j,2}, \sum_i Tat_{i,2})}, j = 1, 2, 3, 4 \quad R_{j,2}(\tau, \tau) = R_{j,1}(\tau) \quad [35]$$

$$\dot{C}_{acc2} = \frac{E(\mu_2, K_2)}{co2c} \quad C_{acc2}(\tau, \tau) = C_{acc1}(\tau) \quad [36]$$

$$\dot{F}_{GHG2} = \psi_1(1 - \mu_2) E_{GHG} - \psi_2 F_{GHG2} \quad F_{GHG2}(\tau, \tau) = F_{GHG1}(\tau) \quad [37]$$

$$\dot{Tat}_{i,2} = \frac{1}{\kappa_i} [\chi_i \cdot TF(R_{j,2}, F_{GHG2}) - Tat_{i,2}] \quad Tat_{i,2}(\tau, \tau) = Tat_{i,1}(\tau) \quad [38]$$

$$\dot{\overline{Tat}}_2 = \bar{\chi} \cdot (\sum_i Tat_{i,2} - \overline{Tat}_2) \quad \overline{Tat}_2(\tau, \tau) = \overline{Tat}_1(\tau) \quad [39]$$

3. Theoretical Analysis

First, we apply the Maximum Principle for age-structured control problem to the 2SOCS (Feichtinger & Wrzaczek, 2024).. Then, we derive the social cost of carbon (SCC) under termination shock risk. The SCC represents the marginal welfare loss from an additional unit of emissions and, in turn, determines the optimal level of abatement. In parallel to the SCC, we introduce the concept of the social cost of radiative forcing (SCF), which measures the marginal welfare cost of an additional unit of radiative forcing.

3.1 Optimality Conditions

Let $x_1(t)$ denote the state variables vector (including capital stock, carbon concentration, temperature warming, shock indicator, etc.) and $u_1(t)$ the vector of control variables for Stage 1 (including saving rate, abatement rate and SRM rate) and similarly, $x_2(t, \tau)$ and $u_2(t, \tau)$ for Stage 2. The system dynamics can be compactly expressed as

$$\dot{x}_1(t) = f_1(t, x_1(t), u_1(t)), \quad x_1(0) = x_0 \quad [40]$$

$$\dot{x}_2(t, \tau) = f_2(t, x_2(t, \tau), u_2(t, \tau)), \quad x_2(\tau, \tau) = b(t, x_1(\tau)) \quad [41]$$

where $b(\cdot)$ is the 'switching conditions' of the state variables from Stage 1 to Stage 2. We will denote by λ_{x_1} the adjoint variable of the state variable for Stage 1 x_1 , and by η_{x_2} the adjoint variable of the state variable for Stage 2 x_2 . Likewise, λ, η denote the co-state vectors of Stage 1 and 2 respectively.

$$\lambda = \lambda(t), \quad \eta = \eta(t, \tau) \quad [42]$$

The current-value Hamiltonian (with simplified notation) is defined as,

$$\begin{aligned} \mathcal{H}(u_1, u_2, x_1, x_2, \lambda, \eta, \zeta) \\ = z_1[LU(c_1) + H] + \lambda^T f_1(t, x_1, u_1) + \eta^T f_2(t, x_2, u_2) + \zeta z_2 LU(c_2) \end{aligned} \quad [43]$$

where ζ is the adjoint variable of H .

First Order Conditions

Due to the concavity of the Hamiltonian with respect to the controls, the control variables do not lie on the boundaries, i.e., they belong to the inner part (0, 1) of the control set [0, 1]. Therefore, for all period such that the optimal savings rate s_1^* and s_2^* is an inner solution (arguments are omitted to lighten the notation),

$$\begin{cases} 0 = \partial_{s_1} \mathcal{H} = \lambda_K - z_1 U_{c_1}(c_1) \\ 0 = \partial_{s_2} \mathcal{H} = \eta_K - z_2 U_{c_2}(c_2) \end{cases} \quad [44]$$

Analogously, whenever s_1^*, s_2^*, μ_1^* and μ_2^* are inner controls,

$$\begin{cases} 0 = \partial_{\mu_1} \mathcal{H} = \lambda_{FGHG} \psi_1 E_{GHG} + \lambda_K \Lambda_{\mu_1} Y_1 + (\sum_j \lambda_{R,j} \xi_j + \lambda_{Cacc}) \frac{\sigma Y_1 + E_{land}}{co2c} \\ 0 = \partial_{\mu_2} \mathcal{H} = \eta_{FGHG} \psi_1 E_{GHG} + \eta_K \Lambda_{\mu_2} Y_2 + (\sum_j \eta_{R,j} \xi_j + \eta_{Cacc}) \frac{\sigma Y_2 + E_{land}}{co2c} \end{cases} \quad [45]$$

yielding,

$$\begin{cases} -\frac{\sum_j (\lambda_{R,j} \xi_j) + \lambda_{Cacc}}{\lambda_K} \cdot \frac{\sigma Y_1 + E_{land}}{co2c} (1 - \mu_1^*) - \frac{\psi_1 (1 - \mu_1^*) \lambda_{FGHG}}{\lambda_K} E_{GHG} = (1 - \mu_1^*) \Lambda_{\mu_1} Y_1 \\ -\frac{\sum_j (\eta_{R,j} \xi_j) + \eta_{Cacc}}{\eta_K} \cdot \frac{\sigma Y_2 + E_{land}}{co2c} (1 - \mu_2^*) - \frac{\psi_2 (1 - \mu_2^*) \eta_{FGHG}}{\eta_K} E_{GHG} = (1 - \mu_2^*) \Lambda_{\mu_2} Y_2 \end{cases} \quad [46]$$

These two equations determine the relationship between Social Cost of Carbon (SCC) and the marginal abatement cost (see next section).

Also, whenever s_1^*, s_2^*, g_1^* and g_2^* are inner controls,

$$\begin{cases} 0 = \partial_{g_1} \mathcal{H} = \lambda_K \Omega_{g_1} Y_1 + \sum_i \lambda_{Tat,i} \frac{\chi_i}{\kappa_i} TF(R_{j,1}, F_{GHG_1}) \\ 0 = \partial_{g_2} \mathcal{H} = \lambda_K \Omega_{g_2} Y_2 \end{cases} \quad [47]$$

yielding

$$\begin{cases} -\frac{\sum_i (\lambda_{Tat,i} \frac{\chi_i}{\kappa_i})}{\lambda_K} = \frac{\Omega_{g_1} Y_1}{TF(t, R_{j,1}, F_{GHG_1})} \\ g_2^* = 0 \end{cases} \quad [48]$$

The first equation determines the relationship between Social Cost of Radiative forcing (SCF) and the optimal SRM level (see next section). The second suggest that the optimal strategy of Stage 2 is to abandon SRM. This is intuitive because SRM has no cooling effect in the second stage beyond cost and damage.

Adjoint Equations

We computed the adjoint equations for all the variables (See Appendix A). Here we focus on the main co-state trajectories appearing in the calculation of the SCC and SCF. For simplicity, dependencies on time and switching (shock) time are omitted here. For 2SOCS, the two-stage adjoint equations are symmetric up to the last term. In general, the shadow value in Stage 1 also depends on the shadow value in Stage 2 as well as on the *switching condition* of the state variables, i.e. $x_2(t, t) = g(t, x_1(t))$.

This dependence reflects the DM's awareness of the termination shock risk.

For the costate equations for capital it holds that,

$$\dot{\lambda}_K = (\rho + \delta_K)\lambda_K - \left[(1 - \Lambda - D - \Omega)\lambda_K + \frac{\sigma}{co2c}(1 - \mu_1^*) \left(\sum_j (\lambda_{R,j}\xi_j) + \lambda_{Cacc} \right) \right] Y_{1K_1} - \eta_K \quad [49]$$

$$\dot{\eta}_K = (\rho + \delta_K)\eta_K - \left[(1 - \Lambda - D - \Omega)\eta_K + \frac{\sigma}{co2c}(1 - \mu_2^*) \left(\sum_j (\eta_{R,j}\xi_j) + \eta_{Cacc} \right) \right] Y_{2K_2} \quad [50]$$

In this case, compared to the scenario without SRM termination shock risk, the additional term in the Stage 1 equation is the last term, $-\eta_K(t, t)$. Under the reasonable assumption that both $\lambda_K, \eta_K > 0$ (shown in the numerical results), it follows that the growth rate and the level of the shadow price of capital in Stage 1 are lower than in the case without shock risk. Intuitively, the shadow value of capital in the first stage tends to decline because termination shocks reduce output and reduce the benefits of capital accumulation. Furthermore, under optimal decision-making, the planner reduces the capital growth rate and overall accumulation level in order to mitigate the impact of potential termination shocks, which is a form of self-insurance. This inference is validated by the numerical results.

Note that $g_2^* = 0$ and $\Omega_{g_2} = 0$. To simplify notation, we define

$$\Phi(t, C_{acc}, \sum_j R_j, \sum_i T_{at_i}) = \frac{1}{\tau_j \psi(C_{acc}, \sum_j R_j, \sum_i T_{at_i})}, \quad [51]$$

and then Φ_1, Φ_2 for each stage. Therefore, for the costate equations for carbon reservoir j ,

$$\dot{\lambda}_{R,j} = \rho \lambda_{R,j} + \Phi_1 \sum_j \left(\lambda_{R,j} \left(1 + \frac{\Delta^M}{\psi^1} R_{j,1} \right) \right) + \frac{\Phi_1}{\ln 2 * M^{at}(\sum_j R_{j,1})} \sum_i \left(\lambda_{T_{at,i}} \frac{\chi_i}{\kappa_i} \right) (1 - g_1^*) - \eta_{R,j} \quad [52]$$

$$\dot{\eta}_{R,j} = \rho \eta_{R,j} + \Phi_2 \cdot \sum_j \left(\eta_{R,j} \left(1 + \frac{\Delta^M}{\psi^1} R_{j,2} \right) \right) + \frac{\Phi_2}{\ln 2 * M^{at}(\sum_j R_{j,2})} \sum_i \left(\eta_{Tat,i} \frac{\chi_i}{\kappa_i} \right) \quad [53]$$

From the costate equations, one can directly see that the dynamics of the shadow price of carbon in Stage 1 differs from Stage 2 in two key respects. First, the temperature feedback term in Stage 1 is multiplied by $(1 - g_1^*)$, reflecting that SRM mitigates the impact of carbon on temperature and hence on damages. However, when SRM is terminated, this mitigation channel disappears, so that the entire forcing effect of the carbon stock is transmitted into temperature, which makes the accumulation of carbon more harmful after the shock.

Second, the Stage 1 equation contains the additional transition term $-\eta_{R,j}$, which stems from the switching condition of the carbon stock across the shock. Carbon stock in Stage 1 persists into Stage 2. Under the reasonable assumption that both $\lambda_{R,j}, \eta_{R,j} < 0$ because carbon is a pollutant, this adjustment implies that Stage 1 shadow prices must already internalize the damages that will occur in Stage 2 and raise the welfare cost of accumulating carbon in Stage 1.

For the costate of temperature box,

$$\dot{\lambda}_{Tat,i} = \left(\rho + \frac{1}{\kappa_i} \right) \lambda_{Tat,i} + \lambda_K(D_{Tat,i,1} Y_1) - \left(\frac{\Delta^T}{\psi^1} \right) \Phi_1 \sum_j (\lambda_{R,j} R_{j,1}) - \bar{\chi} \lambda_{\overline{Tat}} - \eta_{Tat,i} \quad [54]$$

$$\dot{\eta}_{Tat,i} = \left(\rho + \frac{1}{\kappa_i} \right) \eta_{Tat,i} + \eta_K(D_{Tat,i,2} Y_2) - \left(\frac{\Delta^T}{\psi^1} \right) \Phi_2 \sum_j (\eta_{R,j} R_{j,2}) - \bar{\chi} \eta_{\overline{Tat}} \quad [55]$$

Those equations include the shadow prices of moving-average temperature $\lambda_{\overline{Tat}}$, which reflects that the rate of short-term warming contributes to climate damages.

Similar to the costate equations of carbon, the frontloading term $-\eta_{Tat,i}$ shows that Stage 1 shadow prices of warming already incorporate the future damages that will occur in Stage 2.

Next, for the costate of moving average warming,

$$\dot{\lambda}_{\overline{Tat}} = (\rho + \bar{\chi}) \lambda_{\overline{Tat}} + \lambda_K(D_{\overline{Tat},1} Y_1) - \eta_{\overline{Tat}} \quad [56]$$

$$\dot{\eta}_{\overline{Tat}} = (\rho + \bar{\chi}) \eta_{\overline{Tat}} + \lambda_K(D_{\overline{Tat},2} Y_1) \quad [57]$$

In practical terms, a higher $\bar{\chi}$ implies a shorter averaging period s , meaning that the system "forgets" past warming more quickly and places greater weight on recent temperature changes. As a result, faster memory decay (with a higher $\bar{\chi}$) weakens the shadow price of moving-average warming $\lambda_{\overline{at}}$, since damages from rapid warming are incorporated and adjusted for more quickly. By contrast, stronger damage sensitivity or higher output amplifies it; the termination shock leads to a higher marginal damage channel. Again, in line with other costate equations, the frontloading term $-\eta_{\overline{at}}$ shows that Stage 1 shadow prices already incorporate the future damages that will occur in Stage 2.

Finally, for the costate of the shock probability,

$$\dot{\lambda}_z = (\rho + \eta)\lambda_z - U(c_1)L - \eta * \eta_z \quad [58]$$

$$\dot{\eta}_z = \rho\eta_z - U(c_2)L \quad [59]$$

where $e^{-\rho s}[\lambda_z(s) - \eta_z(s, s)]$ measures that, compared to the case of a termination shock that happened in period s , the additional welfare of no termination shock happened before period s . For all derivations of the co-state dynamics, we refer to the Appendix.

3.2 SCC and SCF

The *Social Cost of Carbon* (SCC) is defined as the Marginal Rate of Substitution (MRS) of consumption for carbon emissions: it measures how much consumption one is willing to give up to reduce CO2 emissions by 1 unit. Then, whenever s_1^*, s_2^* are inner controls,

$$SCC_1 = -\frac{1}{c_{o2c}} \cdot \frac{\sum_j (\lambda_{Rj} \cdot \xi_j) + \lambda_{cacc}}{\lambda_K}, \quad SCC_2 = -\frac{1}{c_{o2c}} \cdot \frac{\sum_j (\eta_{Rj} \cdot \xi_j) + \eta_{cacc}}{\eta_K} \quad [60]$$

for each stage, respectively (proved in Appendix).

Correspondingly, we can define the *Social Cost of abatable non-CO2 GHGs emission*, $SCC_{E_{GHG}}$, as the marginal rate of substitution between consumption and abatable non-

CO2 GHGs emission (the E_{GHG} after abatement), i.e., the amount of consumption society is willing to forgo to reduce one unit of E_{GHG} ,

$$SCC_{E_{GHG},1} = -\psi_1 \frac{\lambda_{F_{GHG}}}{\lambda_K}, \quad SCC_{E_{GHG},2} = -\psi_1 \frac{\eta_{F_{GHG}}}{\eta_K} \quad [61]$$

for each stage, respectively (proved in Appendix).

Recall the first-order conditions (FOC) of μ , with marginal abatement cost $MAC = \frac{Y}{\sigma Y + E_{land}} \Lambda_\mu + \frac{\lambda_{F_{GHG}}}{\lambda_K} \cdot \frac{\psi_1 E_{GHG}}{\sigma Y + E_{land}}$ (proved in Appendix), which measures the optimal carbon tax (Barrage, 2020; Gerlagh & Liski, 2018a, 2018b), i.e., the economic cost of reducing emissions by one additional marginal unit of CO2. We obtain the relationship between MAC and SCC (inner controls).

$$MAC_1 = SCC_1 \quad [62]$$

$$MAC_2 = SCC_2 \quad [63]$$

In the same spirit, we define the *Social Cost of Radiative forcing*, SCF , as the marginal rate of substitution between consumption and radiative forcing (after SRM), i.e., the willingness to pay in terms of consumption to reduce radiative forcing by one marginal unit.

$$SCF_1 = -\frac{\sum_j \left(\lambda_{Tat_{i,1}} \frac{\chi_i}{\kappa_i} \right)}{\lambda_K}, \quad SCF_2 = -\frac{1}{co2c} \cdot \frac{\sum_j \left(\eta_{Tat_{i,2}} \frac{\chi_i}{\kappa_i} \right)}{\eta_K} \quad [64]$$

Based on FOC of g , for Stage 1, we also derive,

$$SCF_1 = \frac{\Omega_{g_1} Y_1}{TF(R_{j,1}, F_{GHG_1})} = MFC_1 \quad [65]$$

Here, $\frac{\Omega_{g_1} Y_1}{TF(R_{j,1}, F_{GHG_1})}$ is equal to MFC , the marginal cost of reducing radiative forcing (proved in Appendix), which measures the economic cost of a one-unit reduction in radiative forcing.

Recalling the FOC of μ and g , for Stage 1, we have,

$$\frac{\Lambda_{\mu_1}}{\Omega_{g_1}} = -\frac{dg_1}{d\mu_1} = \frac{(\sigma Y_1 + E_{land}) \cdot SCC_1 + E_{GHG} \cdot SCC_{E_{GHG},1}}{TF(R_{j,1}, F_{GHG_1}) \cdot SCF_1} \quad [66]$$

Recalling the definition of SCC , $SCC_{E_{GHG}}$, and SCF , then gives

$$SCF = \frac{\partial C}{\partial F} = \frac{\partial C}{\partial E_{co2}} \cdot \frac{\partial E_{co2}}{\partial F} = SCC \cdot \frac{\partial E_{co2}}{\partial F} \quad \text{and} \quad SCF = \frac{\partial C}{\partial F} = \frac{\partial C}{\partial E_{GHG}} \cdot \frac{\partial E_{GHG}}{\partial F} = SCC_{E_{GHG}} \cdot \frac{\partial E_{GHG}}{\partial F} \quad [67]$$

where $\frac{\partial F}{\partial E_{CO2}}$ and $\frac{\partial F}{\partial E_{GHG}}$ represent the marginal contribution to radiation from an additional unit of CO2 and an additional unit of non-CO2 GHGs, respectively. Then,

$$-\frac{dg_1}{d\mu_1} = \frac{\frac{\partial F_1}{\partial E_{CO2,1}}(\sigma Y_1 + E_{land}) + \frac{\partial F_1}{\partial E_{GHG}}E_{GHG}}{TF(R_{j,1}, F_{GHG_1})} \quad [68]$$

yielding

$$TF(R_{j,1}, F_{GHG_1})dg_1 = \frac{\partial F_1}{\partial E_{CO2,1}}(\sigma Y_1 + E_{land})d\mu_1 + \frac{\partial F_1}{\partial E_{GHG}}E_{GHG}d\mu_1 \quad [69]$$

This condition equates the marginal effect of one unit of SRM on radiative forcing (left-hand side) to the marginal effect of one unit of abatement through reductions in CO2 and non-CO2 GHG emissions (right-hand side). It provides a clear interpretation of SRM and abatement as substitutable instruments in reducing radiative forcing at the margin.

4. Numerical Results

We start with a “*myopic*” case, where the decision-maker (DM) does not account for the termination shock risk. This case is compared against an *Abatement-Only* scenario (i.e., no SRM deployment), to illustrate how a short-sighted planner allocates mitigation and SRM efforts when future risks are ignored. The *Abatement-Only* scenario corresponds to the standard solution of the DICE model. In this case, there is no termination shock risk at all because the SRM is not available.

We then turn to the forward-looking *smart* DM, where the planner internalizes the termination shock risk and proactively adjusts both abatement and SRM strategies accordingly. Although the myopic scenario has been featured in previous literature (Kosugi, 2011), those studies typically overlook an important endogenous channel of shock impacts: the scale of SRM deployment before the shock directly determines the magnitude of damages after. The more heavily SRM is used, the more severe the warming and economic losses once it is abruptly terminated. Besides, the *myopic* case provides a comparable benchmark for evaluating the “*smart*” case.

Figures 1 to 4 display the key outcomes along time from 2020 to 2100: The red dashed line with square marks represents the *Abatement-Only* solution; the black solid line is the optimal solution of the model without a risk of an SRM termination shock, which is also the *pre-shock path* (stage 1 model) under *myopic* case; the *grey-shaded area* captures the uncertainty range of solutions, i.e. the area envelopes of the solution paths for different times of the termination shock; the blue solid line in each figure shows the *expected path* of stochastic model under *myopic* and *smart* case, respectively; *percentile* bands highlight the timing of potential termination shocks. Here, the dashed lines in each panel mark quantiles for when the termination shock is likely to occur. For example, the purple dashed line in Figure 1a shows a 10% shock probability by 2040 (under the assumption of a 0.5% shock probability each year). The percentile paths also show the dynamics before and after the shock. For example, abandoning SRM after a shock occurs suddenly.

4.1 Myopic Case

Before introducing the stochastic shock outcomes, we begin by comparing the myopic pre-shock path with the *abatement-only* scenario. This comparison shows how the short-sighted behavior of the myopic DM heavily relies on SRM, which corresponds to existing findings in the SRM literature assuming no shock (Heutel et al., 2018).

Before the shock, the myopic DM increases SRM deployment continuously, reaching about 60% by 2100, while the level radiative forcing offset by SRM is about 3.6 W/m^2 (**Figure 1**). Emission control under the myopic path remains significantly lower than under abatement-only for most of the century. The abatement-only path shows a steady decline from around 35 Gt CO₂/year. In contrast, the myopic DM steadily increases emissions year by year over the entire century, from 35 Gt CO₂/year to 80 Gt CO₂/year, if no shock happens. However, due to the aggressive SRM masking of warming, the myopic case maintains lower net radiative forcing caused by GHGs and lower warming than abatement-only, near 3 W/m^2 and 1.5°C throughout the century,

despite emitting more. Besides, the myopic DM improves economic output by avoiding early mitigation costs and utilizing cheap SRM, ending the century with about 807 trillion USD, compared to abatement-only.

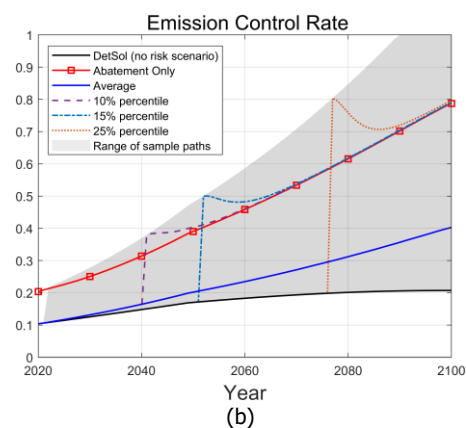
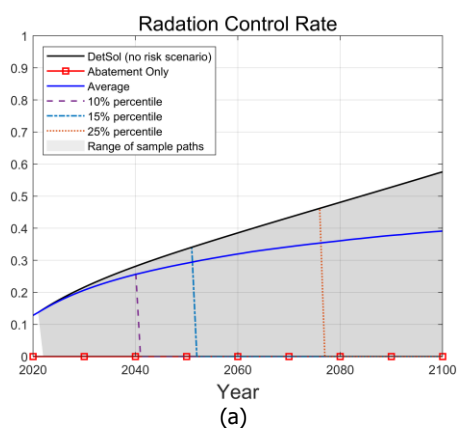
The pre-shock trajectory creates a fragile path for humanity that depends entirely on uninterrupted SRM. **Once the termination shock occurs**, the climate- economy system undergoes an abrupt and destabilizing adjustment. Upon shock, the radiation control rate drops to zero instantly, exposing the previously masked warming. The shaded area in the figure shows the ranges of possible post-shock paths. The shock leads to an abrupt increase in radiative forcing, triggering immediate warming and reactive mitigation. The severity of warming mainly depends on the amount of atmospheric carbon accumulated and the level of SRM deployed at the moment of the shock. Under the myopic path, SRM usage increases rapidly over time, masking ever-growing radiative forcing. As a result, a later shock induces a more severe climatic and economic response.

The emission control rates surge as the DM is forced to compensate for the loss of SRM. If the termination occurs near the end of the century, the system has become so dependent on SRM that it must immediately transition to net-zero emissions in order to mitigate the sudden warming damages of shock. In the decades that follow, emissions gradually converge toward the abatement-only path, as the economy adjusts to a new regime of SRM-free under strict mitigation. However, even after this adjustment, the post-shock world without SRM results in higher warming and lower economic output compared to the abatement-only scenario.

Temperature, previously stabilized near 1.5°C through SRM, now reacts to the unmasked radiative forcing. The global temperature increases rapidly after the shock within the 10 years, clearly exceeding the abatement-only path and pushing the system into a high climate risk regime. Peak warming can reach 2.8°C by 2100. The damages reflect both the legacy of higher cumulative carbon emissions under the myopic case and the economic costs of abrupt low-carbon transition, including overcompensated mitigation and lost output from climate shocks.

Economically, the termination shock induces a significant and persistent loss in output. While GDP was higher under the myopic path before the shock, the sudden policy and climate adjustment leads to a sharp output drop. Although it gradually recovers, it never catches up to the abatement-only path by 2100. The post-shock distribution of GDP under the myopic DM spans from 665 to 807 trillion USD.

From the perspective of the expected path, the myopic DM consistently emits more CO₂ than the abatement-only scenario, with the annual emissions gap steadily widening from about 6 Gt CO₂ at the beginning of the period to around 40 Gt CO₂ by 2100. Despite this sustained excess in emissions, the average global warming remains lower under the myopic path, due to the effect of SRM in masking the warming. The expected economic output under the myopic DM initially lies slightly above the abatement-only scenario, reflecting avoided early mitigation costs and relative-low early shock damages. However, as termination shock risk accumulates and the cumulative probability of abrupt SRM failure increases, these initial gains are gradually eroded. Over time, the expected GDP path converges toward the abatement-only trajectory, and by the end of the century it falls below the abatement-only level. This dynamic illustrates the long-run instability and adjustment cost of short-term gains of SRM reliance: while SRM temporarily allows for relaxed mitigation and lower temperatures, the risk of abrupt failure and the associated economic adjustment erode the earlier benefits.



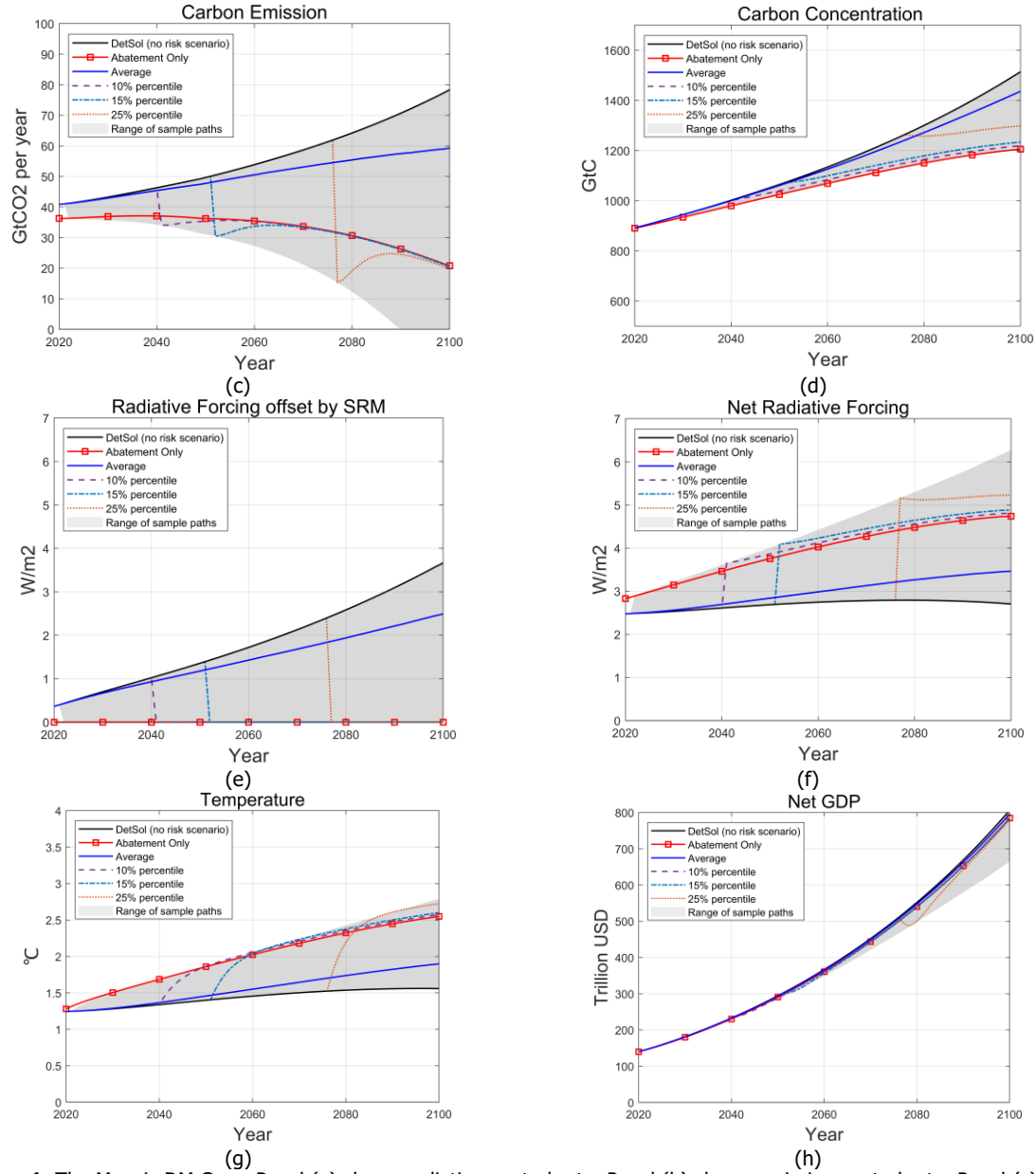


Figure 1. The Myopic DM Case. Panel (a) shows radiation control rate. Panel (b) shows emission control rate. Panel (c) shows annual carbon emissions (Gt CO₂). Panel (d) shows atmosphere carbon concentration (Gt C). Panel (e) shows the level of radiative forcing offset by SRM (W/m²). Panel (f) shows net radiative forcing caused by GHGs under SRM deployment (W/m²). Panel (g) shows temperature anomaly (°C) (increase in atmospheric temperature above the pre-industrial level). Panel (h) shows net output available for consumption or savings (Trillion USD). In each panel, the red dashed line with mark represents the Abatement-Only solution. The black line represents the solution before shock under myopic decision maker who ignores the shock risk, which is also the optimal solution under the deterministic model without the shock risk. The blue line represents the expected path from the stochastic model. The grey-shaded area shows the range of myopic paths from simulations of the optimal stochastic model's solution. The percentile paths measure the cumulative probabilities of the shock event by a certain period, as well as the optimal response after the shock event. For example, the 10% quantile path indicates a 10% probability of a shock having occurred by 2040.

Figure 2 illustrates how the four marginal indicators jointly determine optimal mix of abatement and SRM. The cheap SRM suppresses both the shadow value of radiative-forcing reduction and the incentive for carbon abatement before the shock, whereas termination removes this substitution mechanism and sharply increases the marginal value of mitigation.

Figure 2 illustrates how the four marginal indicators jointly determine the optimal mix of abatement and SRM. **Before termination**, starting from the Abatement-Only allocation, once SRM becomes available, the DM increases SRM deployment as long as its marginal radiation-control cost (MFC) remains below the social cost of radiative forcing (SCF). SRM lowers the marginal value of additional radiative-forcing reduction, reducing SCF and, through its physical linkage with CO₂ forcing, also SCC, weakening the incentive for abatement. The resulting reduction in abatement, in turn, raises emissions and climate damages, increasing both SCC and SCF potentially. Finally, the new equilibrium under the Myopic DM is established until $MFC = SCF$ and $MAC = SCC$, while SCC and SCF remain physical relationship. In 2020, the new equilibrium SCC and MAC are about US\$18/tCO₂, compared with roughly US\$60/tCO₂ under Abatement-Only. Similarly, the equilibrium SCF and MFC are about 0.4 trillion US\$/W/m², slightly below the Abatement-Only SCF of about 0.5 trillion US\$/W/m².

After termination, SRM is no longer available, so the $MFC = SCF$ condition disappears. And the system shifts from a two-policies equilibrium to an abatement-only adjustment. The loss of SRM exposes previously masked radiative forcing, causing SCF and SCC to rise sharply. The higher SCC strengthens the optimal incentive for abatement, while stronger mitigation subsequently limits emissions and future damages. In the new post-shock equilibrium, only the condition $MAC = SCC$ remains operative. Depending on shock timing, termination raises the SCF to roughly trillion US\$ 1–50 /W/m² and the SCC/MAC to about US\$ 60–600/t CO₂.

In summary, the termination shock reveals the fragility of the myopic strategy: what appears to be a low-cost and climate-effective situation pre-shock turns into a high-risk, high-volatility scenario once SRM fails, with greater post-shock losses than in abatement-only strategy.

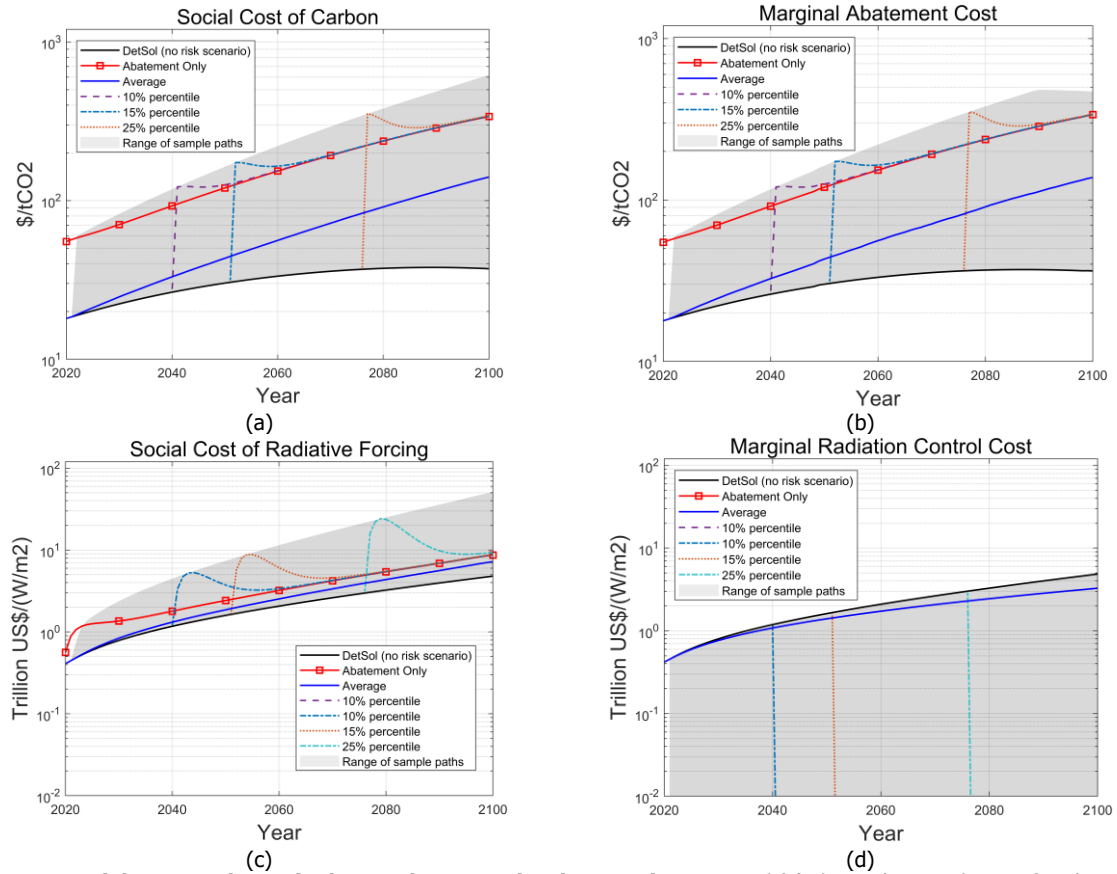


Figure 2. Social costs and marginal control costs under the Myopic DM. Panel (a) shows the Social Cost of Carbon (SCC, $\$/tCO_2$). Panel (b) shows the Marginal Abatement Cost (MAC, $\$/tCO_2$). Panel (c) shows the Social Cost of Radiative Forcing (SCF, trillion $US\$/W/m^2$). Panel (d) shows the Marginal Radiation Control Cost (MFC, trillion $US\$/W/m^2$). The red line with square markers represents the Abatement-Only scenario, the black line represents the deterministic solution without termination risk, and the blue line represents the expected path under stochastic termination. The grey area shows the range of simulated paths, while the dashed lines indicate selected percentile shock paths. The interpretation of the percentile paths follows Figure 1.

4.2 Smart Case

Next, we turn to the *smart* case. Compared to the myopic case, the Smart DM, who recognizes that excessive SRM would amplify the severity of a potential termination shock, deploys moderate levels of SRM before shock and implements stronger carbon abatement as a form of insurance against the shock risk. By actively restraining SRM deployment, the Smart DM limits the magnitude of potential warming spikes and economic disruptions in the event of shock. Although this approach requires higher near-term abatement efforts and accepts somewhat higher pre-shock warming compared to the myopic path, it ensures a more stable long-run outcome and narrows the range of post-shock losses.

Figure 3 shows the smart DM employs SRM in a restrained and cautious manner. **Before the shock**, the radiative forcings offset by SRM rise slowly, reaching less than 2.5 W/m^2 by 2100 (about 45% of the total radiative forcing from GHGs), substantially lower than in the myopic case (about 3.7 W/m^2). **After the shock**, SRM deployment level drops immediately to zero under both Smart and myopic paths. The difference is that Smart DM's more restrained SRM use beforehand makes the adjustment much less severe compared to the myopic case.

The Smart DM implements much stronger emission control than the myopic DM **before the shock**. While the myopic path keeps control rates low, the Smart DM steadily increases abatement, reaching about 50% by 2100. Nevertheless, this rate remains below the abatement-only scenario in the pre-shock period, due to the moderate SRM deployment. **After the shock**, however, the Smart DM's abatement rate also quickly jumps above the level of abatement-only, in order to compensate for the cooling loss from SRM termination. Importantly, the peak warming (the upper bound of uncertainty range) under the Smart DM remains lower than in the myopic case, indicating that proactive abatement reduces the climate volatility.

As a result, Smart DM emissions are consistently much lower than myopic DM emissions. Although still slightly above abatement-only **before the shock**, emissions fall sharply **after the shock** and often undershoot the abatement-only level. From expected path perspective, the smart DM also emits more CO_2 than the abatement-only scenario, with the annual emissions gap widening from about 2 to 20 Gt CO_2 , roughly half the increase observed in the myopic case (6–40 Gt CO_2 annually under myopic DM). Abatement efforts limit the rate of atmospheric carbon accumulation, less than 1350 Gt C by 2100 before the shock, compared with over 1,500 Gt C in the myopic case.

Despite stronger abatement, the reduced reliance on SRM means that net radiative forcing under the Smart DM path is higher than in the myopic case before the shock, since less of the GHG forcing is artificially masked. This requires accepting somewhat higher pre-shock warming. **Before the shock**, the warming rises to about 1.7°C ,

compared with the 1.5°C in the myopic case. After the shock, peak warming under the Smart DM is near the abatement-only scenario (about 2.5°C). However, this preventive strategy also delivers a crucial benefit: compared with the myopic case, the higher pre-shock temperature path, together with lower atmospheric carbon concentration that results in a lower possible upper bound of warming, significantly narrows the magnitude of the temperature surge and the associated damages in post-shock period. The temperature surge caused by termination shock under the Smart DM is limited to about 0.8 °C even in the late century, whereas the myopic DM suffers a much larger increase of over 1.2 °C. Thus, the Smart DM deliberately accepts a sort of “temperature” smoothing insurance, i.e., a modest increase in pre-shock warming in exchange for substantially reducing post-shock damages.

The Smart DM’s GDP trajectory is also more optimistic. Although output declines following the termination shock, the economy has the potential to recover to a level close to that of the abatement-only scenario — an outcome that is not achieved under the myopic strategy. In other words, the Smart DM sacrifices a modest share of short-term economic gains to secure greater long-run resilience. Moreover, the uncertainty range of GDP under the Smart DM is narrower, spanning 718 to 803 trillion USD by 2100, compared with the much wider range of 665 to 807 trillion USD and lower expected level in the myopic case. This reduced range is a direct consequence of the smaller temperature surge of the shock, which significantly lowers the need for abrupt and costly adjustments.

From the perspective of the expected path, the Smart DM’s economic output is roughly comparable to that of the abatement-only scenario. This reflects the benefits of proactive action before the shock, although proactive mitigation entails higher near-term costs, it reduces long-run risks and volatility, leaving overall output close to the abatement-only level. The Smart DM thus secures both a more stable climate state and economic efficiency, underscoring the value of prevention against termination shock.

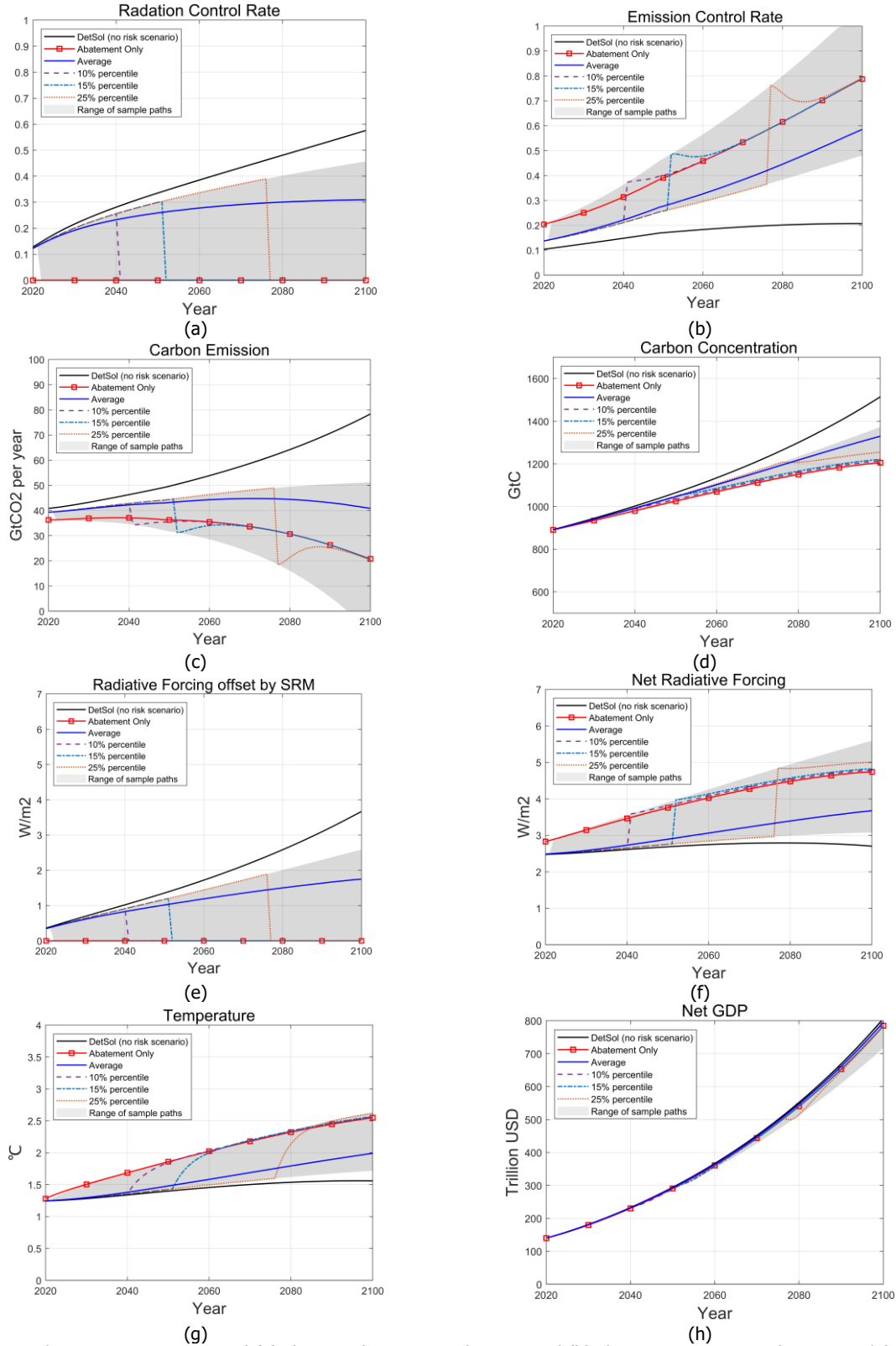
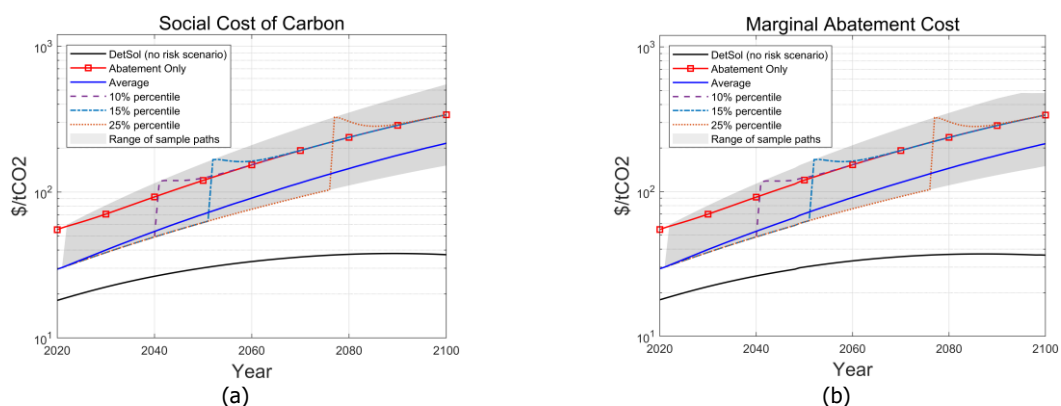


Figure 3. The Smart DM Case. Panel (a) shows radiation control rate. Panel (b) shows emission control rate. Panel (c) shows annual carbon emissions (Gt CO₂). Panel (d) shows atmosphere carbon concentration (Gt C). Panel (e) shows the level of radiative forcing offset by SRM (W/m²). Panel (f) shows net radiative forcing caused by GHGs under SRM deployment (W/m²). Panel (g) shows temperature anomaly (°C) (increase in atmospheric temperature above the pre-industrial level). Panel (h) shows net output available for consumption or savings (Trillion USD). In each panel, the red dashed line with mark represents the Abatement-Only solution. The black line represents the solution before shock under myopic decision maker who ignores the shock risk, which is also the optimal solution under the deterministic model without the shock risk. The blue line represents the expected path from the stochastic model. The grey-shaded area shows the range of smart paths from 10,000 simulations of the optimal stochastic model's solution. The percentile paths measure the cumulative probabilities of the shock event by a certain period, as well as the optimal response after the shock event.

Introducing termination risk changes the relative marginal value of the two climate policies. **Before termination**, starting from the Myopic equilibrium, internalizing termination risk raises the expected marginal damages represented by SCC and SCF, disturbing the original Myopic equilibrium, leading $SCC > MAC$ and $SCF > MFC$.

Under termination risk, unlike in the Myopic case, additional SRM no longer necessarily lowers SCC and SCF and even increase them. This is because SRM has two opposing effects. Its cooling effect tends to reduce SCC and SCF before termination, whereas greater SRM also increases masked warming and exposure to a future termination shock, thereby raising SCC and SCF. By contrast, additional abatement raises MAC while permanently reducing emissions and radiative forcing in both the pre- and post-termination states, thereby lowering SCC and SCF. Thus, greater abatement raises MAC while reducing SCC and SCF, and the associated decline in emissions and residual radiative forcing reduces the need for SRM. Lower SRM deployment, in turn, reduces MFC, SCC and SCF. The Smart DM adjusts both controls toward more abatement and less SRM until the new equilibrium conditions are restored, with higher SCC/MAC (US\$30/tCO₂ initially) but slightly lower SCF/MFC than under the Myopic equilibrium.

After termination, SRM is no longer available. The loss of radiative-forcing control causes that SCC and SCF rise sharply following the shock, but the increases are generally weaker under the Smart DM. The lower post-shock SCC therefore implies a lower optimal level of abatement relative to the Myopic case.



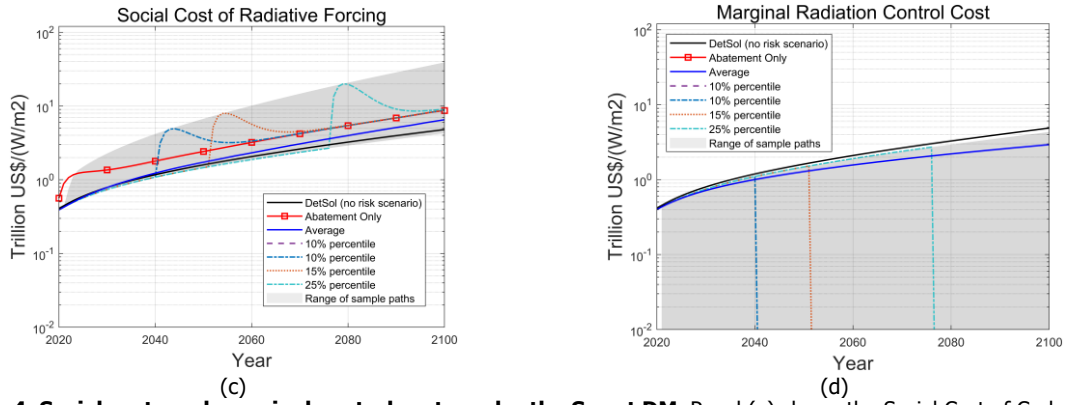


Figure 4. Social costs and marginal control costs under the Smart DM. Panel (a) shows the Social Cost of Carbon (SCC, \$/tCO₂). Panel (b) shows the Marginal Abatement Cost (MAC, \$/tCO₂). Panel (c) shows the Social Cost of Radiative Forcing (SCF, trillion US\$/W/m²). Panel (d) shows the Marginal Radiation Control Cost (MFC, trillion US\$/W/m²). The red line with square markers represents the Abatement-Only scenario, the black line represents the deterministic solution without termination risk, and the blue line represents the expected path under stochastic termination of smart DM. The grey area shows the range of simulated paths, while the dashed lines indicate selected percentile shock paths. The interpretation of the percentile paths follows Figure 3.

5. Sensitivity Analysis

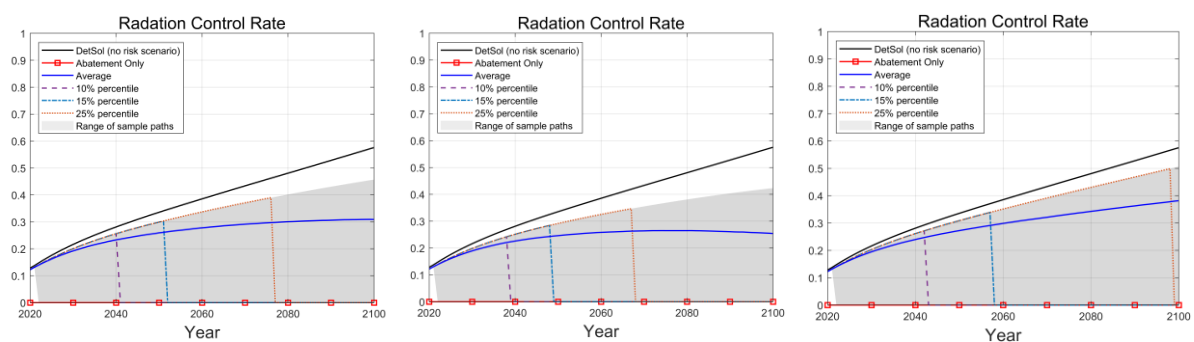
In our baseline setup, the termination shock is modeled as a stochastic event with an exogenous hazard rate. The hazard rate of a termination shock and the magnitude of the damages following such an event jointly determine the severity and impact of termination risk, and thus play a critical role in shaping the optimal climate policy. However, since neither the probability of occurrence nor the scale of post-shock damages can be empirically observed, both parameters are treated here as modeling assumptions and remain uncertain (Ekholm, 2018; Tang & Kemp, 2021). Therefore, we conduct sensitivity analyses by testing alternative assumptions about the likelihood of a shock (hazard rate parameters η) and the severity of damages (parameters β_1 , β_2), examining how these variations influence the outcomes under smart DM.

5.1 The Shock Probability

We consider three alternative specifications for the hazard rate of the termination shock. (a) **constant hazard rate** case (baseline), $\eta = 0.5\%$, where the hazard rate of shock is fixed at 0.5%/yr over time. (b) **increasing hazard rate** case, $\eta = 0.5\% + 0.45\% \cdot (1 - e^{-0.01t})$, where the hazard rate starts at 0.5%/yr and reaches about 0.75% per year by 2100. (c) **decreasing hazard rate** case, $\eta = 0.5\% + 0.45\% \cdot e^{-0.01t}$, where the hazard rate starts from 0.5%/y and reaches about 0.25%/y by 2100).

These three hazard rate trends represent alternative assumptions about how DM may perceive the evolution of termination shock risk over time. The constant hazard reflects a belief that the risk remains unchanged throughout the century. Both the increase and decrease in the hazard rate are related to some fundamentals. The decreasing hazard corresponds to the view that the likelihood of a shock declines as technology or governance capacity improves— reflecting a learning-by-doing process where cumulative SRM deployment experience serves as a proxy for enhanced capability— eventually approaching a lower bound. By contrast, the increasing hazard captures the perception that the risk accumulates over time, rising as large-scale human interventions in the Earth system deepen dependence on SRM and potentially generate hidden vulnerabilities.

The sensitivity analysis for hazard rate shows that the Myopic and Smart DM’s strategy is robust in its overall structure (Figure 5). Changes in the hazard rate have no effect on the Myopic DM, because it does not anticipate termination risk when making current decisions. For the Smart DM, before the shock, a higher perceived termination risk leads to more abatement and less SRM relative to the baseline case, reducing the temperature rebound and associated damages. After the shock, the Smart DM generally undertakes less abatement than in the baseline case, because the shock has already reduced GDP and therefore limits the economic resources available for additional abatement. A lower hazard rate produces the opposite response.



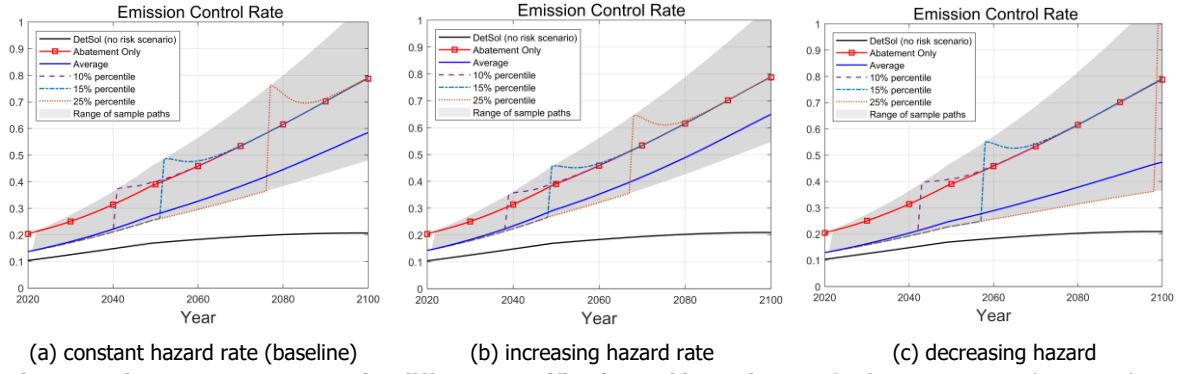


Figure 5. The Smart DM case under different specifications of hazard rate. The first row reports the optimal SRM deployment (radiation control rate), and the second row reports the corresponding optimal abatement effort (emission control rate). Panel (a) shows the baseline case with a constant hazard rate, where the hazard rate of shock is fixed over time. Panel (b) illustrates an increasing hazard rate case, where the hazard rate rises gradually with time. Panel (c) presents a decreasing hazard rate case, where the hazard rate declines over time. The legend follows the same convention as baseline of Smart DM (Figure 2): In each panel, the red dashed line with mark represents the Abatement-Only solution. The black line represents the solution before shock under myopic decision maker who ignores the shock risk, which is also the optimal solution under the deterministic model without the shock risk. The blue line represents the expected path from the stochastic model. The grey-shaded area shows the range of paths from simulations of the optimal stochastic model's solution. The percentile paths measure the cumulative probabilities of the shock event by a certain period.

5.2 The Shock Damage

To evaluate the sensitivity of outcomes to assumptions about termination shock damages, we distinguish between **two components of damages**. The first is **short-term damages**, which occur only during the years immediately following the shock, reflecting the catastrophic impacts of a sudden temperature surge. These damages are severe but transitory. The second is **long-term damages**, which capture the persistent economic losses associated with the higher warming level after the shock.

Here, in order to provide a more comprehensive picture, we analyze these two dimensions simultaneously by comparing four combinations according to the severity of long-term and short-term damages. This approach allows us to compare how the combination of short-term termination shock damages and long-term climate damages shapes the overall climate–economy outcomes, thereby offering a panoramic view of the risks faced by different SRM deployment strategies.

The *first element* of each label refers to long-term damages, and the *second element* refers to short-term damages. **(a) Low-Low**: the baseline case, where the long-term damage parameter (β_1) follows the standard DICE specification and the short-term damage parameter (β_2) is set to its benchmark value 1%. **(b) Low-High**, while the

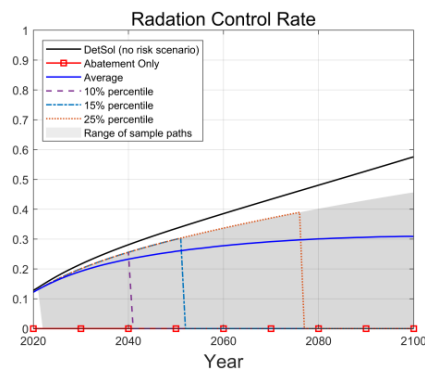
short-term parameter is set to **twice** the baseline level. **(c) High–Low:** the long-term damage parameter is set to **three times** the DICE specification. **(d) High–High:** both long-term and short-term damages are scaled up, with β_1 equal to three times DICE and β_2 equal to two times the baseline. This design allows us to disentangle the effects of catastrophic but temporary damages from abrupt warming, and the persistent damages of global warming, while also exploring their combined impact.

The sensitivity analysis for shock damages further highlights the difference between the Myopic and Smart DMs. For the **Myopic DM**, in general, higher long-term damages induce less abatement and greater SRM deployment before the shock, reflecting a stronger reliance on SRM, while higher short-term damages have little effect before the shock. After termination, however, higher damages lead to stronger abatement efforts to reduce post-shock losses.

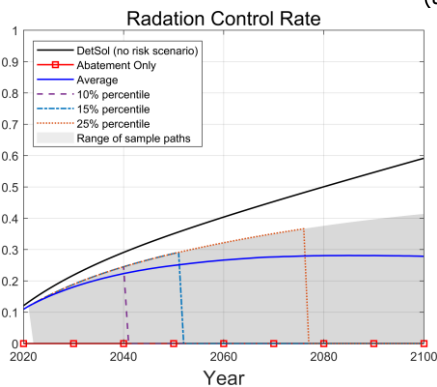
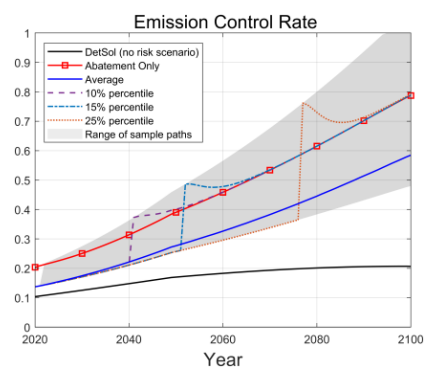
We then focus on the **Smart DM**, whose strategy remains robust across alternative damage assumptions (Figure 6). Comparing Panel (a) (low–low, baseline) with Panel (b) (low–high) isolates the effect of higher short-term damages. Before the shock, the Smart DM deploys less SRM and increases abatement than in the baseline case, which means that the Smart DM becomes more cautious in SRM deployment when short-term damages are higher. The deterministic paths are almost unchanged, indicating that the higher short-term damages have virtually no effect on results of the model assuming no termination shock (such as the standard DICE).

Comparing Panel (a) with Panel (c) (high–low) isolates the effect of higher long-term damages. Relative to the baseline, before the shock, SRM deployment rises more sharply than in the baseline case, while abatement also increases and approaches net-zero earlier. This is because higher long-term damages raise the economic cost of warming before and after shock, strengthening the incentive to use both policies. After the shock, abatement also increases relative to the baseline to limit the larger post-shock damages. Thus, both higher short-term and higher long-term damages strengthen abatement, but their effects on SRM differ: higher short-term damages reduce SRM deployment, whereas higher long-term damages increase it.

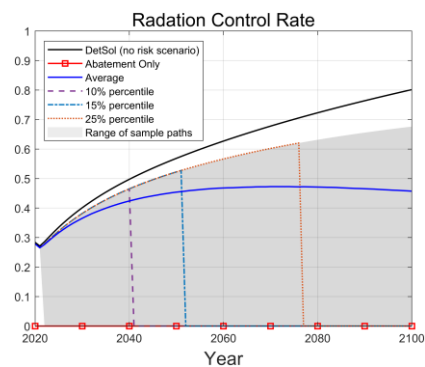
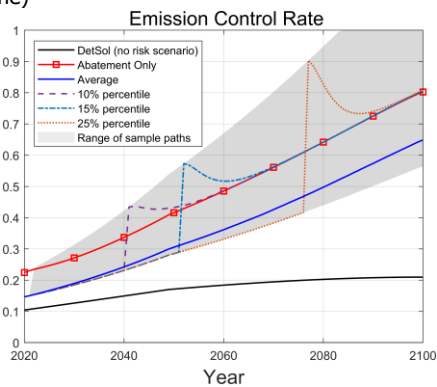
Panel (d) (high–high) combines higher short-term and long-term damages. It resembles the high–low case more closely, with the Smart DM relying more intensively on both SRM and abatement. Since both damage components independently strengthen the incentive for mitigation, the abatement rate rises more sharply before and after the shock. SRM deployment also increases relative to the baseline before and after the shock, indicating that the long-term damage effect dominates the SRM response.



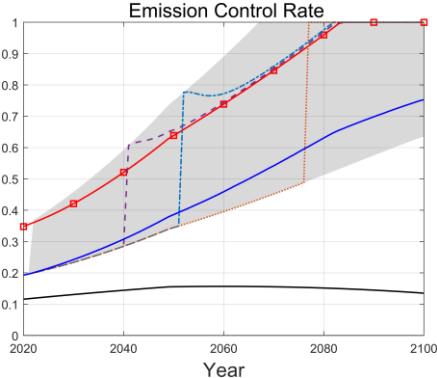
(a) low-low (baseline)



(b) low-high



(c) high-low



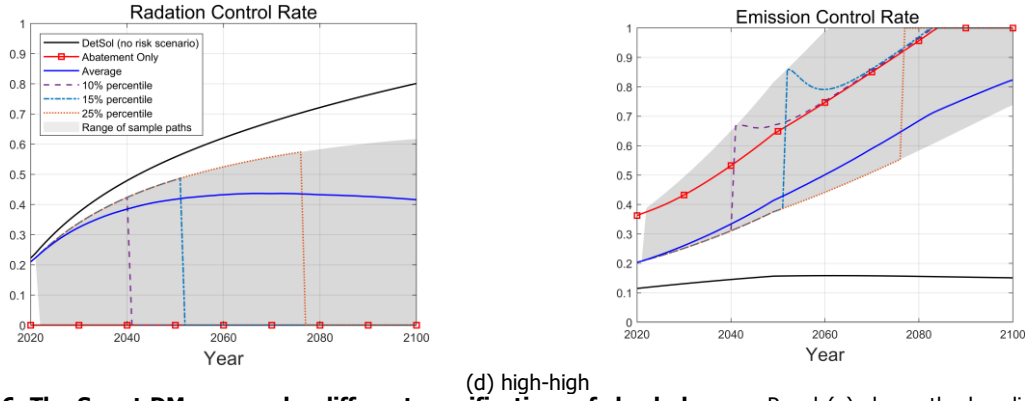


Figure 6. The Smart DM case under different specifications of shock damage. Panel (a) shows the baseline low–low case, where both long-term damages and short-term damages take their benchmark values. Panel (b) illustrates the low–high case, where the short-term damage parameter is doubled relative to the baseline while the long-term damage parameter remains at the DICE benchmark. Panel (c) presents the high–low case, where the long-term damage parameter is scaled up to three times the DICE value while the short-term damage remains at the baseline level. Panel (d) presents the high–high case, where both long-term and short-term damages are scaled up, with long-term damage parameter equal to three times DICE and short-term damage parameter equal to two times the baseline. The legend follows the same convention as baseline of Smart DM (Figure 2): In each panel, the red dashed line with mark represents the Abatement-Only solution. The black line represents the solution before shock under myopic decision maker who ignores the shock risk, which is also the optimal solution under the deterministic model without the shock risk. The blue line represents the expected path from the stochastic model. The grey-shaded area shows the range of paths from simulations of the optimal stochastic model's solution. The percentile paths measure the cumulative probabilities of the shock event by a certain period.

Overall, the key differences across damage cases arise mainly before the shock. Higher short-term damages have little effect on the Myopic DM but induce the Smart DM to increase abatement and reduce SRM, while higher long-term damages increase SRM deployment for both DMs but have opposite effects on abatement. After the shock, both DMs rely on stronger abatement to limit post-shock damages. Those results highlight distinct roles of short-term and long-term damages in shaping optimal SRM deployment.

6. Summary

Rapid global warming has intensified interest in SRM. However, SRM does not address the root cause of climate change (GHGs accumulation) and introduces a critical concern, i.e., the risk of a termination shock. This proposal aims to explore the impacts of such a risk and to identify the optimal response strategy. This paper develops a dynamic stochastic integrated assessment model by endogenizing the termination shock risk SRM, which captures the dynamics of both pre-shock and post-shock. Consequently, we analyze how forward-looking planners should balance SRM and abatement when facing uncertain termination shock risks. We further incorporate

sensitivity analyses with respect to the shock probability and damage specifications, disentangling the roles of short-term catastrophic impacts and long-term persistent damages. In doing so, the paper provides a unified framework to evaluate how different risk specifications shape optimal climate policy before and after the shock.

In the theoretical analysis, we apply the Maximum Principle for age-structured optimal control problems to the proposed IAM that endogenizes termination shock risk. This framework reveals how termination shocks in **Stage 2** influence prior decisions and system dynamics in **Stage 1**. Furthermore, by deriving the social cost of carbon (SCC) and the social cost of radiative forcing (SCF) under termination shock risk, we identify the substitutive relationship between the optimal SRM policy and the abatement policy. Specifically, the analysis shows that the marginal reduction in radiative forcing achieved by one unit of SRM should be equated with the marginal reduction in radiative forcing from one unit of abatement through cuts in CO₂ and non-CO₂ GHG emissions.

The numerical results show that a myopic decision-maker, who ignores the risk, relies heavily on SRM and delays abatement, creating a fragile situation that collapses into severe warming and economic losses once the shock occurs. By contrast, a forward-looking Smart DM, who recognizes the risk, proactively deploys less SRM and increases abatement efforts. Although this strategy accepts somewhat higher warming in pre-shock period, it substantially reduces the magnitude of the post-shock temperature surge and associated economic losses. The contrast between the two cases carries an important policy message: better information and anticipation of termination risk can be interpreted as a shift from myopic to smart decision-making, which effectively allows decision-makers to self-insure against termination shock by adjusting SRM and abatement policies in a precautionary way.

Sensitivity analyses further highlight how outcomes depend on assumptions about the shock probability and damage specification. Varying the probability and damages of the termination shock does not alter the qualitative structure of the Smart DM's strategy. Compared to a myopic DM, the Smart DM relies on limited SRM only and

rather on stronger abatement prior to the shock, while both DMs respond to SRM termination by swiftly intensifying abatement to compensate the lost cooling effect. Varying the probability and damages shifts the timing and intensity of the policy adjustments. Overall, changes in termination probability have little effect on the Myopic DM but induce the Smart DM to increase abatement and reduce SRM when the perceived risk is higher. Higher short-term damages show a similar pattern, mainly affecting the Smart DM through more abatement and less SRM. In contrast, higher long-term damages increase SRM deployment for both DMs, while reducing abatement for the Myopic DM and increasing abatement for the Smart DM.

Finally, some limitations of this study should be acknowledged. First, our model does not incorporate the potential costs of rapid decarbonization. In the aftermath of termination, the Smart DM responds with a dramatic and immediate low-carbon transition. Due to the structural limitations of the DICE framework, we abstract from the economic and technological frictions that such a drastic adjustment would entail. Incorporating these costs may further strengthen earlier abatement effort. Second, our modeling of termination shocks assumes the most extreme case: a full and permanent loss of SRM's cooling effect once the shock happens. In some discussion, SRM infrastructures may experience only partial failures and could potentially be restored. The partial failure scenario is more appropriately modeled within a regional model. Future work should therefore explore how to respond to intermittent shocks. We do not model such intermittent failures, as they involve the complexity of time scales. Very short interruptions may not qualify as termination shocks, whereas prolonged suspensions could generate effects similar to those analyzed in this paper. Future research should therefore examine how different assumptions about the timing and persistence of SRM interruptions influence pre-shock strategies. Furthermore, introducing decision-making methods to address deep uncertainty in SRM is beneficial (Berger et al., 2017; Hu et al., 2012; Marinacci, 2015; Rudik, 2020).

In summary, our findings underscore the importance of anticipating termination shock risk when considering SRM as part of climate policy. While SRM can temporarily mask

warming and lower short-run abatement costs and climate damage, overreliance on it creates fragility that exposes the climate and economy to abrupt and costly adjustments once termination occurs. By contrast, incorporating termination risk into decision-making encourages preventive abatement and restrained SRM deployment, which transform short-term sacrifices into long-term resilience. The policy implication is clear. SRM, if ever considered, should remain a supplementary and strictly temporary measure, embedded within a broader strategy of deep decarbonization. Ultimately, robust climate policy must prioritize sustained emissions reductions, as only mitigation can provide durable protection against climate change.

Acknowledgments

We gratefully acknowledge the support of the National Natural Science Foundation of China (grant nos. 724B2005, 72488101, and 72293603). We also gratefully acknowledge funding from International Institute for Applied Systems Analysis (IIASA) and the National Member Organizations that support the institute. Views and opinions expressed are those of the author(s) only. We would like to sincerely thank the faculty members at CEEP-BIT and IIASA for their valuable guidance and continuous support. We are also grateful to participants of the 7th World Congress of Environmental and Resource Economists (WCERE 2026) and the 2026 RFF and Harvard SRM Social Science Research Workshop for their constructive comments and valuable suggestions.

References

- Acemoglu, D., & Rafey, W. (2023). Mirage on the horizon: Geoengineering and carbon taxation without commitment. *Journal of Public Economics*, 219, 104802.
- Barnett, M. (2023). Climate change and uncertainty: An asset pricing perspective. *Management science*, 69(12), 7562-7584.
- Barrage, L. (2020). Optimal dynamic carbon taxes in a climate–economy model with distortionary fiscal policy. *The Review of Economic Studies*, 87(1), 1-39.
- Barrage, L., & Nordhaus, W. (2024). Policies, projections, and the social cost of carbon: results from the DICE-2023 model. *Proceedings of the National Academy of Sciences*, 121(13), e2312030121.
- Baum, S. D., Maher Jr, T. M., & Haqq-Misra, J. (2013). Double catastrophe: Intermittent stratospheric geoengineering induced by societal collapse. *Environment Systems & Decisions*, 33(1), 168-180.
- Belaia, M., Moreno-Cruz, J. B., & Keith, D. W. (2021). Optimal climate policy in 3D: Mitigation, carbon removal, and solar geoengineering. *Climate Change Economics*, 12(03), 2150008.
- Bellamy, R., & Lezaun, J. (2017). Crafting a public for geoengineering. *Public Understanding of Science*, 26(4), 402-417.
- Berger, L., Emmerling, J., & Tavoni, M. (2017). Managing catastrophic climate risks under model uncertainty aversion. *Management science*, 63(3), 749-765.
- Bickel, J. E., & Agrawal, S. (2013). Reexamining the economics of aerosol geoengineering. *Climatic Change*, 119(3), 993-1006.
- Bickel, J. E., & Lane, L. (2009). An analysis of climate engineering as a response to climate change. *An analysis of climate engineering as a response to climate*

change, 59 pp.

http://fixthecclimate.com/uploads/tx_templavoila/AP_Climate_Engineering_Bickel_Lane_v.5.0.pdf

- Cai, Y. (2019). Computational methods in environmental and resource economics. *Annual Review of Resource Economics*, 11, 59-82.
- Cai, Y., & Lontzek, T. S. (2019). The social cost of carbon with economic and climate risks. *Journal of political Economy*, 127(6), 2684-2734.
- Collin-Dufresne, P., Johannes, M., & Lochstoer, L. A. (2016). Parameter learning in general equilibrium: The asset pricing implications. *American Economic Review*, 106(3), 664-698.
- Diffenbaugh, N. S., & Field, C. B. (2013). Changes in ecologically critical terrestrial climate conditions. *Science*, 341(6145), 486-492.
- Ekholm, T. (2018). Climatic cost-benefit analysis under uncertainty and learning on climate sensitivity and damages. *Ecological Economics*, 154, 99-106.
- Emmerling, J., & Tavoni, M. (2018). Climate engineering and abatement: A 'flat' relationship under uncertainty. *Environmental and Resource Economics*, 69(2), 395-415.
- Feichtinger, G., Tragler, G., & Veliov, V. M. (2003). Optimality conditions for age-structured control systems. *Journal of Mathematical Analysis and Applications*, 288(1), 47-68.
- Feichtinger, G., & Wrzaczek, S. (2024). The optimal momentum of population growth and decline. *Theoretical Population Biology*, 155, 51-66.
- Freiberger, M., Kuhn, M., & Wrzaczek, S. (2023). Modeling health shocks.
- Gerlagh, R., & Liski, M. (2018a). Carbon prices for the next hundred years. *The Economic Journal*, 128(609), 728-757.
- Gerlagh, R., & Liski, M. (2018b). Consistent climate policies. *Journal of the European Economic Association*, 16(1), 1-44.
- Ghosh, S., Bohra, A., & Dutta, S. (2021). The Texas Freeze of February 2021: Event and Winterization Analysis Using Cost and Pricing Data. 2021 IEEE Electrical Power and Energy Conference (EPEC),
- Goes, M., Tuana, N., & Keller, K. (2011). The economics (or lack thereof) of aerosol geoengineering. *Climatic Change*, 109(3), 719-744.
- Golosov, M., Hassler, J., Krusell, P., & Tsyvinski, A. (2014). Optimal taxes on fossil fuel in general equilibrium. *Econometrica*, 82(1), 41-88.
- Gramstad, K., & Tjøtta, S. (2010). Climate engineering: cost benefit and beyond.
- Hellman, M., & Cerf, V. (2021). An existential discussion: What is the probability of nuclear war. *Bulletin of the Atomic Scientists*, 18.
- Heutel, G., Moreno-Cruz, J., & Shayegh, S. (2016). Climate tipping points and solar geoengineering. *Journal of Economic Behavior & Organization*, 132, 19-45.
- Heutel, G., Moreno-Cruz, J., & Shayegh, S. (2018). Solar geoengineering, uncertainty, and the price of carbon. *Journal of Environmental Economics and Management*, 87, 24-41.
- Hu, Z., Cao, J., & Hong, L. J. (2012). Robust simulation of global warming policies using the DICE model. *Management science*, 58(12), 2190-2206.

- Iverson, T., & Karp, L. (2021). Carbon taxes and climate commitment with non-constant time preference. *The Review of Economic Studies*, 88(2), 764-799.
- Jones, A., Haywood, J. M., Alterskjær, K., Boucher, O., Cole, J. N., Curry, C. L., Irvine, P. J., Ji, D., Kravitz, B., & Egill Kristjánsson, J. (2013). The impact of abrupt suspension of solar radiation management (termination effect) in experiment G2 of the Geoengineering Model Intercomparison Project (GeoMIP). *Journal of Geophysical Research: Atmospheres*, 118(17), 9743-9752.
- Karabarbounis, L., & Neiman, B. (2014). The global decline of the labor share. *The quarterly journal of economics*, 129(1), 61-103.
- Kelly, D. L., Heutel, G., Moreno-Cruz, J. B., & Shayegh, S. (2023). Solar Geoengineering, Learning, and Experimentation. *Journal of the Association of Environmental and Resource Economists*, 0. <https://doi.org/https://doi.org/10.1086/729608>
- Kemp, L., Xu, C., Depledge, J., Ebi, K. L., Gibbins, G., Kohler, T. A., Rockström, J., Scheffer, M., Schellnhuber, H. J., & Steffen, W. (2022). Climate endgame: Exploring catastrophic climate change scenarios. *Proceedings of the National Academy of Sciences*, 119(34), e2108146119.
- Kosugi, T. (2011). Climate-economy Modeling Considering Solar Radiation Management and its Termination Risk. SIMULTECH,
- Kosugi, T. (2013). Fail-safe solar radiation management geoengineering. *Mitigation and Adaptation Strategies for Global Change*, 18(8), 1141-1166.
- Kravitz, B., & MacMartin, D. G. (2020). Uncertainty and the basis for confidence in solar geoengineering research. *Nature Reviews Earth & Environment*, 1(1), 64-75.
- Lemoine, D., & Traeger, C. (2014). Watch your step: optimal policy in a tipping climate. *American Economic Journal: Economic Policy*, 6(1), 137-166.
- Lempert, R. J., Schlesinger, M. E., Bankes, S. C., & Andronova, N. G. (2000). The impacts of climate variability on near-term policy choices and the value of information. *Climatic Change*, 45, 129-161.
- MacMartin, D. G., Ricke, K. L., & Keith, D. W. (2018). Solar geoengineering as part of an overall strategy for meeting the 1.5 C Paris target. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2119), 20160454.
- Malm, A. (2023). The future is the termination shock: on the antinomies and psychopathologies of geoengineering. Part two. *Historical Materialism*, 31(1), 3-61.
- Marinacci, M. (2015). Model uncertainty. *Journal of the European Economic Association*, 13(6), 1022-1100.
- Martin, I. W., & Pindyck, R. S. (2015). Averting catastrophes: The strange economics of Scylla and Charybdis. *American Economic Review*, 105(10), 2947-2985.
- Mattauch, L., Matthews, H. D., Millar, R., Rezai, A., Solomon, S., & Venmans, F. (2020). Steering the climate system: Using inertia to lower the cost of policy: Comment. *American Economic Review*, 110(4), 1231-1237.

- Matthews, H. D., & Caldeira, K. (2007). Transient climate–carbon simulations of planetary geoengineering. *Proceedings of the National Academy of Sciences*, 104(24), 9949-9954.
- McCusker, K. E., Armour, K. C., Bitz, C. M., & Battisti, D. S. (2014). Rapid and extensive warming following cessation of solar radiation management. *Environmental Research Letters*, 9(2), 024005.
- Moreno-Cruz, J. B., & Keith, D. W. (2013). Climate policy under uncertainty: a case for solar geoengineering. *Climatic Change*, 121(3), 431-444.
- Nordhaus, W. (2019). Climate change: The ultimate challenge for economics. *American Economic Review*, 109(6), 1991-2014.
- Okullo, S. J. (2020). Determining the social cost of carbon: Under damage and climate sensitivity uncertainty. *Environmental and Resource Economics*, 75(1), 79-103.
- Parker, A., & Irvine, P. J. (2018). The risk of termination shock from solar geoengineering. *Earth's Future*, 6(3), 456-467.
- Rudik, I. (2020). Optimal climate policy when damages are unknown. *American Economic Journal: Economic Policy*, 12(2), 340-373.
- Schneider, T., Kaul, C. M., & Pressel, K. G. (2020). Solar geoengineering may not prevent strong warming from direct effects of CO₂ on stratocumulus cloud cover. *Proceedings of the National Academy of Sciences*, 117(48), 30179-30185.
- Smith, W. (2020). The cost of stratospheric aerosol injection through 2100. *Environmental Research Letters*, 15(11), 114004.
- Sovacool, B. K., Baum, C. M., & Low, S. (2023). Risk–risk governance in a low-carbon future: Exploring institutional, technological, and behavioral tradeoffs in climate geoengineering pathways. *Risk analysis*, 43(4), 838-859.
- Tang, A., & Kemp, L. (2021). A fate worse than warming? Stratospheric aerosol injection and global catastrophic risk. *Frontiers in Climate*, 3, 720312.
- Traeger, C. P. (2023). ACE—Analytic climate economy. *American Economic Journal: Economic Policy*, 15(3), 372-406.
- Traeger, C. P., & Meier, F. D. (2023). *Uncertain Remedies to Fight Uncertain Consequences: The Case of Solar Geoengineering* (RFF Working Paper Series, Issue. <https://ideas.repec.org/p/rff/dpaper/dp-23-37.html>)
- Trisos, C. H., Amatulli, G., Gurevitch, J., Robock, A., Xia, L., & Zambri, B. (2018). Potentially dangerous consequences for biodiversity of solar geoengineering implementation and termination. *Nature Ecology & Evolution*, 2(3), 475-482.
- UNEP. (2023). *One Atmosphere: An Independent Expert Review on Solar Radiation Modification Research and Deployment*. <https://wedocs.unep.org/20.500.11822/41903>
- Van den Bremer, T. S., & Van der Ploeg, F. (2021). The risk-adjusted carbon price. *American Economic Review*, 111(9), 2782-2810.
- van der Ploeg, F. (2014). Abrupt positive feedback and the social cost of carbon. *European Economic Review*, 67, 28-41.

- Victor, N., & Nichols, C. (2024). Impact of carbon dioxide removal technologies on deep decarbonization: EMF37 MARKAL–NETL modeling results. *Energy and Climate Change*, 5, 100143.
- Wheatley, S., Sovacool, B., & Sornette, D. (2017). Of Disasters and Dragon Kings: A Statistical Analysis of Nuclear Power Incidents and Accidents. *Risk analysis*, 37(1), 99-115. <https://doi.org/10.1111/risa.12587>
- Wrzaczek, S., Kuhn, M., & Frankovic, I. (2020). Using age structure for a multi-stage optimal control model with random switching time. *Journal of Optimization Theory and Applications*, 184(3), 1065-1082.
- Zarnetske, P. L., Gurevitch, J., Franklin, J., Groffman, P. M., Harrison, C. S., Hellmann, J. J., Hoffman, F. M., Kothari, S., Robock, A., & Tilmes, S. (2021). Potential ecological impacts of climate intervention by reflecting sunlight to cool Earth. *Proceedings of the National Academy of Sciences*, 118(15), e1921854118.
- Zhao, Y., Basu, A., Lontzek, T. S., & Schmedders, K. (2023). The social cost of carbon when we wish for full-path robustness. *Management science*, 69(12), 7585-7606.
- Zheng, G.-L., Liao, H., & Cai, Y. (2026, 29 June-03 July 2026). *Optimal Solar Geoengineering for Managing the Tipping Risk of Coral Reef Ecosystems* 7th World Congress of Environmental and Resource Economists (Conference Paper), Carcavelos, Portugal, 29 June-03 July 2026.
- Zheng, G.-L., Liao, H., Cai, Y., Xu, J.-J., Yang, Z., Wu, J.-Y., & Wei, Y.-M. (2025). *Integrated Climate-Economic Assessment of Solar Radiation Modification and the Breakup Risk of Marine Stratocumulus Clouds* The 14th China Academic Annual Conference on Energy Economics and Management, Xi'an, China, 21-22 November, 2025.
- Zheng, G. L., Wu, J. Y., Zhang, T., & Liao, H. (2026). Robust Solar Radiation Modification Strategy for Achieving Temperature Targets. *Risk analysis*, 46(8), e70312.

Appendix

A: Details of the stochastic model

Objective function

Recalling the objective function (detailed in Section 2.2), the DM adjusts economic policy (the savings rate s), abatement policy (the abatement rate μ) and SRM policy (the radiative forcing control rate g) to maximize expected total welfare over time, which is constituted by the flow of household consumption affected by the termination shock risk.

$$\max_{s_t, \mu_t, g_t \in [0,1]} \mathbb{E}_{\omega \in \Omega} \left\{ \int_0^T e^{-\int_0^t \rho(t') dt'} L(t) \cdot U\left(c(t; I(t, \omega))\right) dt \right\} \quad [\text{A.1}]$$

For notational simplicity, we denote the stochastic process as $I(t)$, omitting explicit dependence on the random realization $\omega \in \Omega$. The process $I(t)$ follows a Poisson jump process with a constant hazard rate η , implying that the probability of a shock occurring in a small interval $[t, t + \Delta t]$ is $\eta \cdot \Delta t + o(\Delta t)$. The waiting time until the shock, denoted τ , follows an exponential distribution with parameter η , which means that the shock event occurs at a random time τ with probability distribution $P(\tau > t) = e^{-\eta t}$. Once the shock occurs, $I(t)$ switches from 1 to 0 permanently and modifies the system dynamics to feature a loss of the SRM cooling effect.

The Economic Dynamics

The total global economic output $Y(K)$ depends on capital K , population (labor) L , and total factor productivity A in the form of a Cobb-Douglas production function.

$$Y(t, K) = A(t) \cdot K^\gamma \cdot L(t)^{1-\gamma} \quad [\text{A.2}]$$

where γ represents the capital share (Karabarbounis & Neiman, 2014).

The net output $Q(\cdot)$ is the total output net of climate damage D , the abatement cost Λ , SRM loss (deployment cost and potential damages) Ω .

Climate change reduces the total economic output. The climate damage D consists of long-term D^L and short-term damage D^S (see the previous section). The long-term damage D^L depends on the increase in atmospheric temperature T^{at} relative to pre-industrial levels, i.e., the sum of the temperature boxes $\sum_i Tat_i$, while the short-term damage D^S depends on the short-term temperature changes $\sum_i Tat_i - \overline{Tat}$. Altogether we have

$$D(\cdot) = D^L(Tat_1, Tat_2) + D^S(Tat_1, Tat_2, \overline{Tat}) = \beta_1 (\sum_i Tat_i)^2 + \beta_2 \left(\frac{\sum_i Tat_i - \overline{Tat}}{\omega * s} \right)^2 \quad [\text{A.3}]$$

Abatement cost of carbon emissions is also deducted from total output. When the reduction ratio of industrial carbon emissions in time t is μ , the corresponding abatement cost Λ is,

$$\Lambda(\mu) = \theta_1 * \mu^{\theta_2} \quad [\text{A.4}]$$

where θ_1 is time-dependent coefficient, which represents the price of the backstop technology (a zero-carbon energy technology that can replace all fossil fuels). It is assumed to decline over time with carbon abatement technological change. And $\theta_2 > 1$ is a parameter.

Deploying SRM is also costly in terms of total output (Sovacool et al., 2023). When the reduction ratio of the radiative forcing caused by GHG in year t is g , the corresponding loss Ω is,

$$\Omega(g) = (\omega_c + \omega_d) \cdot g^2 \quad [\text{A.5}]$$

where $\omega_c = 0.0027$ and $\omega_d = 0.03$, representing the deployment costs and potential damage (such as ozone depletion) of SRMs, respectively. Parameter values are adopted from Heutel et al. (2018).

Therefore, net output Q is.

$$Q(\mu, g, K, Tat_i, \overline{Tat}) = (1 - D(Tat_1, Tat_2, \overline{Tat}) - \Lambda(\mu) - \Omega(g)) * Y(K) \quad [\text{A.6}]$$

A share s of total output is saved and invested in capital accumulation. As capital depreciates at rate δ_k , total capital K follows the dynamics presented below.

$$\dot{K} = s \cdot Q(\mu, g, K, Tat_i, \overline{Tat}) - \delta_k K \quad [\text{A.7}]$$

The rest of the net output is used for consumption:

$$C = (1 - s) \cdot Q(\mu, g, K, Tat_i, \overline{Tat}) \quad [A.8]$$

Production activities generate carbon emissions, $E_{ind}(\mu, K)$, but carbon intensity σ declines over time, which is an exogenous function of time that is linked to technological advancement.

$$E_{ind}(\mu, K) = \sigma Y(K) \quad [A.9]$$

The CO2 emissions also arise from land-use changes E_{land} , which are assumed to be exogenous. The actual industrial carbon emissions (in gigaton of carbon dioxide) depend on the abatement efforts, i.e., the emission control rate μ . Emissions reduction efforts μ can be seen as the share of zero-carbon energy technology usage. That is,

$$E_{co2}(\mu, K) = (1 - \mu)(E_{ind}(\mu, K) + E_{land}) \quad [A.10]$$

CO2-equivalent emissions from other gases are not included in the emissions term. But the backstop technologies for reducing CO2 emissions have synergistic effects in reducing other GHG emissions. This synergistic effect is taken into account in the radiative forcing dynamics, as discussed below (Eq. 24 and 25).

The Carbon Cycle

The geophysical model describes how CO2 emissions E_{co2} affect atmospheric carbon concentration M^{at} , radiative forcing F , and atmospheric temperature T^{at} .

Carbon emissions increase atmospheric carbon concentration M^{at} (in gigaton of carbon) through reservoir j (R_j , $j=1, 2, 3, 4$) in period t .

Cumulative carbon emissions amount to,

$$\dot{C}_{acc} = \frac{1}{c_{o2c}} E_{co2}(\mu, K) \quad [A.11]$$

The dynamics of the carbon reservoir is based on a linear four-reservoir impulse-response model of the response of CO2 concentrations to emissions. Here the “reservoirs” have no physical or structural interpretation but are variables in reduced-form dynamic equations.

$$\dot{R}_j = \xi_j \frac{1}{co2c} E(\mu, K) - \frac{1}{\tau_j \psi(C_{acc}, \sum_j R_j, \sum_i T at_i)} R_j, \quad j = 1, 2, 3, 4 \quad [A.12]$$

where ξ_j is the emissions share into reservoir j , and $\sum_j \xi_j = 1$. Coefficient $co2c = 11/3$ converts the mass of CO₂ to the mass of carbon. τ_j is the carbon decay time constant for R_j .

A key innovation in the DICE-2023's climate dynamics is the structural parameter, ψ , which increases the fraction of total CO₂ emissions that resides in the atmosphere as cumulative CO₂ emissions increase.

$$\psi(C_{acc}, \sum_j R_j, \sum_i T at_i) = \psi^0 \exp\left(\frac{\psi_*^{100} + \Delta^M [C_{acc} - \sum_j R_j] + \Delta^T \sum_i T at_i}{\psi^1}\right) \quad [A.13]$$

where ψ_*^{100} is the 100-year integrated impulse response value in the pre-industrial period, Δ^M represents the increase in ψ with cumulative carbon uptake (years per gigaton of carbon), and Δ^T represents the increase in ψ with warming (years per degree K).

The atmospheric carbon concentration M^{at} (in gigaton of carbon) is the sum of four Reservoirs;

$$M^{at}(\sum_j R_j) = M_*^{at} + \sum_j R_j \quad [A.14]$$

The Climate Dynamics

The atmospheric carbon M^{at} and other GHGs increase radiative forcing, i.e., the greenhouse effect. Although the main driver of this phenomenon is the atmospheric CO₂, forcing from other GHGs (e.g., aerosols and ozone) is also present.

$$TF(\sum_j R_j, F_{GHG}) = \varpi \cdot \log_2\left(\frac{M^{at}(\sum_j R_j)}{M_*^{at}}\right) + F_{GHG} + F_{exo} \quad [A.15]$$

where ϖ is increasing the level of forcings of equilibrium CO₂ doubling; M_*^{at} is the equilibrium concentration of CO₂ in the atmosphere. $\eta \cdot \log_2\left(\frac{M^{at}(\sum_j R_j)}{M_*^{at}}\right)$ is the change in radiative forcing due to carbon. F_{exo} is the exogenous radiative forcing from non-abatable GHGs or other sources. Finally, F_{GHG} is the radiative forcing from abatable

non-CO2 greenhouse gases (GHG), where emission reduction efforts μ exhibit synergistic effects.

$$\dot{F}_{GHG} = (1 - \mu)\psi_1 E_{GHG} - \psi_2 F_{GHG} \quad [A.16]$$

where ψ_2 and ψ_1 are geophysical parameters, $E_{GHG}(t)$ is exogenous baseline GHG emission from abatable non-CO2.

Deploying SRM offsets radiative forcing caused by GHGs. The actual radiative forcing F depend on the rate of radiation control g . However, this cooling effect may dissipate as the termination shock I occurs.

$$F(g, \sum_j R_j, F_{GHG}, I) = TF(\sum_j R_j, F_{GHG}) \cdot (1 - I \cdot g) \quad [A.17]$$

Here, I follows a Poisson jump process with a constant hazard rate η . Let us denote by τ the instant at which the shock event occurs. Once the shock occurs, $I(t)$ switches from 1 to 0 permanently.

$$\lim_{dt \rightarrow 0^+} \frac{\mathbb{P}(\tau \leq t + dt | \tau > t)}{dt} = \eta \quad [A.18]$$

The radiative forcing increases atmospheric temperature T^{at} through temperature box Tat_i (for $i=1, 2$);

$$\dot{Tat}_i = \frac{1}{\kappa_i} [F(g, \sum_j R_j, F_{GHG}, I) \cdot \chi_i - Tat_i], i = 1, 2 \quad [A.19]$$

$$T^{at}(\sum_i Tat_i) = \sum_i Tat_i \quad [A.20]$$

for $i = 1, 2$, where χ_i represents the thermal equilibration parameter (diffusion rates) for box i , and κ_i represents the thermal response timescale (time lags) for box i .

The temperature moving average can be expressed in the following continuous-time form:

$$\dot{\overline{Tat}} = \bar{\chi} \cdot (\sum_i Tat_i - \overline{Tat}) \quad [A.21]$$

The parameter $\bar{\chi}$ is calibrated by simulation data from the myopic solution over the period 2020–2100, with a termination shock in 2060 (the median year). We then choose the parameter value $\bar{\chi} = 0.49$ that minimizes the squared difference between the model-implied moving average and the simulated 5-year moving average.

When the shock occurs, the indicator I transitions from 1 to 0, indicating SRM failure and thus the inability to counteract radiative forcing, whose level is $TF(\sum_j R_j, F_{GHG}) \cdot g$. A sudden increase in radiative forcing would accelerate atmospheric temperatures dynamics, with the magnitude of the increase primarily determined by variables such as the extent of SRM deployment (g), GHGs concentrations (R_j, F_{GHG}) and climate dynamic parameters. Its economic impact is captured by the climate damage functions D^L and D^S .

B: Adjoint Equations

Recall the 2SOCS formulation of stochastic problems, and a Hamiltonian (with simplified notation) is defined as,

$$\begin{aligned}
\mathcal{H} = & z_1 \left[L \cdot U((1 - s_2)Q(\cdot)) + \int_0^t z_2 \cdot L \cdot U((1 - s_2)Q(\cdot)) d\tau \right] \\
& - \lambda_z \cdot \eta z_1 + \lambda_K \cdot (s_1 Q(\cdot) - \delta_k K_1) + \sum_j \left[\lambda_{R,j} \cdot \left(\xi_j \frac{E_{CO2}(t, \mu_1, K_1)}{CO2C} - \frac{R_{j,1}}{\tau_j \psi(\cdot)} \right) \right] + \\
& \lambda_{CACC} \frac{E_{CO2}(t, \mu_1, K_1)}{CO2C} + \lambda_{F_{GHG}} (\psi_1(1 - \mu_1)E_{GHG} - \psi_2 F_{GHG_1}) + \lambda_{\overline{Tat}} \cdot \\
& [\bar{\chi} \cdot (\sum_i Tat_{i,1} - \overline{Tat}_1)] + \sum_i \left[\lambda_{T,i} \cdot \frac{1}{\kappa_i} \cdot (\chi_i \cdot TF(t, R_{j,1}, F_{GHG_1})(1 - g_1) - \right. \\
& \left. Tat_{i,1}) \right] \\
& - \eta_z * 0 + \eta_K * [s_2 Q(\cdot) - \delta_k K_2] + \sum_j \left[\eta_{R,j} \cdot \left(\xi_j \frac{E_{CO2}(t, \mu_2, K_2)}{CO2C} - \frac{R_{j,2}}{\tau_j \psi(\cdot)} \right) \right] + \eta_{CACC} \cdot \\
& \frac{E_{CO2}(t, \mu_2, K_2)}{CO2C} + \eta_{F_{GHG}} \cdot (\psi_1(1 - \mu_2)E_{GHG} - \psi_2 F_{GHG_2}) + \eta_{\overline{Tat}} \cdot [\bar{\chi} \cdot \\
& (\sum_i Tat_{i,2} - \overline{Tat}_2)] + \sum_i \left[\eta_{T,i} \cdot \frac{1}{\kappa_i} \cdot (\chi_i \cdot TF(t, R_{j,2}, F_{GHG_2})(1 - 0 * g_2) - Tat_{i,2}) \right] \\
& + \zeta * [z_2 \cdot L \cdot U((1 - s_2)Q(\cdot))]
\end{aligned}$$

[B.1]

And the adjoint equations are:

$$\dot{\lambda}_z = \rho\lambda_z - U(c_1)L + \lambda_z\eta + \eta_z \cdot (-\eta) \quad [\text{B.2}]$$

$$\begin{aligned} \dot{\lambda}_K &= \rho\lambda_K + \delta_K\lambda_K - \lambda_K \left(s_1(1 - \Lambda - D - \Omega)Y_{1K_1} \right) - z_1U_{c_1} \left((1 - s_1)(1 - \Lambda - D - \right. \\ &\quad \left. \Omega)Y_{1K_1} \right) - \frac{1}{\text{co2c}} \sum_j (\lambda_{R,j}\xi_j) \cdot [(1 - \mu_1)\sigma Y_{K_1}] - \frac{1}{\text{co2c}} \lambda_{Cacc} [(1 - \mu_1)\sigma Y_{K_1}] - \eta_K \\ &= (\rho + \delta_K)\lambda_K - \left[(1 - \Lambda - D - \Omega)\lambda_K + \frac{\sigma}{\text{co2c}} (1 - \mu_1^*) (\sum_j (\lambda_{R,j}\xi_j) + \lambda_{Cacc}) \right] Y_{1K_1} - \eta_K \end{aligned} \quad [\text{B.3}]$$

$$\begin{aligned} \dot{\lambda}_{R,j} &= \rho\lambda_{R,j} + \frac{\partial \left(\sum_j \left[\lambda_{R,j} \left(\xi_j \frac{E_{\text{co2}}(t, \mu_1, K_1)}{\text{co2c}} - \frac{R_{j,1}}{\tau_j \psi(\cdot)} \right) \right] \right)}{\partial R_{j,1}} - \\ &\quad \frac{\partial \left(\sum_i \left[\lambda_{T,i} \frac{1}{\kappa_i} (\chi_i \cdot TF(t, R_{j,1}, F_{GHG_1})(1 - g_1) - T_{at,i,1}) \right] \right)}{\partial R_{j,1}} - \eta_{R,j} \\ &= \rho\lambda_{R,j} + \frac{\sum_j \left(\lambda_{R,j} \left(1 + \frac{\Delta^M}{\psi^1} R_{j,1} \right) \right)}{\tau_j \psi(\cdot)} + \frac{1}{\tau_j \psi(\cdot)} \frac{1}{\ln 2 * Mat(\sum_j R_{j,1})} \sum_i \left(\lambda_{Tat,i} \frac{\chi_i}{\kappa_i} \right) (1 - g_1^*) - \eta_{R,j} \end{aligned} \quad [\text{B.4}]$$

$$\dot{\lambda}_{Cacc} = \rho\lambda_{Cacc} - \frac{\Delta^M}{\psi^1} \sum_j \lambda_{R,j} R_{j,1} \frac{1}{\tau_j \psi(\cdot)} - \eta_{Cacc} \quad [\text{B.5}]$$

$$\begin{aligned} \dot{\lambda}_{F_{GHG}} &= \rho\lambda_{F_{GHG}} + \psi_2\lambda_{F_{GHG}} - \sum_i \lambda_{T,i} \frac{\chi_i}{\kappa_i} \frac{\partial TF(R_{j,1}, F_{GHG_1})}{\partial F_{GHG_1}} (1 - g_1) - \eta_{F_{GHG}} \\ &= (\rho + \psi_2)\lambda_{F_{GHG}} - (1 - g_1) \sum_i \lambda_{T,i} \frac{\chi_i}{\kappa_i} - \eta_{F_{GHG}} \end{aligned} \quad [\text{B.6}]$$

$$\begin{aligned} \dot{\lambda}_{Tat,i} &= \rho\lambda_{Tat,i} - \lambda_K (-s_1 D_{Tat,i,1} Y_1) + \sum_j \lambda_{R,j} R_{j,1} \left(-\frac{\Delta^T}{\psi^1} \right) \frac{1}{\tau_j \psi(\cdot)} + \frac{1}{\kappa_i} \lambda_{Tat,i} - \bar{\chi} \lambda_{\overline{Tat}} - \\ &\quad z_1 U_{c_1} \left((1 - s_1) D_{Tat,i,1} Y_1 \right) - \eta_{Tat,i} \\ &= \left(\rho + \frac{1}{\kappa_i} \right) \lambda_{Tat,i} + \lambda_K (D_{Tat,i,1} Y_1) - \left(\frac{\Delta^T}{\psi^1} \right) \frac{1}{\tau_j \psi(\cdot)} \sum_j (\lambda_{R,j} R_{j,1}) - \bar{\chi} \lambda_{\overline{Tat}} - \eta_{Tat,i} \end{aligned} \quad [\text{B.7}]$$

$$\begin{aligned} \dot{\lambda}_{\overline{Tat}} &= \rho\lambda_{\overline{Tat}} + \bar{\chi} \lambda_{\overline{Tat}} - z_1 U_{c_1} \left((1 - s_1) D_{\overline{Tat},1} Y_1 \right) - \lambda_K (s_1 D_{\overline{Tat},1} Y_1) - \eta_{\overline{Tat}} \\ &= (\rho + \bar{\chi}) \lambda_{\overline{Tat}} + \lambda_K (D_{\overline{Tat},1} Y_1) - \eta_{\overline{Tat}} \end{aligned} \quad [\text{B.8}]$$

$$\dot{\eta}_z = \rho\eta_z - U(c_2)L \quad [\text{B.9}]$$

$$\begin{aligned} \dot{\eta}_K &= \rho\eta_K + \delta_K\eta_K - \eta_K \left(s_2(1 - \Lambda - D - D_g)Y_{2K_2} \right) - z_2U_{c_2} \left((1 - s_2)(1 - \Lambda - D - \right. \\ &\quad \left. D_g)Y_{2K_2} \right) - \sum_j \eta_{R,j} \xi_j \frac{1}{\text{co2c}} * \left[(1 - \mu_2)\sigma Y_{2K_2} \right] - \eta_{Cacc} \frac{1}{\text{co2c}} \left[(1 - \mu_2)\sigma Y_{2K_2} \right] \end{aligned}$$

$$= (\rho + \delta_k)\eta_K - \left[(1 - \Lambda - D - \Omega)\eta_K + \frac{\sigma}{c o 2 c} (1 - \mu_2^*) (\sum_j (\eta_{R,j} \xi_j) + \eta_{Cacc}) \right] Y_{2K_2} \quad [\text{B.10}]$$

$$\begin{aligned} \dot{\eta}_{R,j} &= \rho \eta_{R,j} + \frac{\eta_{R,j}}{\tau_j \psi^0} \frac{\partial}{\partial R_{j,2_1}} \left[R_{j,2} e^{-\frac{(iURF_*^{100} + \Delta^M [Cacc_2 - \sum_j R_{j,2}] + \Delta^T \sum_i T_{at,i,2})}{\psi^1}} \right] - \\ &\quad \sum_i \eta_{T,i} \frac{\chi_i}{\kappa_i} \frac{\partial [F(R_{j,2}, F_{GHG_2})]}{\partial R_{j,2_1}} \\ &= \rho \eta_{R,j} + \frac{1}{\tau_j \psi(\cdot)} \cdot \sum_j \left(\eta_{R,j} \left(1 + \frac{\Delta^M}{\psi^1} R_{j,2} \right) \right) + \frac{1}{\tau_j \psi(\cdot)} \frac{1}{\ln 2 * Mat(\sum_j R_{j,2})} \sum_i \left(\eta_{Tat,i} \frac{\chi_i}{\kappa_i} \right) \end{aligned} \quad [\text{B.11}]$$

$$\dot{\eta}_{Cacc} = \rho \eta_{Cacc} - \frac{\Delta^M}{\psi^1} \sum_j \eta_{R,j} R_{j,2} \frac{1}{\tau_j \psi(\cdot)} \quad [\text{B.12}]$$

$$\begin{aligned} \dot{\eta}_{F_{GHG}} &= \rho \eta_{F_{GHG}} + \psi_2 \eta_{F_{GHG}} - \sum_i \eta_{T,i} \frac{\chi_i}{\kappa_i} \frac{\partial}{\partial F_{GHG_2}} [TF(R_{j,2}, F_{GHG_2})] \\ &= (\rho + \psi_2) \eta_{F_{GHG}} - \sum_i \eta_{T,i} \frac{\chi_i}{\kappa_i} \end{aligned} \quad [\text{B.13}]$$

$$\begin{aligned} \dot{\eta}_{Tat,i} &= \rho \eta_{Tat,i} - \eta_K (-s_2 D_{Tat,i,2} Y_2) + \left(\sum_j \eta_{R,j} R_{j,2} \frac{1}{\tau_j \psi(\cdot)} \right) \left(-\frac{\Delta^T}{\psi^1} \right) + \frac{1}{\kappa_i} \eta_{Tat,i} - \bar{\chi} \eta_{\bar{T}at} - \\ &\quad z_2 U_{c_2} \left((1 - s_2) D_{Tat,i,2} Y_2 \right) \\ &= \left(\rho + \frac{1}{\kappa_i} \right) \eta_{Tat,i} + \eta_K (D_{Tat,i,2} Y_2) - \left(\frac{\Delta^T}{\psi^1} \right) \sum_j \left(\eta_{R,j} R_{j,2} \frac{1}{\tau_j \psi(\cdot)} \right) - \bar{\chi} \eta_{\bar{T}at} \end{aligned} \quad [\text{B.14}]$$

$$\begin{aligned} \dot{\eta}_{\bar{T}at} &= \rho \eta_{\bar{T}at} + \bar{\chi} \eta_{\bar{T}at} - z_2 U_{c_2} \left((1 - s_2) D_{\bar{T}at,2} Y_2 \right) - \eta_K (s_2 D_{\bar{T}at,2} Y_2) \\ &= (\rho + \bar{\chi}) \eta_{\bar{T}at} + \lambda_K (D_{\bar{T}at,2} Y_1) \end{aligned} \quad [\text{B.15}]$$

$$\zeta = 1 \quad [\text{B.16}]$$

C: Proof for costate of shock probability

Define the value function of stage 2 (post-shock),

$$V_2(t, \mathbf{X}_1(t)) = \max_{c, \mu, g} \int_t^T e^{-\rho(\theta-t)} L \cdot U(c_2(\theta)) d\theta \quad [\text{C.1}]$$

subject to:

$$\dot{\mathbf{X}}_2(s) = f_2(s, \mathbf{X}_2(s), c_2(t), \mu_2(t), g_2(t)) \quad s > t \quad [\text{C.2}]$$

$$\mathbf{X}_2(t) = \varphi(\mathbf{X}_1(t)) \quad [\text{C.3}]$$

We got:

$$\begin{aligned} \eta_z(t, s) &= \int_t^T e^{-\int_t^{\tau} \rho ds'} [U(c_2(\tau, s)) L(\tau, s)] d\tau \\ &= \int_t^T e^{-\rho(\tau-t)} [U(c_2(\tau, s)) L(\tau, s)] d\tau = V_2(t, \mathbf{X}_2^*(t, s)) \end{aligned} \quad [\text{C.4}]$$

and, in particular,

$$\eta_z(t, t) = V_2(t, \mathbf{X}_2^*(t, t)) = V_2(t, \varphi(\mathbf{X}_1^*(t))) \quad [\text{C.5}]$$

So that

$$\begin{aligned} \lambda_z(t) &= \int_t^T e^{-(\rho+\eta)(\tau-t)} [L \cdot U(c_1(\tau)) + \eta V_2(t, \varphi(\mathbf{X}_1^*(t)))] d\tau \\ &= \int_t^T e^{-\eta(\tau-t)} e^{-\rho(\tau-t)} [L \cdot U(c_1(\tau)) + \eta V_2(t, \varphi(\mathbf{X}_1^*(t)))] d\tau \\ &= \int_t^T \frac{z_1(\tau)}{z_1(t)} e^{-\rho(\tau-t)} [L \cdot U(c_1(\tau)) + \eta V_2(t, \varphi(\mathbf{X}_1^*(t)))] d\tau \end{aligned} \quad [\text{C.6}]$$

Consider two cases,

Case1: tipped just happened in time s , the welfare is:

$$\begin{aligned} W_1 &= \int_0^s e^{-\rho t} L \cdot U(c_1(t)) dt + \int_s^T e^{-\rho t} L \cdot U(c_2(t)) dt \\ &= \int_0^s e^{-\rho t} L \cdot U(c_1(t)) dt + e^{-\rho s} V_2(t, \mathbf{X}_1^*(t)) \\ &= \int_0^s e^{-\rho t} L \cdot U(c_1(t)) dt + e^{-\rho s} \eta_z(s, s) \end{aligned} \quad [\text{C.7}]$$

the probability of Case1 is $z_1(t) * \eta$.

Case 2: no tipped before time s , and the expected value for the rest of periods (s, T) is:

$$\begin{aligned}
 W_2 &= \int_0^s e^{-\rho t} L \cdot U(c_1(t)) dt \\
 &\quad + \mathbb{E}_\tau \left[\int_s^\tau e^{-\rho t} L \cdot U(c_1(t)) dt + \int_\tau^T e^{-\rho t} L \cdot U(c_2(t, \tau)) dt \right] \\
 &= \int_0^s e^{-\rho t} L \cdot U(c_1(t)) dt + \int_s^T \mathbf{p}_1(t) e^{-\rho t} \left[L \cdot U(c_1(t)) + \eta V_2(t, \varphi(\mathbf{X}_1^*(t))) \right] dt \quad [\text{C.8}]
 \end{aligned}$$

Subject to

$$\dot{\mathbf{X}}_1(t) = f_2(t, \mathbf{X}_1(t), c_1(t), \mu_1(t), g_1(t)) \quad [\text{C.9}]$$

$$\dot{p}_1(t) = -\eta p_1(t) \quad p_1(s) = 1 \quad [\text{C.10}]$$

Here, $p_1(s) = 1$, since the decision maker learn the information that no shock occurs until before s , therefore,

$$\begin{aligned}
 W_2 &= \int_0^s e^{-\rho t} L \cdot U(c_1(t)) dt + \int_s^T \frac{z_1(t)}{z_1(s)} e^{-\rho t} \left[L \cdot U(c_1(t)) + \eta V_2(t, \varphi(\mathbf{X}_1^*(t))) \right] dt \\
 &= \int_0^s e^{-\rho t} L \cdot U(c_1(t)) dt + e^{-\rho s} \lambda_z(s) \quad [\text{C.11}]
 \end{aligned}$$

Therefore, compared to the case of termination shock happened in period s , the additional welfare of no termination shock happened before period s , is:

$$W_2 - W_1 = e^{-\rho s} [\lambda_z(s) - \eta_z(t, t)] \quad [\text{C.12}]$$

D: Derivation of SCC and SCF, etc.

Derivation of SCC

The Social Cost of Carbon (SCC) is defined as the Marginal Rate of Substitution (MRS) of consumption for emissions:

$$SCC = -\frac{dC}{dE_{CO2}} \quad [D.1]$$

Let V be the Value function depending on total consumption and emissions. Since V is constant on the level set, which yields

$$SCC = -\frac{\partial V}{\partial E_{CO2}} * \frac{\partial C}{\partial V} \quad [D.2]$$

Since E_{CO2} and C are not state variables and V does not depend on them explicitly, we identify which quantities that are directly impacting V they influence, compute the derivatives of such quantities with respect to E_{CO2} and C , and then apply the chain rule to $\frac{\partial V}{\partial E_{CO2}}$ and $\frac{\partial V}{\partial C}$.

C enters the utility function $U(c) = U(C/L)$

$$\frac{\partial V}{\partial C} = \frac{\partial V}{\partial U} \frac{dU}{dC} \frac{1}{L} \quad [D.3]$$

Here, the derivative $\frac{dV}{dU}(t)$ is to be intended in the sense of Fréchet, i.e. the variation of $V(t)$ that is produced by $\Delta U = 1$ in the small interval preceding t and 0 otherwise ("blip"),

$$\Delta V(t) = \int_t^{t+\varepsilon} e^{-\rho(s-t)} [z_1(s)L(s)] ds \quad [D.4]$$

$$\frac{\partial V}{\partial U}(t) = \lim_{\varepsilon \rightarrow 0^+} \frac{\Delta V(t)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{\int_t^{t+\varepsilon} e^{-\rho(s-t)} [z_1(s)L(s)] ds}{\varepsilon} = z_1(t)L(t) \quad [D.5]$$

Then, if saving rate is an inner control,

$$\frac{\partial V}{\partial C} = z_1 \frac{dU(c)}{dc} = \lambda_K \quad [D.6]$$

E_{CO2} enters the dynamics of both R_j and C_{acc} , therefore,

$$\frac{\partial V}{\partial E_{CO2}} = \sum_j \frac{\partial V}{\partial R_j} \frac{\partial R_j}{\partial E_{CO2}} + \frac{\partial V}{\partial C_{acc}} \frac{\partial C_{acc}}{\partial E_{CO2}} \quad [D.7]$$

Recall the dynamics of R_j .

$$\dot{R}_j(t) = \xi_j \frac{1}{c_{02c}} E_{CO2}(t, \mu, K) - \frac{1}{\tau_j \psi(C_{acc}, \sum_j R_j, \sum_i T a t_i)} R_j, \quad j = 1, 2, 3, 4 \quad [D.8]$$

Specifically, we assume an exponential form:

$$\varpi(t, C_{acc}, \sum_j R_j) = \frac{1}{\tau_j \psi^0} e^{-\frac{iURF_*^{100} + \Delta T_* \sum_i T a t_i}{\psi^1}} e^{-\frac{\Delta M}{\psi^1} (C_{acc1} - \sum_j R_j)} = A \cdot e^{-B \cdot (C_{acc1} - \sum_j R_j)} \quad [D.9]$$

Then we can rewrite:

$$\dot{R}_j(t) = \xi_j \frac{1}{co2c} E_{co2}(t, \mu, K) - \varpi(t, C_{acc}, \sum_j R_j) R_j, \quad j = 1, 2, 3, 4 \quad [D.10]$$

To investigate the marginal effect of carbon emissions on the system, we introduce a small perturbation δE in emissions at time t . This perturbation induces corresponding changes ΔR_j in each reservoir and ΔC_{acc} in the accumulated carbon. define ϖ^ϵ as the value of ϖ at the perturbed state:

$$\varpi^\epsilon = \varpi(t, C_{acc}, \sum_j R_j) \quad [D.11]$$

After applying the emission perturbation, the dynamics becomes:

$$\frac{d}{dt} (R_j + \Delta R_j) = \xi_j \frac{1}{co2c} (E_{co2} + \delta E) - \varpi^\epsilon \cdot (R_{j,1} + \Delta R_{j,1}) \quad [D.12]$$

Subtracting the original unperturbed system, we isolate the dynamics of the perturbation:

$$\frac{d}{dt} \Delta R_j = \xi_j \cdot \frac{1}{co2c} \cdot \delta E - \varpi \Delta R_j - (\varpi^\epsilon - \varpi) R_j + o(\delta E) \quad [D.13]$$

To proceed, we expand ϖ^ϵ using a first-order Taylor approximation around $(C_{acc}, \sum_j R_{j,1})$

$$\varpi^\epsilon = \varpi \cdot e^{-B \cdot (\Delta C_{acc} - \sum_j \Delta R_{j,1})} \approx \varpi \cdot \left(1 - B \cdot (\Delta C_{acc} - \sum_j \Delta R_{j,1}) \right) + o(\delta E) \quad [D.14]$$

Substituting this back, we obtain

$$\frac{d}{dt} \Delta R_j = \xi_j \cdot \frac{1}{co2c} \cdot \delta E - \varpi \Delta R_j + \varpi \cdot B \cdot R_j \cdot (\Delta C_{acc} - \sum_j \Delta R_{j,1}) + o(\delta E) \quad [D.15]$$

To solve for ΔR_j we use integrating factors and assume ϖ remains approximately constant over a small interval $[t - \epsilon, t]$. Solving the linear differential equation, then,

$$\Delta R_j(\epsilon) = \frac{\xi_j}{co2c} \cdot \delta E \cdot \left[\frac{1 - e^{-\varpi \epsilon}}{\varpi} \right] + o(\delta E) \quad [D.16]$$

Dividing both sides by δE and taking the limit as $\delta E \rightarrow 0$ and $\epsilon \rightarrow 0$, we get the sensitivity of each reservoir to emissions:

$$\frac{dR_j}{dE_{co2}} = \frac{\xi_j}{co2c} \quad [D.17]$$

E_{co2} also enters the dynamics of C_{acc}

$$\dot{C}_{acc}(t) = \frac{1}{co2c} E_{co2}(t, \mu, K) \quad [D.18]$$

Then, the derivative $\frac{dC_{acc}}{dE_{CO2}}$ is again a Fréchet derivatives, and

$$\begin{aligned}
\frac{d}{dt}(C_{acc} + \Delta C_{acc}) &= \frac{1}{CO2c}(E_{CO2} + 1) \\
\Rightarrow \frac{d(\Delta C_{acc})}{dt} &= \frac{1}{CO2c} \text{ with } \Delta C_{acc}(t - \varepsilon) = 0 \\
\Rightarrow \Delta C_{acc}(t) &= \int_{t-\varepsilon}^t \frac{d(\Delta C_{acc})}{dt} ds = \frac{1}{CO2c} \varepsilon \\
\Rightarrow \frac{dC_{acc}}{dE_{CO2}}(t) &= \lim_{\varepsilon \rightarrow 0^+} \frac{\Delta C_{acc}}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{1}{CO2c} = \frac{1}{CO2c}
\end{aligned} \tag{D.19}$$

So that,

$$SCC = -\frac{1}{CO2c} \left(\sum_j (\lambda_{R_j} \xi_j) + \lambda_{C_{acc}} \right) * \frac{1}{\lambda_K} \tag{D.20}$$

Derivation of $SCC_{E_{GHG}}$

Correspondingly, the *Social Cost of abatable non-CO2 GHGs emission*, $SCC_{E_{GHG}}$, as the marginal rate of substitution between consumption and abatable non-CO2 GHGs emission (the E_{GHG} after abatement),

$$SCC_{E_{GHG}} = -\frac{dC}{d((1-\mu)E_{GHG})} = -\frac{dV}{d(E_{GHG})} \cdot \frac{1}{(1-\mu)} \cdot \frac{1}{\lambda_K} \tag{D.21}$$

E_{GHG} enters the dynamic of F_{GHG}

$$\dot{F}_{GHG} = (1 - \mu)\psi_1 E_{GHG} - \psi_2 F_{GHG} \tag{D.22}$$

The derivative $\frac{dF_{GHG}}{dE_{GHG}}$ is again a Fréchet derivatives. Refer to the previous derivation,

$$\frac{dF_{GHG}}{dE_{GHG}} = \psi_1 \tag{D.23}$$

So that,

$$SCC_{E_{GHG}} = -\psi_1 \frac{\lambda_{F_{GHG}}}{\lambda_K} \tag{D.24}$$

Derivation of SCF

We now turn to the derivation of the Social Cost of Forcing, SCF , which represents the marginal rate of substitution between consumption and radiative forcing (after SRM).

$$SCF = -\frac{dC}{d((1-g) \cdot TF(R_j, F_{GHG}))} = -\frac{dV}{d((1-g) \cdot TF(R_j, F_{GHG}))} * \frac{1}{\lambda_K} \quad [D.25]$$

The radiative forcing enters the dynamic of Tat_i ($i = 1, 2$), then

$$\frac{dV}{d((1-g) \cdot TF(R_j, F_{GHG}))} = \left(\sum_i \frac{dV}{dTat_i} \frac{dTat_i}{d((1-g) \cdot TF(R_j, F_{GHG}))} \right) \quad [D.26]$$

Recall the dynamics of Tat_i ,

$$\dot{Tat}_i = \frac{1}{\kappa_i} [\chi_i \cdot TF(t, \sum_j R_j, F_{GHG}) \cdot (1 - I \cdot g) - Tat_i] \quad [D.27]$$

Here, $TF(R_{j,1}, F_{GHG_1})$ is the radiative forcing from greenhouse gases, g denotes the fraction of forcing offset by SRM. We introduce a small perturbation $\delta F = (1 - I \cdot g) \cdot \delta TF$, corresponding to a marginal increase in effective radiative forcing.

$$\frac{d}{dt}(Tat_i + \Delta Tat_i) = \frac{1}{\kappa_i} [\chi_i \cdot ((1 - I \cdot g) \cdot TF(R_j, F_{GHG}) + \delta F) - Tat_i - \Delta Tat_i] \quad [D.28]$$

Subtracting the unperturbed system, the dynamics of the perturbation are:

$$\frac{d}{dt} \Delta Tat_i = \frac{\chi_i}{\kappa_i} \delta F - \frac{1}{\kappa_i} \Delta Tat_{i,1} \quad [D.29]$$

With the boundary condition $\Delta Tat_i(t - \varepsilon) = 0$,

$$\Delta Tat_i(t) = \int_{t-\varepsilon}^t e^{-\frac{t-s}{\kappa_i}} \cdot \left(\frac{\chi_i}{\kappa_i} \right) \delta F ds \quad [D.30]$$

For sufficiently small ε , the integral reduces to

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\Delta Tat_i}{\delta F \cdot \varepsilon} = \frac{\chi_i}{\kappa_i} \quad [D.31]$$

Therefore,

$$SCF_1 = -\frac{1}{\lambda_K} \cdot \sum_j \left(\frac{\chi_i}{\kappa_i} \cdot \lambda_{Tat_{i,1}} \right) \quad [D.32]$$

Derivation of MAC

The marginal abatement cost MAC measures the economic out cost of reduce one more unit CO2 emission,

$$MAC = -\frac{\partial Q}{\partial E_{CO2}} \quad [D.33]$$

and note the abatement efforts μ impacts Q , E_{CO2} and F_{GHG} , then for $\mu = \mu^*$,

$$\begin{aligned}
\frac{\partial V}{\partial \mu} &= \frac{\partial V}{\partial Q} \frac{\partial Q}{\partial \mu} + \frac{\partial V}{\partial E_{co2}} \frac{\partial E_{co2}}{\partial \mu} + \frac{\partial V}{\partial F_{GHG}} \frac{\partial F_{GHG}}{\partial \mu} = 0 \\
\Rightarrow -\frac{\partial V / \partial E_{co2}}{\partial V / \partial Q} &= \frac{\partial Q / \partial \mu}{\partial E_{co2} / \partial \mu} + \frac{\partial V / \partial F_{GHG}}{\partial V / \partial Q} \frac{\partial F_{GHG} / \partial \mu}{\partial E_{co2} / \partial \mu} \\
\Rightarrow -\frac{\partial Q}{\partial E_{co2}} &= \frac{Y}{\sigma Y + E_{land}} \Lambda_{\mu} + \frac{\lambda_{F_{GHG}}}{\partial V / \partial Q} \cdot \frac{\psi_1 E_{GHG}}{\sigma Y + E_{land}}
\end{aligned}
\tag{D.34}$$

Note the abatement efforts Q impacts C and K ,

$$\frac{\partial V}{\partial Q} = s \cdot \frac{\partial V}{\partial C} + (1 - s) \cdot \lambda_K \tag{D.35}$$

Then, $\frac{\partial V}{\partial C} = \lambda_K$, if saving rate is an inner control,

Therefore,

$$MAC = \frac{Y}{\sigma Y + E_{land}} \Lambda_{\mu} + \frac{\lambda_{F_{GHG}}}{\lambda_K} \cdot \frac{\psi_1 E_{GHG}}{\sigma Y + E_{land}} \tag{D.36}$$

Derivation of MFC

The marginal radiative forcing cost MFC measures the economic out cost of reduce one more unit radiative forcing (after SRM),

$$MAC = -\frac{\partial Q}{\partial F} \tag{D.37}$$

and note the abatement efforts g impacts Q and F , then for $g = g^*$,

$$\begin{aligned}
\frac{\partial V}{\partial g} &= \frac{\partial V}{\partial Q} \frac{\partial Q}{\partial g} + \frac{\partial V}{\partial F} \frac{\partial F}{\partial g} = 0 \\
\Rightarrow -\frac{\partial V / \partial F}{\partial V / \partial Q} &= \frac{\partial Q / \partial g}{\partial F / \partial g}
\end{aligned}
\tag{D.38}$$

Therefore,

$$MFC = -\frac{\partial Q}{\partial F} = \frac{\Omega_g Y}{TF(R_j, F_{GHG})} \tag{D.39}$$