The Social Cost of Carbon with Economic and Climate Risk*

Yongyang Cai, Kenneth L. Judd, and Thomas S. Lontzek

Economics Working Paper 18113

HOOVER INSTITUTION
434 GALVEZ MALL
STANFORD UNIVERSITY
STANFORD, CA 94305-6010

December 1, 2017

There is great uncertainty about both future economic and future climate conditions; uncertainty that substantially affect the choice of policies for managing interactions between the climate and the economy. A critical index of this interaction is the social cost of carbon (SCC)—that is, the present value of the marginal damage to economic output caused by carbon emissions. We show that SCC is substantially affected by both economic risks, such as shocks to

* We thank Kenneth Arrow, Buz Brock, Varadarajan V. Chari, Jesus Fernandez-Villaverde, Larry Goulder, Lars Peter Hansen, Tom Hertel, Larry Karp, Tim Lenton, Robert Litterman, Alena Miftakhova, Karl Schmedders, Christian Traeger, Rick van der Ploeg, and Ole Wilms for comments on earlier versions of this paper. We thank the anonymous referees for their helpful comments. We are grateful for comments from audiences at the 2016 North American Summer Meeting of the Econometric Society, the 2016 EAERE conference, the 2014 Minnesota Conference on Economic Models of Climate Change, the 2014 Stanford Institute for Theoretical Economics, the 2013 INFORMS conference, the 2012 IIES Conference on Climate and the Economy, the 2012 AERE summer conference, the 2011 ASSA annual meeting, and many seminars. Yongyang Cai acknowledges support from the Hoover Institution at Stanford University, and National Science Foundation grants (SES-0951576 and SES-1463644). Financial support for Thomas Lontzek was provided by the Zürcher Universitätsverein, the University of Zurich, and the Ecosciencia Foundation. This research is part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the State of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications. This research was also supported in part by NIH through resources provided by the Computation Institute and the Biological Sciences Division of the University of Chicago and Argonne National Laboratory, under grant 1S10OD018495-01. We give special thanks to the HTCondor team of the University of Wisconsin-Madison for their support. Earlier versions of this paper include "The Social Cost of Stochastic and Irreversible Climate Change" (NBER working paper 18704), "DSICE: A Dynamic Stochastic Integrated Model of Climate and Economy" (RDCEP working paper 12-02), and "Tipping Points in a Dynamic Stochastic IAM" (RDCEP working paper 12-03).

growth rates, and climate risks, such as irreversible changes in sea levels. Moreover, SCC is itself a stochastic process with significant variation, a factor of substantial importance in evaluating any decision with a real option dimension. Furthermore, we have only imprecise information about what parameter values best approximate reality. Our response is to examine a plausible but wide range of values for critical parameters and find results that are robust. More generally, this work shows that large-scale computing makes it possible to analyze policies in models substantially more complex and realistic than usually used in the integrated assessment models literature.

The Social Cost of Carbon with Economic and Climate Risk
Yongyang Cai, Kenneth L. Judd, and Thomas S. Lontzek
Economics Working Paper 18113
December 1, 2017

JEL Code: C63, Q54, H21
Keywords: Climate policy, social cost of carbon, climate tipping process, Epstein–Zin preferences, stochastic growth, long-run risk

Yongyang Cai
Department of Agricultural, Environmental and Development Economics,
The Ohio State University
Columbus, OH 43210, USA
cai.619@osu.edu

Kenneth L. Judd
Hoover Institution, Stanford University & NBER
Stanford, CA 94305, USA
kennethjudd@mac.com

Thomas S. Lontzek
RWTH Aachen University, School of Business and Economics
52056 Aachen, Germany
lontzek@econ.rwth-aachen.de

1 Introduction

Global warming has been recognized as a growing potential threat to economic well-being. Determining which policies should be implemented requires analyses that incorporate models of both the climate and the economy, and how they interact; this is the purpose of *integrated assessment models* (IAMs). This paper expands the scope of IAMs by adding uncertainties and risks to a canonical model of the economic and climate systems. and shows that such risks and uncertainties significantly affect optimal climate policy. Almost all IAMs assume that both the climate and economic systems are deterministic, and that economic agents are myopic. Studies of those IAMs focus on uncertainty about the key parameters. This study, instead, incorporates economic and climate risks into models where economic agents have rational expectations concerning the future of the economy and of the climate. This allows us to use ideas from economic growth, investment, and consumption CAPM theories to evaluate alternative climate change policies. As expected, explicit treatment of economic and climate risks significantly changes the results.

The impact of carbon emissions on society is measured by the social cost of carbon (SCC), defined as the marginal economic loss caused by an extra metric ton of atmospheric carbon.[1] The importance of SCC for policy decisions has led to a large effort to estimate its value. One prominent study was produced by the Interagency Working Group on Social Cost of Carbon (IWG)—a joint effort involving several United States federal agencies. The IWG report (IWG 2010) was based on three well-known economic integrated assessment models: FUND (Anthoff and Tol 2014), PAGE (Hope 2011), and DICE (Nordhaus 2008). The IWG report gives a wide range of plausible values for SCC with a median of \$51/tC.[2] The Stern report (Stern 2007) presents an alternative analysis, focusing on parameter uncertainty. Using the probabilistic integrated assessment model PAGE (Hope, 2011), the report addresses parametric uncertainty regarding the causes and impacts of climate change and argues for substantially higher SCC values.

The models behind the IWG and Stern reports use simplified descriptions of the climate. Many other analyses are based on far more complex climate models and much more detailed economic

[1] We express SCC in US dollars (\$) per ton of carbon and denoted \$/tC. SCC per ton of CO2 equals12/44 times the SCC per ton of carbon.

[2] The IWG report is based on solving FUND, PAGE, and DICE for thousands of values of the critical parameters, where those parameter choices are based on empirical data and expert opinion. The models express SCC in terms of US dollars per ton of CO2. After converting to \$/tC, they report that the 5th, 25th, 50th, 75th, and 95th percentile values are -\$33, \$15, \$51, \$102, and \$238, respectively.

models, and are included in the summary produced by the Intergovernmental Panel on Climate Change (IPCC). IPCC (2007) reports that estimates of SCC vary across peer-reviewed studies with an average estimate of $43 per ton of carbon but with a large range.

All of these studies are based on IAMs with deterministic models of economy and climate, with all economic agents knowing current and future output and climate conditions. Recent reviews, such as Pindyck (2013) and IPCC (2014), criticize the fact that SCC estimates are based on IAMs that ignore the considerable risk and uncertainty in both the economic and the climate system, and their interactions.

This study presents Dynamic Stochastic Integration of Climate and the Economy (DSICE), a computational, dynamic, stochastic general equilibrium framework for studying global models of both the economy and the climate. We apply it to the specific issue of how SCC depends on stochastic features of both the climate and the economy, and use specifications of tastes consistent with the observed willingness to pay to reduce economic risk. This study focuses on the SCC as an example of the flexibility of the DSICE framework and its ability to examine a wide range of issues. DSICE builds on the canonical "Dynamic Integrated Climate–Economy" (DICE) model (Nordhaus, 1992, 2008) by adding both economic and climate risks to the DICE framework. We choose to extend DICE because it is one of the few prominent IAM models that are based on dynamic models of agent decision-making.[3]

We first examine how economic risks affect SCC. Specifically, we assume that factor productivity growth implies consumption growth rates that display long-run risk as modeled by Bansal and Yaron (2004) and Beeler and Campbell (2011). We combine this with Epstein–Zin utility specifications of dynamic preferences (Epstein and Zin, 1989). Current empirical analyses do not give us precise estimates of critical parameters, particularly risk-aversion and the inter-temporal elasticity of substitution (IES). We solve DSICE for parameter values covering the range of empirical estimates in the macroeconomic literature. We find that the 2005 SCC ranges from $59 to $99 per ton of carbon

[3]Manne and Richels (2005) and Nordhaus and Yang (1996) are also based on dynamic models of agent decision-making. Many IAMs say that they have "recursive" dynamic models of the economy. This use of the term "recursive" was common in dynamic CGE modeling in the 1970s. However, those models were recursive only in the sense that economic decisions today create the state of the economy tomorrow. Today we call those models "myopic", because dynamic decisions, such as investment, were simple functions of contemporaneous prices and states. Today, the description "recursive models" refers to models where agents have rational expectations; see, for example, Stokey and Lucas (1989) and Ljungqvist and Sargent (2004). The antiquated use of the term "recursive" in the IAM literature (such as Babiker et al. 2005) is misleading to economists familiar with modern dynamic economic analysis.

over a plausible range of parameter values. These results demonstrate that we should be skeptical of studies that give one number relying on a single parameterization. In this case, sensitivity analysis does support the case for a significant carbon tax.

Of equal interest and greater novelty are our results on the dynamics of the future SCC. Factor productivity shocks create riskiness in future output and carbon emissions, which in turn makes SCC a random process. A common interpretation of results in deterministic models is that they represent the expected path in stochastic models. In many cases, we find that the mean path for the SCC is close to that implied by deterministic models. DSICE, however, can also determine the stochastic features of the SCC process. SCC is the shadow price of a state variable and, as we expect, is approximately a random walk. When we quantify the SCC process we find that it displays substantial variance. For example, in our benchmark case the median SCC is $286 in 2100 but with a 10 percent chance of exceeding $700 and a 1 percent chance of exceeding $1,200. In general, the standard deviation of SCC grows faster than its mean.

Some studies (e.g., Weitzman (2009), Pindyck (2011), and Scheffer et al. (2001)) use the possibility of very high damage caused by a very low probability catastrophic event to advocate aggressive mitigation policies. Estimating the probability of tail events is very difficult, a fact that reduces the force of these arguments. However, it is recognized that temperature increases[4] may cause substantial irreversible damage to the climate. The IAM literature has recently studied the importance of *climate tipping points*, which refer to "a critical threshold at which a tiny perturbation can qualitatively alter the state or development of the climate system", and *tipping elements,* which are defined as "large-scale components of the Earth system that may pass a tipping point" (both definitions Lenton et al. 2008). A key feature of a tipping element is that current temperature affects the likelihood of a tipping element experiencing a *tipping event*—that is, a transition to an irreversible climate process, called a *tipping process*. Examples of tipping processes include the irreversible melting of the Greenland ice sheet, the collapse of the west Antarctic ice sheet, and the weakening of the Atlantic thermohaline circulation. The incorporation of tipping elements in DSICE allows it to model a range of potentially irreversible climate processes, including but not limited to global

[4]When quantifying temperature increases throughout this paper, we are referring to the year 1900 temperature level as the baseline level. Thus, our term "temperature change" is analogous to the term "temperature anomaly", which is often used in the climate literature and typically denotes the difference from some baseline temperature level.

catastrophes. These processes are uncertain, and give DSICE a second source of risky damage to economic productivity.[5]

There is a consensus that SCC specifically and optimal policy in general are sensitive to the choice of discount rate used by the social planner. Standard IAMs are deterministic, treat mitigation expenditures as a way of increasing consumption in the future, and use the same discount rate for all issues. Some have argued that mitigation has an insurance component. Schneider (1989), in his testimony to the Committee on Energy and Commerce in 1989, argued that investing in climate change mitigation is like "buying insurance against the real possibility of large and potentially catastrophic climate change." More recently, former Secretary of State George Shultz (Shultz, 2013) , who convinced President Reagan to support the Montreal Protocol, argued

> "We all know there are those who have doubts about the problems presented by climate change. But if these doubters are wrong, the evidence is clear that the consequences, while varied, will be mostly bad, some catastrophic. So why don't we follow Reagan's example and take out an insurance policy?"

R&D is an example of insurance spending. The new technologies arising from today's R&D spending will arrive only with significant lags, and, if we are lucky, we may not need to use them, just as we do not file a claim to our insurer if we experience no bad events. Insurance has a negative expected rate of return, but we buy it because of its option value. Similarly, simple discounting rules motivated by a deterministic model are not valid in a dynamic stochastic context because they would ignore the real option aspect of R&D investments.

The results in this study demonstrate the importance of adding both economic and climate risk to IAMs. First, economic risks imply that there is great uncertainty about future SCC and the value of future climate change policies. In particular, the possibility of climate change causing substantial damage is much higher in DSICE than in models that ignore economic risks. Second, while SCC today is sensitive to parameter choices regarding preferences, we do arrive at a robust finding that

[5]Catastrophic climate change—in our case a climate tipping process, is just one of many possible types of critical events that might ultimately affect economic productivity. A more general concept of ecological thresholds of the Earth is the concept of *planetary boundaries* (Rockström et al. 2009, Steffen et al. 2015). Ocean acidification, loss of biodiversity and ozone depletion are examples of Earth system processes that are believed to exhibit specific boundaries—that is to say, thresholds that lead to irreversible and abrupt environmental change. Furthermore, possible catastrophic events that affect the global economy are, of course, not limited ecological ones. Most recently, Martin and Pindyck (2015, 2017) consider a broad variety of potential catastrophes such as a mega-virus, a nuclear or bioterrorism attack or a climate catastrophe. They study the willingness to pay to avoid these extreme events which, interdependently, do not only affect consumption dynamics but also might be fatal to the population.

SCC is nontrivial, with \$40–\$100 being the range implied by the various opinions regarding dynamic preferences. Third, there is no simple discounting rule to apply to climate change policy decisions. The IAM literature argues over what is the right discount rate for valuing future damages from climate change (see, e.g., Stern and Taylor 2007, Nordhaus 2007). DSICE shows that there is no one discount rate. As standard finance theory teaches, the appropriate discount rate depends on covariance with consumption. In this study, we show that the damages arising from tipping events should be discounted at a much lower rate than damages from temperature increases. This may seem surprising, because we follow Nordhaus's approach and use the market equilibrium stochastic discount factor instead of some planner discount rate unrelated to the market pricing of risks. The intuition is clear: the damages from a permanent shock to productivity have little covariance with short-run fluctuations in consumption. Therefore, damages from tipping events should be discounted less than damages from temperature (which are proportional to output). Fourth, DSICE examines cases where both economic and climate risks are present and shows that the results differ significantly from a separate examination of these two sources of risk. More generally, this study shows that it is important to examine these issues in a model that incorporates multiple kinds of uncertainty.

Many studies express great skepticism about what is computationally feasible, and use this as a reason to look at simple models. Traeger's (2014) claim that the curse of dimensionality limits what can be done is typical:

> "The quantitative analysis of optimal mitigation policy under uncertainty requires a recursive dynamic programming implementation of integrated assessment models. Such implementations are subject to the curse of dimensionality. Every increase in the dimension of the state space is paid for by a combination of (exponentially) increasing processor time, lower quality of the value or policy function approximations, and reductions of the uncertainty domain."

The computational mathematics literature has developed many methods that avoid the curse of dimensionality,[6] and DSICE uses them to solve problems generally considered intractable. This study solves nine-dimensional stochastic dynamic programming problems over multiple centuries. The non-stationary character of the problems makes value function iteration the only possible approach.

[6] See Judd (1998) for an elementary introduction to some of those methods.

The specifications of risks make these problems among the most computationally demanding ever solved in economics. However, DSICE is not limited by any curse of dimensionality because it uses efficient multivariate methods to approximate value functions. Even though DSICE solves millions (billions in some examples) of optimization problems, this is possible because it uses reliable optimization methods and parallelization to solve the Bellman equations (Cai and Judd 2010; Cai et al. 2015). The "curse of dimensionality" is not a valid excuse for oversimplified modeling.

Any numerical computation has numerical errors, and the fact that DSICE must solve billions of small numerical problems means that we need to subject any results to stringent accuracy checks. In recent years, the scientific computing community has addressed this issue in its literature on Verification, Validation, and Uncertainty Quantification (VVUQ); see Oberkampf and Roy (2010) for a comprehensive discussion of this literature. The "Verification" part of the VVUQ literature develops methods for verifying the accuracy of the numerical results. We develop such a test for each value function iteration and find that every value function computed by DSICE passes demanding verification tests. This gives us confidence that numerical errors do not affect our economic conclusions. Some of the cases we present below required tens of thousands of core hours, and sensitivity analysis demanded that we examine hundreds of cases to determine the robustness of results across empirically plausible parameter values. This study required the use of a few million core hours on Blue Waters, a modern supercomputer. DSICE is based on general-purpose numerical methods, implying that many economics problems with similar computational requirements can now be solved.

We present the climate model in Section 2 and the economic model in Section 3. In Section 4, we formulate the dynamic programming problem and outline our solution method. Sections 5, 6, and 7 present and discuss the implications for SCC from stochastic specifications of factor productivity growth and a Markov chain process; first each in isolation and then combined. Section 8 concludes. This paper uses a large number of variables, reflecting the complexity of DSICE. For the convenience of the reader, all variables and their definitions appear in tables.

2 The Climate Model

Our climate model contains three modules. First is the carbon system, describing how carbon diffuses across the atmosphere, the upper ocean, and the deep ocean. Second is the temperature system, which describes how heat energy diffuses between the atmosphere, ocean, and space. The carbon and temperature systems are not closed, because economic and biological activity injects carbon into the atmosphere, carbon interacts with solar radiation to cause heating, and some heat is lost through radiation to space. We follow the specification for the carbon and temperature modules used in DICE-2007 (Nordhaus, 2008).

We add a third system that models climate conditions other than carbon and temperature. A simple example of such a climate state is sea level. Prolonged periods of warm atmospheric temperatures will (likely) melt ice in glaciers (on land), which will then lead to higher sea levels. This makes sea levels and the condition of glaciers dependent on past temperatures. We will focus on what are called tipping elements, which model irreversible changes. For example, if warming causes the melting of glaciers, even if temperatures fall back to pre-industrial levels, glaciers reappear, if at all, only after millions of years (IPCC, 2014).

Climate change affects economic productivity in different ways. Higher temperatures will likely reduce agricultural output, raise expenditures on cooling, and facilitate the spread of disease. There is evidence that higher temperatures will increase the likelihood of severe droughts and floods (Wuebbles, 2016). Higher sea levels will worsen coastal flooding, and may even cause some land to disappear. We provide details on the productivity effects below when we describe the economic model in DSICE.

2.1 Carbon System

We assume two sources for carbon emissions at each time t, an industrial source, $E_{\text{Ind},t}$, related to economic activity, and an exogenous source, $E_{\text{Land},t}$, arising from biological processes on the ground[7]. Total emissions are denoted by

$$E_t \equiv E_{\text{Ind},t} + E_{\text{Land},t}. \tag{1}$$

[7]These emissions depend e.g. on many biological processes and are also subject to uncertainty. However, here we adopt the DICE specification of land emissions which are assumed to be exogenous and deterministic.

The details of $E_{\text{Ind},t}$ will be presented when we discuss the economic model. We follow DICE-2007 and aggregate the distribution of carbon in the world into three "boxes"– atmosphere, upper ocean, and lower ocean. The three-dimensional vector $\mathbf{M}_t = (M_{\text{AT},t}, M_{\text{UO},t}, M_{\text{LO},t})^\top$ represents the mass of carbon in the atmosphere, upper levels of the ocean, and lower levels of the ocean, respectively (in gigatons of carbon (GtC)). The carbon concentrations evolve over time according to the physics of diffusion, and are represented by the linear dynamical system

$$\mathbf{M}_{t+1} = \mathbf{\Phi}_M \mathbf{M}_t + (E_t, 0, 0)^\top, \tag{2}$$

where

$$\mathbf{\Phi}_M = \begin{bmatrix} 1-\phi_{12} & \phi_{21} & 0 \\ \phi_{12} & 1-\phi_{21}-\phi_{23} & \phi_{32} \\ 0 & \phi_{23} & 1-\phi_{32} \end{bmatrix}. \tag{3}$$

The coefficient ϕ_{ij} in $\mathbf{\Phi}_M$ is the rate at which carbon diffuses from level i to level j, where $i, j = 1, 2, 3$ represent the atmosphere, upper ocean, and lower ocean, respectively. If $E_t = 0$, then this would be a closed system, implying that the column sums of $\mathbf{\Phi}_M$ must be unity. Also note that there is no diffusion between the atmosphere and the lower ocean. Table A.4 lists the parameter values for the carbon system. The exogenous processes and calibration of the module appear in Appendices A.1 and A.3.

2.2 Temperature System

The DICE family of models is based on the continuous-time differential equation system in Schneider and Thompson (1981). DSICE will also use one-year time periods in all examples in this paper. The temperature system tracks the temperatures of the atmosphere (T_{AT}) and the ocean (T_{OC}), measured in °C above the pre-industrial level. That system is governed by the diffusion of heat, is represented by the vector $\mathbf{T}_t = (T_{\text{AT},t}, T_{\text{OC},t})^\top$, and evolves according to

$$\mathbf{T}_{t+1} = \mathbf{\Phi}_T \mathbf{T}_t + (\xi_1 \mathcal{F}_t (M_{\text{AT},t}), 0)^\top, \tag{4}$$

where

$$\mathbf{\Phi}_T = \begin{bmatrix} 1-\varphi_{21}-\xi_2 & \varphi_{21} \\ \varphi_{12} & 1-\varphi_{12} \end{bmatrix}. \tag{5}$$

The coefficient φ_{ij} is the heat diffusion rate from level i to level j, where $i, j = 1, 2$ represent the atmosphere and ocean, respectively. The coefficient ξ_2 is the rate of cooling arising from infrared radiation to space (Schneider and Thompson 1981), and ξ_1 represents heating due to radiative forcing. Atmospheric temperature rises through two sources of radiative forcing: an exogenous level, $F_{\mathrm{EX},t}$, and endogenous forcing due to carbon in the atmosphere. Total radiative forcing at t is

$$\mathcal{F}_t\left(M_{\mathrm{AT},t}\right) = \eta \log_2\left(M_{\mathrm{AT},t}/M^*_{\mathrm{AT}}\right) + F_{\mathrm{EX},t}, \tag{6}$$

where M^*_{AT} is the pre-industrial atmospheric carbon concentration, and η is the radiative forcing parameter. Table A.5 lists the parameter values for the temperature system. The exogenous processes and calibration of the module appear in Appendices A.1 and A.3.

2.3 Tipping Element System

The temperature and carbon systems in DSICE are deterministic and evolve smoothly over time. Recent work in the IAM literature has drawn attention to tipping points, which we defined earlier in our introduction. Tipping points are not necessarily irreversible, but they are essentially irreversible for the planning horizons related to the examples considered in this paper.

Let J represent some feature of the climate other than temperature or carbon. It will have a finite set of possible values, and will represent the state of a tipping element; we will refer to J as the tipping state. Changes in J will be modeled by a Markov chain where transition probabilities depend on the vector of all climate states, ($\mathbf{T}$,$\mathbf{M}$,J). The Markov transition process is denoted as

$$J_{t+1} = g_J(\mathbf{T}_t, \mathbf{M}_t, \omega_{J,t}),$$

where $\omega_{J,t}$ is one serially independent stochastic process.

The key properties of any tipping element include the likelihood of tipping events, the expected duration of the tipping process, the mean and variance of the long-run impacts on economic produc-

tivity, and how all of these depend on ($\mathbf{T}$,$\mathbf{M}$,J). There is substantial uncertainty about all of these properties. Our examples will be based on expert opinion expressed in the climate science literature on various tipping elements. We use Lenton's (2010) summary of the findings from Kriegler et al. (2008) and other expert elicitation studies to calibrate the likelihood of transitions in tipping elements. We will also rely on damage estimates used in the IAM literature (e.g., IPCC 2014; Smith et al. 2009; Stern 2007; Nordhaus 2008; and Hope 2011). The impact on economic productivity is included in our description of the economic system, below. Section 6 examines some specific cases and will precisely describe the Markov process for J. Our calibration of the tipping element system appears in Appendices A.1 and A.4.

2.4 Comparisons with Other IAM Climate Systems

DSICE, as in DICE-2007, uses a five-dimensional system (two dimensions for temperature and three for the carbon cycle) to compute paths of world average carbon concentrations and temperature levels that are close to the results from much more complex models. DSICE uses a modified version of the temperature module in DICE-2007. Cai et al. (2012a) point out that the computer code in Nordhaus (2008) has temperature increases in each period depend on carbon emissions ten years in the future. Cai et al. (2012a) use differential equation methods to solve the continuous-time physical model in Schneider and Thompson (1981), and find that applying the Euler method to the differential equations in Schneider and Thompson with one-year time steps produces the discrete-time system defined above in Equations (2) and (4), which successfully matches the MAGICC scenarios (Meinshausen et al., 2011) that DICE-2007 aimed to match. The differential equation approach produced results significantly different from those in DICE-2007. To avoid confusion about the different solution methods, DICE-2007 will refer to the model and results in Nordhaus (2008), and DICE-CJL will refer to the model specified in Cai et al. (2012a). The mathematical details of this calibration method are presented in Appendix A.3.

DICE is regarded as the canonical climate system in much of the IAM literature, but few studies use it, often citing tractability issues. Golosov et al. (2014) track only carbon concentration in the atmosphere to represent the whole climate system, implying that temperature is proportional to carbon concentrations. Jensen and Traeger (2014) also use a one-box climate model with only

atmospheric carbon concentration. The models in Golosov et al. (2014) and Jensen and Traeger (2014) ignore the lag between CO2 emissions and their impact on temperature lag climate scientists tell us is in the order of decades. Lemoine and Traeger (2014) use a two-box climate model, tracking only atmospheric carbon concentration and temperature, ignoring the impact of oceans on atmospheric temperature and CO2 concentrations.

Others have studied tipping elements but make simplifying assumptions that do not conform with physical evidence. A common approach is to assume that there is a threshold (perhaps unknown to the planner) such that a tipping event will occur immediately when temperature crosses that threshold (see, e.g., Keller et al. 2004). With economic uncertainty, temperatures can fall or rise. If temperature were to cross a threshold, the tipping event would immediately happen even if the temperature quickly fell below that threshold.

More recent studies assume that the full impact of a tipping event is immediate *and* that its level is known (see, e.g., Lemoine and Traeger 2014). In our framework, that is equivalent to assuming there are only two climate states related to tipping: pre-tipping and post-tipping. Our tipping element specification is also the first, and so far only one, to address the recent critique by Kopits et al. (2014) of these assumptions. In DSICE, the timing of a tipping event is unknown, even conditional on knowing the full state of the climate system, and the timing of the transitions after the tipping event is unknown, and—furthermore—the impact of the Markov chain is unknown prior to the tipping event (Lenton et al. 2008; National Research Council 2013). DSICE is so far the only model that can handle these characteristics of Markov chain points. Other applications of DSICE have considered various transition times of Markov chain points (Cai et al. 2015, 2016; Lontzek et al. 2015).

Earlier studies ignore expert opinion and climate science when they calibrate the process producing tipping events. Lemoine and Traeger (2014), for example, keep track of the historically maximum temperature level, and assume that a tipping event cannot be triggered as long as temperature is below that maximum level. This assumption ignores the inertia in the climate system, implying that long-lasting heating dynamics that might be present in the climate system could trigger climate tipping events. It also rules out the possibility that tipping might be triggered during periods of decreasing temperature trends. DSICE, instead, relies on expert opinion (Kriegler et al. 2009 and Lenton et al. 2008) to calibrate its tipping elements, and assumes that a tipping event

is a random event with probability depending on the state of the climate system, assuming that the hazard rate of the tipping point event is increasing with global warming. Lontzek et al. (2015) present a detailed discussion of studies on climate tipping in stochastic IAMs.

Several recent studies look at analytically tractable, climate–economy formulations. Subsequent to our work, Anderson et al. (2014), for example, conduct a robustness analysis of model uncertainty in a simple integrated assessment model. Van der Ploeg and de Zeeuw (2016) decompose analytically SCC in the face of catastrophic climate events. Gerlagh and Liski (2016) present an analytically tractable model of learning about impacts from climate change. Other numerical integrated assessment models (e.g., Kelly and Kolstad (1999), Kelly and Tan (2015) and Hwang et al. (2017)) deal with uncertainty by incorporating Bayesian learning about the parameter describing climate sensitivity—that is to say, the equilibrium temperature change in response to changes of the radiative forcing (by carbon dioxide).[8] More generally, Bayesian methods could be applied to integrate model uncertainty into IAMs.[9] While this approach can generate insightful results regarding the evaluation of different policies, it is not the focus of our study. Instead, we apply an uncertainty quantification analysis to determine the robustness of our results across empirically plausible parameter values.

3 The Economic Model

DSICE merges a basic dynamic stochastic general equilibrium model with the commonly used box model for climate in DICE-2007. This allows us to explore how economic risks and climate risks interact and affect the evaluation of climate change policies.

3.1 Pre-Damage Output

The economic side of DSICE is a simple stochastic growth model where production produces greenhouse gas emissions and output is affected by the state of the climate. We assume time is discrete, with each period equal to one year. Let K_t be the world capital stock in trillions of dollars at time t, and L_t be the world population in millions at time t.

[8] See, e.g., Roe and Baker (2007).

[9] See, e.g., Brock et al. (2007) and Cogley et al. (2011) for an analysis in macroeconomics.

The deterministic parts of our economic model are taken directly from the calibrated DICE-2007 model in Nordhaus (2008). This includes the production function, population growth, productivity growth, the carbon intensity of output, and the damage due to temperature levels.[10] In the absence of any climate damage, the gross world product is the Cobb–Douglas production function

$$f(K, L, \widetilde{A}_t) = \widetilde{A}_t K^{\alpha} L^{1-\alpha},$$

where $\alpha = 0.3$ (as in Nordhaus, 2008) and $\widetilde{A}_t$ is productivity at time t. Productivity is decomposed into two pieces: a deterministic trend A_t, and a stochastic productivity state ζ_t; that is to say, $\widetilde{A}_t \equiv \zeta_t A_t$. The deterministic trend A_t is taken from Nordhaus (2008) and denoted as

$$A_t = A_0 \exp\left(\alpha_1 (1 - e^{-\alpha_2 t})/\alpha_2\right), \tag{7}$$

where α_1 is the 2005 growth rate and α_2 is the decline rate of the growth rate.

We add a stochastic component, ζ_t, to the productivity process so that we can examine how uncertainty about productivity affects climate change policies. Our specification of ζ_t uses ideas from the long-run-risk literature (e.g., Bansal and Yaron 2004, Hansen et al. 2008). We let χ_t represent the persistence of ζ_t, and use the formulation introduced in Bansal and Yaron (2004):

$$\log(\zeta_{t+1}) = \log(\zeta_t) + \chi_t + \varrho \omega_{\zeta,t} \tag{8}$$

$$\chi_{t+1} = r\chi_t + \varsigma \omega_{\chi,t}, \tag{9}$$

where $\omega_{\zeta,t}, \omega_{\chi,t} \sim i.i.d.\, \mathcal{N}(0,1)$ and ϱ, r, and ς are parameters.

For theoretical and computational reasons, we need to change some details of the Bansal and Yaron (2004) specifications. Bansal and Yaron (2004) assume that $\omega_{\zeta,t}, \omega_{\chi,t} \sim i.i.d.\, \mathcal{N}(0,1)$. As Gaussian disturbances are unbounded and would imply the possibility of arbitrarily large growth rates and output. The unbounded character of the state space makes it difficult to even prove that expected utility exists. This creates computational challenges. Even if we could overcome

[10] The DSICE framework is flexible enough to allow for various functional forms of its components. However, to maximize comparability with the DICE model that is currently used in research for the design of US regulatory policy, we retain the assumptions of the DICE model.

these theoretical and computational challenges, our results for SCC could be driven by highly unlikely tail events. Others have examined the impact of very bad outcomes that have very small probabilities. An example of this is the 'dismal' theorem of Weitzman (2009), which states that the risk premium could be infinite for unboundedly distributed uncertainties. We want to avoid existence issues and avoid repeating insights regarding extreme tail events. We also want to use numerically stable computational methods that allow us to verify our results. We achieve these goals by constructing a time-dependent, finite-state Markov chain for (ζ_t, χ_t) with parameter values implying conditional and unconditional moments of consumption processes observed in market data. The Markov transition processes are denoted

$$\zeta_{t+1} = g_\zeta(\zeta_t, \chi_t, \omega_{\zeta,t})$$

$$\chi_{t+1} = g_\chi(\chi_t, \omega_{\chi,t}),$$

where $\omega_{\zeta,t}$, and $\omega_{\chi,t}$ are two serially independent stochastic processes. This approach also makes it possible to directly apply reliable numerical methods for solving dynamic programming problems.

The approximation of the stochastic growth process requires a careful formulation of the Markov chains for the productivity shock ζ_t and the rate of its growth persistence χ_t. Markov chains with only a few states cannot represent the kind of persistence properties observed in Bansal and Yaron (2004). After examining various possibilities, we choose $n_\zeta = 91$ values of ζ_t and $n_\chi = 19$ values of χ_t at each time t; the time dependence is required due to the fact that the variance of consumption levels grows over time. Appendix A.2 describes the discretization procedure in greater detail.

We calibrate the stochastic factor productivity growth so that the resulting consumption process is statistically close to empirical data. Calibrating the productivity process presents some computational challenges because we need to solve the economic model in DSICE for various values of ϱ, r, and ς, and choose those values, which imply a consumption process that matches US data on per capita consumption growth.[11] For our calibration, we solve versions of DSICE without stochastic climate impact because the consumption data we use comes from the twentieth century when climate damage to productivity was negligible. Appendix A.2 describes the calibration procedure in

[11]We are grateful to Ravi Bansal for providing us with the annual per capita data on real consumption used in Bansal et al. (2012) and obtained from the Bureau of Economic Analysis website.

greater detail, and shows that the statistics of simulation paths of our consumption growth from our calibrated parameters are close to those of the empirical data. The results of the calibration gave us the following values:

$$\varrho = 0.035,\ r = 0.775,\ \varsigma = 0.008.$$

3.2 Damage Function and Emissions

DSICE models two potential ways in which output is affected by the climate: global average temperature T_{AT}, and the Markov chain state denoted by J. The function $\Omega\left(T_{\mathrm{AT},t}, J_t\right)$, referred to as the damage function, represents the impact of climate on output in that gross world product equals

$$Y_t \equiv \Omega\left(T_{\mathrm{AT},t}, J_t\right) f\left(K_t, L_t, \zeta_t A_t\right),$$

where

$$\Omega\left(T_{\mathrm{AT},t}, J_t\right) = \Omega_T\left(T_{\mathrm{AT},t}\right)\Omega_J\left(J_t\right) = \frac{1}{1+\pi_1 T_{\mathrm{AT},t} + \pi_2 (T_{\mathrm{AT},t})^2}(1 - D(J_t))$$

decomposes the damage function as the product of Ω_T (damage due to temperature rise), and Ω_J (damage related to the climate conditions implied by the Markov process, J. When $D(J_t) = 0$, our damage function reduces to $\Omega_T\left(T_{\mathrm{AT},t}\right)$ for which we use the damage function in Nordhaus (2008).[12] $D(J)$ equals the impact of state J on productivity. This study generalizes the damage function to include effects of the tipping state and associated past cumulative effects along with the impact of temperature on productivity.

We assume that industrial emissions are proportional to output, with the proportionality factor σ_t, which is referred to as the carbon intensity of output. The social planner can mitigate (i.e., reduce) emissions by a factor μ_t, $0 \leq \mu_t \leq 1$. Annual industrial carbon emissions (billions of metric tons of carbon) equal

$$E_{\mathrm{Ind},t} = \sigma_t(1-\mu_t) f\left(K_t, L_t, \zeta_t A_t\right). \tag{10}$$

[12] The damage function in DSICE is, of course, highly aggregated, and only to be considered as an approximation of smooth damages from global warming. Recently, several studies have attempted to quantify climate-related damages on a disaggregated scale: e.g., Deschenes and Greenstone (2011) and Dell et al. (2014) make use of weather data to assess the impact of temperature on a local scale. While this approach certainly generates important insight, more research is needed, e.g., to quantify longer-term impacts of temperature increase. Here, we make these choices to keep the model simple, and to facilitate comparisons with Nordhaus's canonical model. The only general requirement is that $\Omega\left(T_{\mathrm{AT},t}, J_t\right) > 0$.

We follow Nordhaus (2008) and assume that the cost of mitigation level μ_t is

$$\Psi_t = \theta_{1,t}\mu_t^{\theta_2} Y_t. \tag{11}$$

World output, net of damage, is allocated across total consumption C_t, mitigation expenditures Ψ_t, and gross capital investment I_t;that is to say,

$$Y_t = C_t + \Psi_t + I_t, \tag{12}$$

and the capital stock evolves according to

$$K_{t+1} = (1-\delta)K_t + I_t, \tag{13}$$

where $\delta = 0.1$ is the annual depreciation rate for capital. In all of our computations, $K_0 = 137$ trillion dollars.

3.3 Epstein–Zin Preferences

The additively separable utility functions commonly used in climate–economy models do not do well in explaining the willingness of people to pay to avoid risk. Epstein–Zin preferences (Epstein and Zin 1989) are used because they better explain observed equity premia. The equity premium tells us about society's willingness to pay to reduce consumption risk, which will in turn affect how much society is willing to pay to reduce the risk of economic damage from climate change.

We assume that agents care only about consumption.[13] Let C_t be the stochastic consumption process. Epstein–Zin preferences recursively define social welfare as

$$U_t = \left\{ (1-\beta)\, u(C_t, L_t) + \beta \left[\mathbb{E}_t \left\{ U_{t+1}^{1-\gamma} \right\} \right]^{\frac{1-1/\psi}{1-\gamma}} \right\}^{\frac{1}{1-1/\psi}}, \tag{14}$$

where $\mathbb{E}_t\{\cdot\}$ is the expectation conditional on the states at time t, and β is the discount factor.

[13] There may be other aspects of climate change that affect social welfare, but they are not included in DSICE, or in the standard Nordhaus (2008) family of models. See, for example, the Cai et al. (2015) version of DSICE including uncertain, non-market impacts such as ecosystem tipping points.

Here,

$$u(C_t, L_t) = \frac{(C_t/L_t)^{1-1/\psi}}{1-1/\psi} L_t$$

is the annual world utility function (assuming that each individual has the same power utility function), ψ is the inter-temporal elasticity of substitution,[14] and γ is the risk-aversion parameter. Epstein–Zin preferences are flexible specifications of decision-makers' preferences regarding uncertainty, and allow us to distinguish between risk preference and the desire for consumption smoothing. Even though we refer to γ as the risk-aversion parameter, equilibrium risk premia will depend on interactions between ψ and γ.

We will be comparing the results of DSICE with results of deterministic models such as DICE. Those models assume inter-temporally separable preferences. It is important to note that in the absence of uncertainty, the risk-aversion parameter, γ, disappears and Epstein–Zin preferences become inter-temporally separable. In deterministic models with time-separable utility, it is common to refer to $1/\psi$ as the "risk-aversion" parameter, but that is misleading and implicitly assumes $\gamma = 1/\psi$. A better way to think about deterministic models is that they use Epstein–Zin preferences with ψ representing the IES, but the risk-aversion parameter, γ, can take any value. This observation will be important when we discuss the impact of adding risks to DICE.

Empirical analyses of macroeconomic data have not given us definitive values for the parameters γ and ψ . We examine a range of parameter values, relying on the literature on long-run risk and asset pricing, in particular Bansal and Yaron (2004), Barro (2009), Bansal and Ochoa (2011), Pindyck and Wang (2013), Epstein et al. (2014) and Schorfheide et al. (2014). Most studies estimate or assume γ to be between 2 and 10, and argue for $\psi > 1$. Table 1 summarizes the range of values for ψ and γ in the literature. Our benchmark parameter specification will be $\psi = 1.5$ and $\gamma = 10$, choices consistent with the majority of opinion. We will also solve DSICE for a broad range of values covering $2 \leq \gamma \leq 15$ and $0.5 \leq \psi \leq 2.0$.

[14] Here we assume $\psi > 1$. When $0 < \psi < 1$, the utility function $u(C_t, L_t)$ is negative, and the formula becomes

$$U_t = -\left\{-(1-\beta)\, u(C_t, L_t) + \beta \left[\mathbb{E}_t \left\{(-U_{t+1})^{1-\gamma} \mid C_t, L_t\right\}\right]^{\frac{1-1/\psi}{1-\gamma}}\right\}^{\frac{1}{1-1/\psi}}.$$

While the standard formulation of Epstein–Zin preferences does not have a denominator of $1 - 1/\psi$ in the annual world utility function $u(C, L)$, here we use this formulation to ensure that it is consistent with our later reformulation of the Bellman equation (15).

	IES	RA
Bansal & Yaron (2004)	1.5	10
Bansal and Ochoa (2011)	1.5	10
Vissing-Jørgensen and Attanasio (2003)	>1	[5, 10]
Barro (2009)	2	4
Pindyck and Wang (2013)	1.5	3.066
Ackerman et al. (2013)	1.5	10
Constantinides and Ghosh (2011)	1.41	9.43
Schorfheide et al. (2014)	1.7	10.8
Epstein et al. (2014)	1.5	7.5
Jensen and Traeger (2014)	1.5	10

Table 1: Default or estimated values of IES and the risk-aversion parameter in the literature.

3.4 Relation to the Literature

Some have used Epstein–Zin preferences to examine SCC, but use much simpler models. Bansal and Ochoa (2011) assume an endowment model, implying that their SCC only tells one how to price the exogenous shocks exerted by the climate on the endowment. DSICE, instead, builds on a Ramsey-type, representative agent, stochastic growth model where the agent at each time must choose how to allocate output across consumption and savings. Savings are split between capital investment and mitigation expenditures, both of which are forms of spending that aim to improve future productivity. Our SCC represents the social tradeoff between spending resources on investment and on mitigation.

IAMs have only recently used long-run risk and Epstein–Zin preferences. Jensen and Traeger (2014) assume much lower levels of volatility than those implied by empirical data. See Appendix A.2 for a detailed comparison. The higher volatility in DSICE (which is calibrated to fit empirical data) presents computational challenges because the state variables cover a much larger region in the $(K, \mathbf{M}, \mathbf{T}, \zeta, \chi)$ space. DSICE is successful because it uses a flexible set of functions for approximations.

4 The Dynamic Programming Problem

We formulate the nine-dimensional, social planner's dynamic optimization problem as a dynamic programming problem. The nine states include six continuous state variables (the capital stock K, the three-dimensional carbon system $\mathbf{M}$, and the two-dimensional temperature vector $\mathbf{T}$) and three

discrete state variables (the climate shock J, the stochastic productivity state ζ, and the persistence of its growth rate χ). Let $\mathbf{S} \equiv (K, \mathbf{M}, \mathbf{T}, \zeta, \chi, J)$ denote the nine-dimensional state variable vector and let $\mathbf{S}_+$ denote its next period's state vector.

The Epstein–Zin utility definition (14) expressed utility in terms of consumption. We make a nonlinear change of variables[15] in terms of utils, $(U_t)^{1-\frac{1}{\psi}}$, and then get the following Bellman equation

$$\begin{aligned} V_t(\mathbf{S}) = \max_{C,\mu} \quad & u(C_t, L_t) + \beta \left[\mathbb{E}_t \left\{ \left(V_{t+1}(\mathbf{S}_+) \right)^{\frac{1-\gamma}{1-1/\psi}} \right\} \right]^{\frac{1-1/\psi}{1-\gamma}}, \\ \text{s.t.} \quad & K^+ = (1-\delta)K + Y_t - C_t - \Psi_t, \\ & \mathbf{M}^+ = \mathbf{\Phi}_M \mathbf{M} + (E_t, 0, 0)^\top, \\ & \mathbf{T}^+ = \mathbf{\Phi}_T \mathbf{T} + (\xi_1 \mathcal{F}_t(M_{\mathrm{AT}}), 0)^\top, \\ & \zeta^+ = g_\zeta(\zeta, \chi, \omega_\zeta), \\ & \chi^+ = g_\chi(\chi, \omega_\chi), \\ & J^+ = g_J(\mathbf{T}, \mathbf{M}, J, \omega_J), \end{aligned} \tag{15}$$

for $t = 0, 1, \ldots, 599$, and any $\psi > 1$.[16] The terminal value function V_{600} is given in Appendix A.5.[17] In the model, consumption C and emission control rate μ are the two control variables.

4.1 Numerical Solution Method

We solve the nine-dimensional problem specified in (15) using value function iteration. Three state variables, (ζ, χ, J), are discretized and their movements are modeled as transitions of finite-state Markov chains. The productivity process states, (ζ, χ), use Markov chains that have enough states to ensure that the resulting consumption processes matches the conditional variance and autocorrelation observed in consumption data. Furthermore, J is calibrated to represent processes discussed in the climate literature on tipping points. At each discrete point in the (ζ, χ, J) space, the

[15] That is, $V_t(\mathbf{S}) = [U_t(\mathbf{S})]^{1-\frac{1}{\psi}} / (1-\beta)$.

[16] When $0 < \psi < 1$, the objective function of the optimization problem is

$$u(C_t, L_t) - \beta \left[\mathbb{E}_t \left\{ \left(-V_{t+1}(\mathbf{S}_+) \right)^{\frac{1-\gamma}{1-1/\psi}} \right\} \right]^{\frac{1-1/\psi}{1-\gamma}}.$$

[17] As in Nordhaus (2008) we assume a model time horizon of 600 periods. The terminal value function should be regarded as an approximation of the value function at the terminal period.

value function has six continuous states, $(K, \mathbf{M}, \mathbf{T})$, and is approximated by multivariate orthogonal polynomials after appropriate nonlinear changes of variables. The range of each continuous state variable is chosen so that all simulation paths stay in that range. This is a large problem, but the use of parallel programming methods and hardware makes it tractable. For example, the largest case has 366 billion optimization problems, but we solved it in less than eight hours using 110,688 cores in parallel on the Blue Waters supercomputer. See Appendices A.2 and A.5 of this paper and Cai et al. (2015) for an extended discussion of the mathematical and computational details.

4.2 A Verification Test

One theme of the Validation, Verification, and Uncertainty Quantification (VVUQ) literature (e.g., Oberkampf and Roy 2010) is the implementation of tests that check the correctness of the computer code; this is called verification. One common test is to apply the code to special cases where the solution is known. If all uncertainty is removed, then DSICE reduces to a deterministic optimal control problem that can be solved using nonlinear programming methods. We compare these optimal control solutions to our value function iteration results to see if the value functions imply an optimal path in line with the results using nonlinear programming. Our tests show that paths implied by the value functions have at least three-digit accuracy, often significantly more. See Appendix A.7 for more details.

4.3 The Social Cost of Carbon

The climate literature interprets the social cost of carbon as a marginal concept—that is to say, the monetized economic loss caused by a one metric ton increase in atmospheric carbon. We follow that notation: in DSICE, the social cost of carbon is the marginal cost of atmospheric carbon expressed in terms of the numeraire good, which can be either consumption or capital as there are no adjustment costs. We define the social cost of carbon (SCC) to be the marginal rate of substitution between atmospheric carbon concentration and capital, as in

$$\mathrm{SCC}_t = -1000 \left(\partial V_t / \partial M_{\mathrm{AT},t}\right) / \left(\partial V_t / \partial K_t\right). \tag{16}$$

It will be important to remember that SCC is a relative shadow price—that is to say, a ratio of two marginal values—and does not express the total social cost of climate damage.[18] For example, as we change economic and/or climate risks, SCC may go up or down because that change in risks will affect the marginal value of carbon, and the marginal value of consumption, and the marginal value of investment. DSICE is a general equilibrium model where the results arise from assumptions about tastes and technology, as well as their equilibrium interactions.

SCC will often be the optimal *carbon tax*. The optimal carbon tax is the tax on carbon that would equate the private and social costs of carbon. In DSICE we also examine the optimal carbon tax, which is the Pigovian tax policy because the externality from carbon emissions can be directly dealt with by a carbon tax and because there are no other market imperfections. The social planner in DSICE chooses mitigation μ_t, which is equivalent to choosing a carbon tax equal to

$$1000\theta_{1,t}\theta_2\mu_t^{\theta_2-1}/\sigma_t \tag{17}$$

in units of dollars per ton of carbon. If $\mu_t < 1$, the carbon tax equals SCC. However, if $\mu_t = 1$, its maximum value, then the carbon tax only equals that level that will drive emissions to zero, and may be far less than SCC. In such cases, mitigation policies have reached their limit of effectiveness. Alternative policies may reduce carbon concentrations directly, as would carbon removal and storage technologies, or reduce temperature directly, as would certain solar geoengineering technologies. We do not explicitly include those technologies in DSICE, but our SCC numbers will identify equilibrium paths along which SCC is so high that these more direct technologies may be competitive. [19]We leave a quantitative analysis of these issues to future studies.

The Business-As-Usual (BAU) case is the solution to the dynamic programming problem (15) but with the restriction that $\mu = 0$, i.e., no mitigation. It is used in the IAM literature as an approximation of the competitive equilibrium in the absence of any carbon tax. We will generally not report any BAU results because they differ little from the optimal mitigation results.

[18]Because K is measured in trillions of dollars and M_{AT} is measured in billions of tons of carbon, the 1,000 factor is needed to express SCC in units of dollars per ton of carbon.

[19]Robock et al. (2009) critically assess geoengineering options in the context of decision-making on climate change and point out severe side effects of geoengineering, as well as the lack of knowledge associated with this technology. Recently, Heutel et al (2016) presents one of the first studies to address geoengineering in a stochastic IAM. That study focuses on the effectiveness of geoengineering to deal with different types of damage from a tipping point. Further research is needed to improve understanding of how different sources of uncertainty affect the effectiveness of geoengineering as an option to combat climate change.

4.4 Simulation Procedures

In Sections 6, 7, and 8, respectively, we analyze the implications for climate policy of only stochastic growth, only stochastic climate, and stochastic growth and stochastic climate combined. In each of these sections we define a benchmark parameter specification and show its implications for SCC and other economic variables. We then perform a sensitivity analysis on some parameters, using tables to report how today's optimal level of SCC is affected by different parameter choices.

All three sections follow the same procedure. First, we solve the dynamic programming problem, computing the value function and policy decision rules at each time t. We then use the decision rules to simulate some paths of the key economic variables. We use 2005 as the first year in order to be comparable with Nordhaus (2008) and similar studies. These paths are affected by shocks to productivity; we simulate 10,000 paths that differ only in the realized shocks. This set of paths allows us to not only compute the expected value of SCC and economic variables, but also compute their distribution at any time t. The large number of simulations allows us to compute even the 0.99 percentile, and gives us a good measure of the important tail events. Each simulation begins with the stochastic growth states equal to $\chi_0 = 0$ and $\zeta_0 = 1$. Other initial values were given in Tables A.4 and A.5. In cases with a tipping element, we assume that a tipping point has not yet been triggered in the year 2005.

5 The Social Cost of Carbon with Stochastic Growth

This section analyzes the impact of stochastic growth and risk preferences on SCC and on the combined climate and economic system. This section excludes tipping elements so that it can focus on the impact of productivity risk. We first present a detailed analysis of a benchmark example based on parameter specifications; we call this our *stochastic growth benchmark* case. The stochastic growth benchmark case assumes Epstein–Zin parameter values $\psi = 1.5$ and $\gamma = 10$. This stochastic growth benchmark case allows us to exposit key features of the resulting dynamic processes, such as consumption, output, productivity, climate states, and SCC. We then perform a sensitivity analysis for empirically plausible alternative preference parameters, focusing on how preference assumptions affect SCC in 2005.

5.1 The Stochastic Growth Benchmark

Figures 1 and 2 display features of the solution to the Bellman equation (15) for the stochastic growth benchmark case. Each panel summarizes the results of our 10,000 simulations. We also display results from two deterministic cases of DICE-CJL. DICE-CJL with $\psi = 0.5$ (the solid red line) represents the DICE-2007 choice of ψ, and DICE-CJL with $\psi = 1.5$ (dashed red line) represents the choice in our stochastic growth benchmark. The gray area represents the 0.01 to 0.99 percentiles of the SCC paths, and the other lines represent various quantiles.

Figure 1 displays the SCC process for 2005–2050. First note that moving from the DICE-2007 choice of $\psi = 0.5$ to our benchmark choice of $\psi = 1.5$ substantially increases SCC. However, when we add uncertainty with $\gamma = 10$, we see a decline in SCC to \$61/tC, which is still higher than that implied by DICE-2007 preferences. During 2005–2050, we see that the no-risk line (the dashed red line) exceeds the average SCC with long-run risk (the black line) by about \$50/tC. Therefore, in the next half century, uncertainty reduces SCC. We shall discuss the intuition for this below when we carry out sensitivity analysis. The productivity risk does substantially increase the range of SCC. By 2050, there is a 25 percent chance of SCC being almost \$200/tC or greater, and a 10 percent chance of it exceeding \$250/tC. The large range is easily explained. If there is a sustained sequence of positive shocks to productivity, then output will rise, which will have two effects: the damage due to temperature will increase proportionately, and emissions will also rise and increase the risk of even greater damage from high temperatures in the future. Recall that the SCC results in Figure 1 represent the solution to the optimal tax and mitigation policy.

Figure 2 describes several variables over the twenty-first century. Panel A describes the SCC process. After 2050, average SCC in DSICE moves closer to SCC in DICE-CJL with $\psi = 1.5$. By 2100, the impact of uncertainty is small. Therefore, while uncertainty substantially lowers SCC prior to 2050, that effect nearly disappears over the following 50 years. After 2050, the variation in SCC in DSICE continues to grow rapidly. The 1–99 percentile range is \$67/tC to \$1,282/tC, and even the 10 percent and 90 percent quantiles in 2100 show a range of \$127/tC to \$667/tC.

Panel B displays the optimal carbon tax and Panel C displays the optimal rate of mitigation. If mitigation is less than 100 percent, then a Pigovian tax policy would be to impose a tax equal to SCC, thereby equating the private and the social costs of carbon. Panels A, B, and C show that

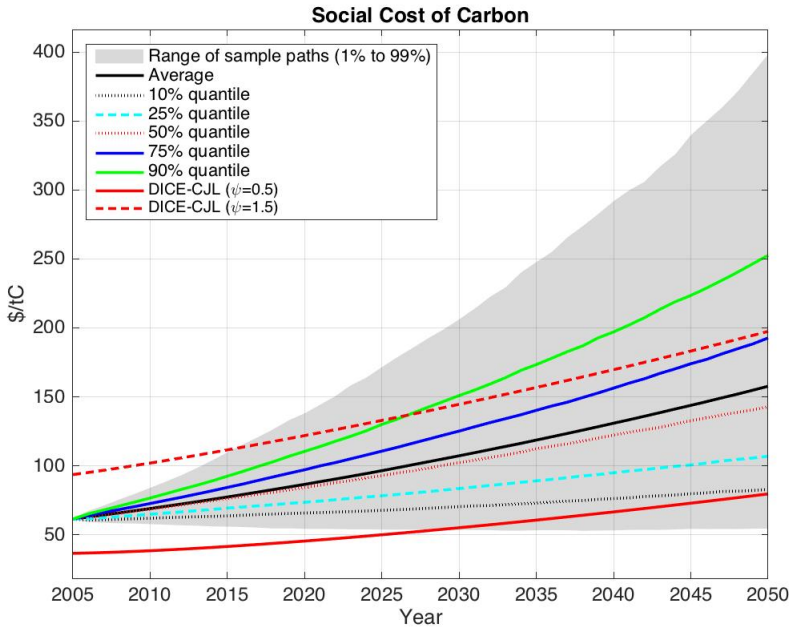

Figure 1: SCC ($/tC): 2005–2050

this is always the case before 2085. After 2085, the emission control rate may hit its limit of $\mu = 1$, and optimal policy requires only that the carbon tax is large enough to eliminate all emissions. In that case, the optimal tax can be much less than SCC, which is true at 2100 in about 3.5 percent of simulation runs. When the optimal tax is less than SCC, the benefits of other carbon policies, such as carbon removal and storage or solar geoengineering technologies, would be valued using the SCC.

Table 2 lists the mean and standard deviation of SCC_t and their statistics on a $\log_{10}$ scale.[20] Table 2 tells us that both the mean and the standard deviation of $\log_{10}(\text{SCC}_t)$ are increasing over time. In fact, the standard deviation of $\log_{10}(\text{SCC}_t)$ is increasing much faster than its mean.

	mean of SCC_t	s.d. of SCC_t	mean of $\log_{10}(\text{SCC}_t)$	s.d. of $\log_{10}(\text{SCC}_t)$
2020	87	18	1.924	0.087
2050	158	71	2.153	0.184
2100	357	247	2.457	0.279

Table 2: SCC statistics from 10,000 simulation paths

Both panels D and E indicate that long-run risk will on average lead to more accumulation of carbon in the atmosphere and higher temperature, a natural consequence of the reduced mitigation

[20] We include the $\log_{10}$ scale because the SCC distribution is more like a log-normal distribution.

efforts. However, the economic risk implies uncertainty in atmospheric temperature in 2100, with the 10–90 percentile range being 0.6°C.

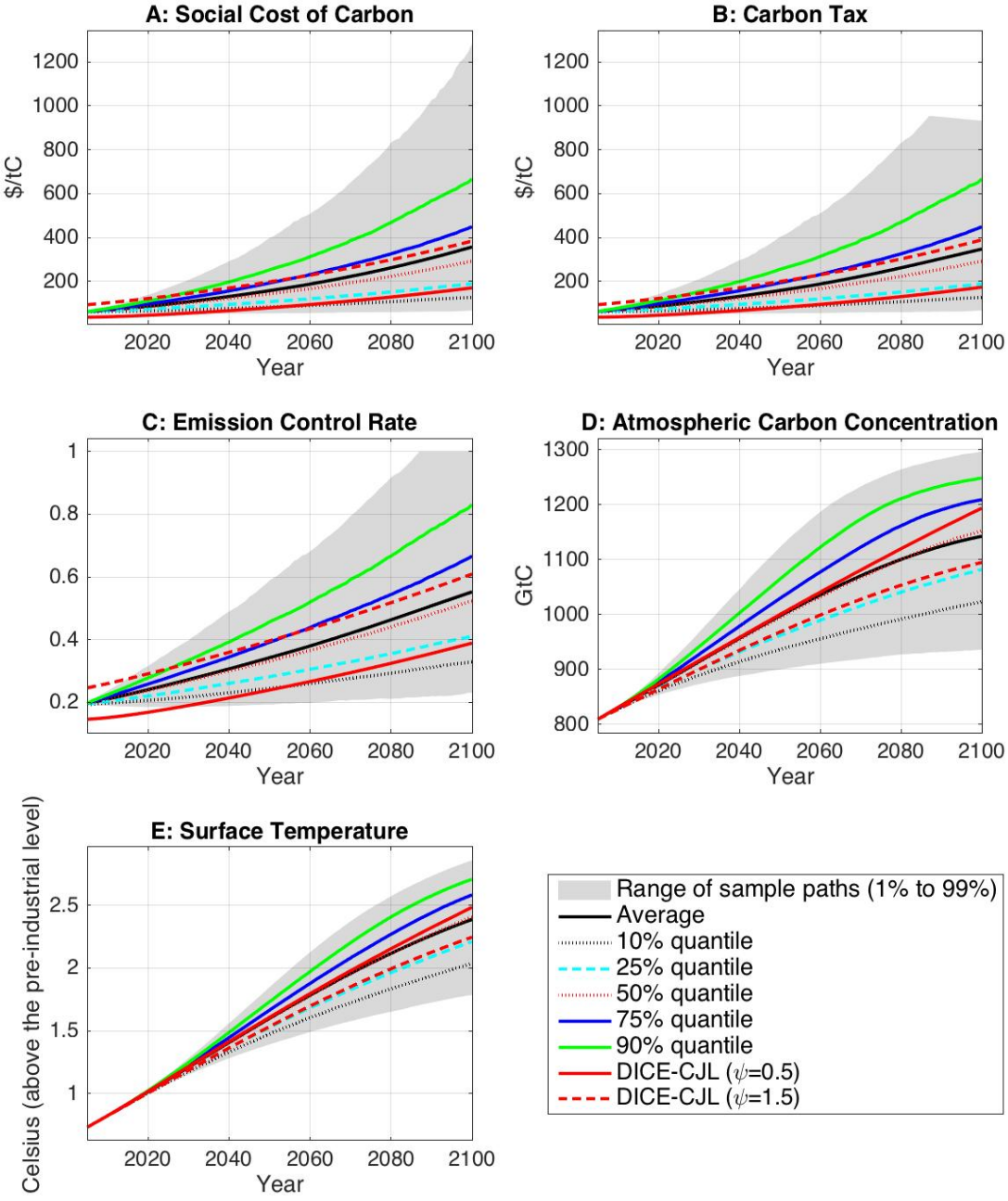

Figure 2: Simulation results of the stochastic growth benchmark—climate system and policies

Our findings point to one very important fact: there is great uncertainty about all aspects of the combined economic and climate system. For many variables, the mean value at each point in time is close to the solution of the purely deterministic model. Tracking the mean is all one can ask of any deterministic model, and in that sense deterministic models can be successful. However, there is great uncertainty about the future value of each key variable. This fact is of particular importance for understanding SCC. SCC is the marginal cost of extra carbon in terms of wealth, making it the marginal rate of substitution between mitigation expenditures and investment expenditures in physical capital. At the margin, these two uses of savings have different impacts on future economic variables, making allocation decisions between mitigation and investment essentially a portfolio choice problem; a large SCC represents the amount of investment in new capital that one is willing to sacrifice to reduce carbon emissions by a gigaton.

In comparison to a model with purely deterministic growth, DSICE implies a lower ratio of consumption to gross world output and higher ratios of investment in capital accumulation and abatement. Figure 3 presents the details, displaying the optimal dynamic distribution of the four ratios for the first 100 years with various quantiles of 10,000 simulations of the solution to the dynamic programming problem (15).

For example, the black dotted lines represent the 10 percent quantiles for each ratio. Similarly, the cyan dashed lines, the red dotted lines, the blue solid lines, and the green solid lines represent the 25%, 50%, 75%, and 90% quantiles at each time respectively. The black solid lines represent the sample mean path. As explained earlier, two cases of DICE-CJL ($\psi = 0.5$ or 1.5) makes it comparable with DICE-2007. We denote these two special deterministic cases by the red solid lines and the red dashed lines, respectively. The lower (upper) edge of the gray areas represent the 1% (99%) quantile; the gray areas represent the 98% probability range of each ratio.[21]

From Figure 3 we see that with more than 90 percent probability I_t/Y_t will be greater than in the case of a purely deterministic model (the red solid line is below the black dotted line). Furthermore, we find that in 2005 I_t/Y_t is at 32 percent, about 8 percent higher than under the deterministic growth assumption and that the expected difference is about 5% toward the end of this century. Overall, the assumption of stochastic factor productivity growth with persistence leads to a significant expected increase in capital investments and thus to a precautionary buildup

[21] We will use the same graphical exposition for all plots describing the distributions of simulation results.

of the capital stock.[22] Also, throughout the 10,000 simulations of DSICE, we found I_t/Y_t to ranged roughly between 22 and 36 percent, expanding the distributional results reported in Table 2.

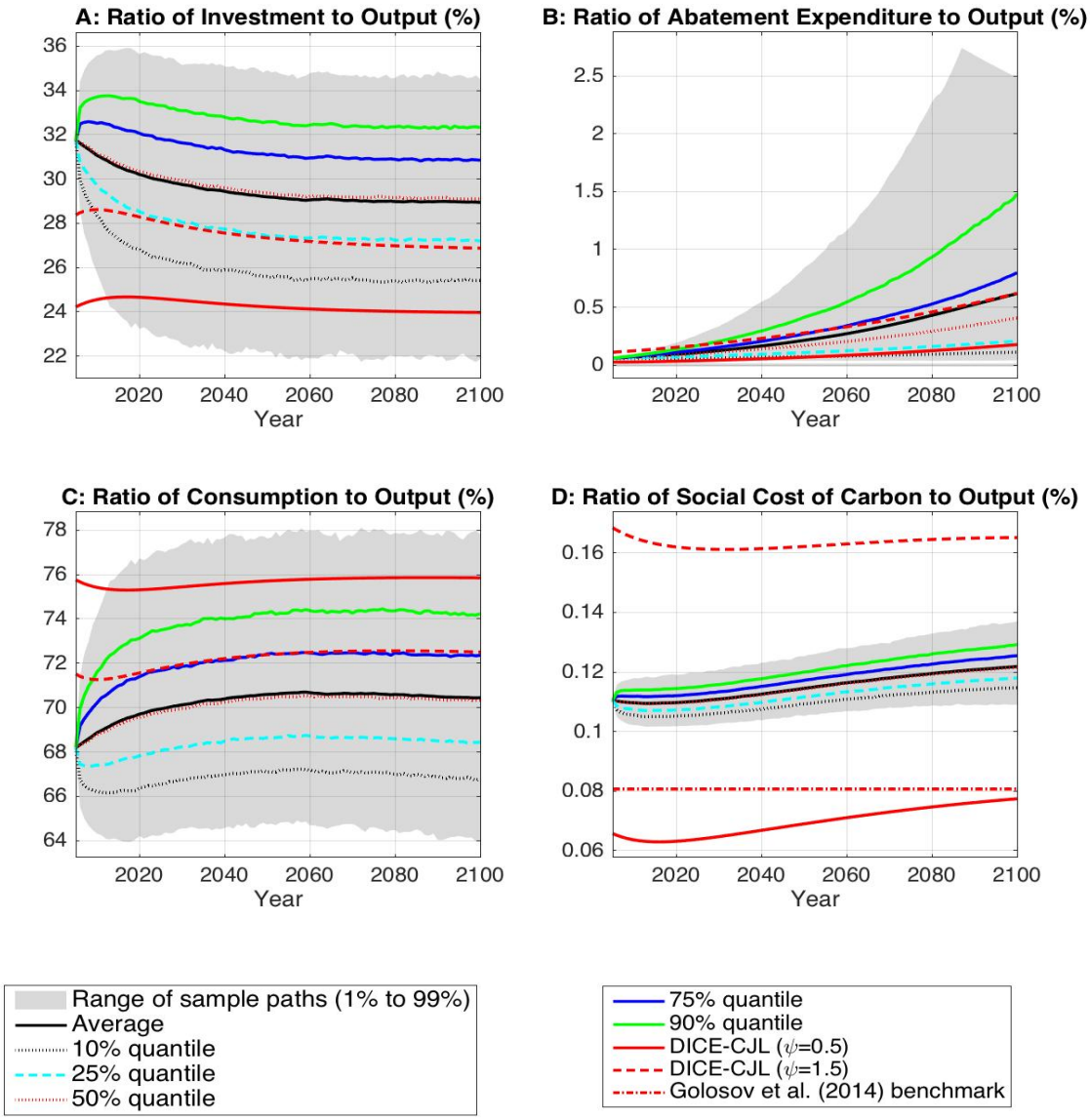

Figure 3: Simulation results of the stochastic growth benchmark—ratios to gross world output

Panel C presents quite the opposite statistical picture for C_t/Y_t. We see that with more than 90 percent probability C_t/Y_t will be lower than in a purely deterministic model, and that toward the end of this century C_t/Y_t appears to stabilize at about 70 percent–a reduction of about 6 percent

[22]The simulation paths of gross world output, capital, and per capita consumption growth are shown in Figure A.2 in Appendix A.8.

over the deterministic model. Overall, this reduction is not fully offset by higher capital investments and, as panel B indicates, that difference is allocated to expenditures on the abatement of emissions. We find that the latter, which is denoted by Ψ_t/Y_t, is generally quite low and does not exceed 0.2 percent in this century in the deterministic case. Yet, as the black solid line in panel B indicates, there is a 50 percent probability that the expenditures on emissions abatement will be at least three times higher by the year 2100 when growth is modeled stochastically. Furthermore, with more than 20 percent probability, more than 1 percent of gross world output should be devoted to mitigation by the year 2100.

Some recent research has argued for simple rules of thumb for SCC. In particular, Golosov et al. (2014) set up a dynamic, forward-looking climate–economy model with logarithmic utility and full capital depreciation to argue that SCC is proportional to output. Barrage (2014) shows that the benchmark in Golosov et al. (2014) implies that the ratio of SCC to decadal gross world output is 8.07×10^{-5} (i.e., $\text{SCC}_t/Y_t = 8.07 \times 10^{-4}$ for annual gross world output Y_t) and constant over time with constant productivity growth, but that it increases over time for the productivity process in Nordhaus (2008), approaching 8.07×10^{-4} from below.[23]

Panel D in Figure 3 uses red dash-dotted lines to represent the results of Golosov et al. (2014). However, our Panel D shows that the SCC_t/Y_t ratio is stochastic and that its mean or quantile paths have an upward trend, while its volatility is also expanding. Its mean path increases from 0.11 percent in 2005 to 0.12 percent in 2100. This implies that long-run risk shows that the constant SCC_t/Y_t in Golosov et al. (2014) does not hold true in general.

5.2 Sensitivity Analysis for Preference Parameters

Empirical work suggests plausible values for ψ and γ, but the data do not give us precise values for the key parameters. We next examine how SCC varies across values of ψ and γ. Each example will differ from the stochastic growth benchmark only in the preference specification, with no change in the stochastic growth productivity process.

Our sensitivity analysis of IES (i.e., ψ) and RA (i.e., γ) in stochastic growth cases will look at consumption, capital investment, and SCC in 2005. Tables 3–6 report the sensitivity of the initial

[23] More precisely, we report the ratio $\text{SCC}_t/1000/Y_t$ in order to have the same units as those used in Golosov et al. (2014).

ratios of consumption to gross world output (C_t/Y_t), capital investment to gross world output (I_t/Y_t), and SCC, respectively, assuming the values of the elasticity of inter-temporal substitution to be $\psi = 1.25$, 1.5, 1.75, and 2.0, and the risk-aversion parameter to be $\gamma = 2$, 5, 10, and 15 (we also include the cases with $\psi = 0.5$ and/or $\gamma = \infty$ for comparison, and $\psi = 0.7$, 0.9, 1.1 for SCC in Table 6). For example, in our benchmark example with $\psi = 1.5$ and $\gamma = 10$, the 2005 C_t/Y_t is 0.682, the 2005 I_t/Y_t is 0.682, and the optimal 2005 SCC is \$61. We see that SCC ranges from \$57 to \$99 for these combinations of $1.25 \leq \psi \leq 2$ and $2 \leq \gamma \leq 15$, about 60 percent to 168 percent larger than the 2005 SCC of the baseline deterministic model with $\psi = 0.5$.

IES (ψ)	DICE-CJL	RA (γ)				
		2	5	10	15	∞
0.50	0.76	0.75	0.73	0.71	0.69	0.58
1.25	0.72	0.71	0.70	0.69	0.68	0.64
1.50	0.72	0.71	0.69	0.68	0.68	0.65
1.75	0.71	0.70	0.69	0.68	0.67	0.66
2.00	0.71	0.69	0.68	0.67	0.67	0.67

Table 3: Initial consumption–output ratio in stochastic growth cases

IES (ψ)	DICE-CJL	RA (γ)				
		2	5	10	15	∞
0.50	0.24	0.29	0.27	0.29	0.31	0.42
1.25	0.28	0.29	0.30	0.31	0.32	0.36
1.50	0.28	0.29	0.31	0.32	0.32	0.35
1.75	0.29	0.30	0.31	0.32	0.33	0.34
2.00	0.29	0.31	0.32	0.33	0.33	0.33

Table 4: Initial investment–output ratio in stochastic growth cases

ψ	DICE-CJL	RA (γ)				
		2	5	10	15	∞
0.5	2.6(−4)	2.9(−4)	4.1(−4)	5.7(−4)	6.6(−4)	8.0(−4)
1.25	9.1(−4)	8.4(−4)	7.0(−4)	5.9(−4)	5.5(−4)	3.9(−4)
1.5	1.1(−3)	9.9(−4)	7.6(−4)	5.9(−4)	5.5(−4)	3.5(−4)
1.75	1.3(−3)	1.1(−3)	8.1(−4)	6.3(−4)	5.5(−4)	3.3(−4)
2.0	1.4(−3)	1.2(−3)	8.6(−4)	6.4(−4)	5.6(−4)	3.1(−4)

Table 5: Initial abatement-expenditure-to-output ratio in stochastic growth cases. Note that $a(-n)$ represents $a \times 10^{-n}$

IES (ψ)	DICE-CJL	RA (γ)				
		2	5	10	15	∞
0.5	37	39	49	60	66	76
0.7	51	52	56	60	62	62
0.9	64	63	61	60	60	55
1.1	75	71	65	60	58	50
1.25	82	77	68	61	58	48
1.5	94	86	71	61	57	45
1.75	103	93	74	62	57	43
2	111	99	76	62	57	41

Table 6: Initial social cost of carbon ($/tC) in stochastic growth cases

Tables 3–6 show that the patterns of the impact of IES and RA in the stochastic growth cases are similar to those of IES and productivity growth in the deterministic growth cases. We use the extreme case of $\gamma = \infty$ to provide some intuition. In that case, the planner focuses on the worst-case scenario, which in our discretization of the productivity process means that decisions are the same as if productivity were deterministic and constantly declining, allowing us to use deterministic optimal control methods for that case; see Appendix A.9 for details. The last column of Table 6 displays how IES affects economic quantities when $\gamma = \infty$. In this case, DSICE is essentially optimal growth with declining productivity. A small IES implies a strong desire to smooth per capita consumption, and save more today to reduce a future decline in consumption. As IES increases, smoothing consumption is less of a priority, implying less savings in both forms, capital and mitigation, and a smaller SCC.

When there is no risk (the "DICE-CJL" column of Table 6) or if risk-aversion is small, then the patterns are reversed. In these cases, the positive future growth in productivity implies that as IES increases, there is less desire for smoothing and more willingness to save for the future, when productivity is higher. SCC and the share of savings going to mitigation rise as IES increases, indicating that the relative importance of avoiding climate damage increases.

Table 6 shows that SCC is decreasing over RA when $\psi \geq 0.9$, and increasing over RA when $\psi \leq 0.7$. From the above discussion, we see that RA in DSICE is related to productivity growth in DICE-CJL: $\gamma = \infty$ corresponds to DICE-CJL with negative productivity growth, and a smaller RA corresponds to DICE-CJL with a higher productivity growth. Appendix A.10 shows that SCC from DICE-CJL with a range of ψ and 2005 productivity growth rate has a pattern similar to that

of Table 6.

5.3 Comparisons with Others' SCC Estimates

Our results differ from other studies on SCC estimates, but usually for understandable reasons. Golosov et al. (2014) obtains a relatively high SCC due to the assumption that all global warming effects from carbon emissions occur instantaneously, ignoring the lags in the climate system. Anthoff and Tol (2014) use their model, FUND, to argue for lower estimates for SCC in the early twenty-first century because their disaggregated damage function takes into account the fact that increases in CO2 plus mild warming can improve agricultural productivity in the upper latitudes. However, FUND does not allow agents to optimize consumption decisions dynamically. Future versions of DSICE will aim to include more sophisticated damage functions.

IWG (2010) finds that lowering the rate of discounting for damages from 3 percent to 2 percent increases SCC by a factor of 1.64. Another major component that affects SCC is the dynamic structure of damages caused by global warming. For example, Dietz and Stern (2015) assume much higher damages due to high temperature than does DICE (when the temperature increase is 4°C, the damage is 50 percent of output in Dietz and Stern (2015), but only 4 percent in DICE-2007). We do not examine these questions extensively in this study, but the limited experimentation we did indicate that DSICE responds to these changes in the same way.

6 Social Cost of Carbon with Stochastic Climate Tipping

We next study how a tipping element in the climate system may affect SCC in the absence of any economic uncertainty. First, we present a Markov chain specification of a representative climate tipping element. We then specify benchmark parameters calibrated to a "representative" scenario, and study the optimal climate policy; we call this our *climate tipping benchmark*. In a final step we present a multidimensional sensitivity analysis. In light of the numbers provided in the few studies available, we conclude that the impact of potential tipping point events should be carefully assessed.

6.1 A Markov Chain Specification of the Climate Tipping Process

We study a simple example of tipping, with uncertainty about the time of tipping and the long-run damage. $\mathcal{J}_0$ will denote the initial state, which we call the *pre-tipping state*, and $D(\mathcal{J}_0) = 0$. "Tipping" is the time and event when the climate leaves state $\mathcal{J}_0$ and moves to some other state in the tipping process. We follow the common assumption that warming alone causes tipping (IPCC 2014; Smith et al. 2009), and assume that the probability that a tipping event occurs at time t is a function of temperature,

$$p_{\text{tip},t} = 1 - \exp\left\{-\lambda \max\left\{0,\, T_{\text{AT},t} - \underline{T_{\text{AT}}}\right\}\right\}, \tag{18}$$

where λ is the hazard rate parameter and $\underline{T_{\text{AT}}}$ is the temperature for which $p_{\text{tip},t} = 0$. The hazard rate parameter, λ, together with the temperature process determines the duration of the initial state $\mathcal{J}_0$.

In this study, we examine a simple tipping process that models the gradual nature of the tipping process and uncertainty about the ultimate damage caused. After a tipping event, the state J follows one of three possible processes, each one also modeled by a Markov chain. Define $\mathcal{M}_i$, for $i = 1$, 2, or 3, to be a Markov chain with states $\{\mathcal{J}_{i,1}, \mathcal{J}_{i,2}, \mathcal{J}_{i,3}, \mathcal{J}_{i,4}, \mathcal{J}_{i,5}\}$. Each process $\mathcal{M}_i$ moves in sequence through transient tipping states $\mathcal{J}_{i,2}$, $\mathcal{J}_{i,3}$, $\mathcal{J}_{i,4}$, and ultimately arrives at the absorbing state $\mathcal{J}_{i,5}$.

The damage factor in state $\mathcal{J}_{i,j}$ is $\mathfrak{D}_{i,j} = D(\mathcal{J}_{i,j})$. The damages at the terminal states, $\{\mathfrak{D}_{1,5}, \mathfrak{D}_{2,5}, \mathfrak{D}_{3,5}\}$ represent three different levels of long-run permanent damages. Let J_∞ denote the random long-run state, and let $\mathfrak{D}_\infty$ denote the random long-run damage from tipping (i.e., $\mathfrak{D}_\infty = D(J_\infty)$).

Therefore, we begin in state $\mathcal{J}_0$, leave $\mathcal{J}_0$ at time t with probability $p_{\text{tip},t}$, and after a tipping event J jumps to one of $\{\mathcal{J}_{1,1}, \mathcal{J}_{2,1}, \mathcal{J}_{3,1}\}$ with equal probability. Before the tipping event, we do not know which $\mathcal{M}_i$ process will be followed after tipping and we do not know the final level of damage. However, at the tipping time, we learn which $\mathcal{M}_i$ governs the post-tipping process and the long-run damage.

This is a stark simplification, but it does allow us to distinguish the expected duration of the transient post-tipping process, $\overline{\mathcal{D}}$, from the uncertainty about the ultimate damage level. The fact

that the uncertainty about $\mathfrak{D}_\infty$ is resolved when the climate tipping process is triggered allows us to make statements about the relative impact of the hazard rate of the tipping point event, the expected duration of the whole transient post-tipping process, and the mean and variance of $\mathfrak{D}_\infty$. The unconditional mean of $\mathfrak{D}_\infty$ is denoted $\overline{\mathfrak{D}}_\infty$, and the variance of $\mathfrak{D}_\infty$ is $q\overline{\mathfrak{D}}_\infty^2$, where q is called the "mean squared-variance ratio". The ratio q is analogous to the square of the Sharpe ratio, a concept used in portfolio theory, and arises naturally in our discussion of results.

In our examples, for each post-tipping Markov chain $\mathcal{M}_i$, we assume that its expected duration from the time when the tipping event occurs to the time when the tipping state reaches the absorbing state is known, and denoted $\overline{\Gamma}$. We also refer to it as the *expected duration of the post-tipping process*. There are four transient stages in the post-tipping process, and we assume that each transient stage from $\mathcal{J}_{i,j}$ to $\mathcal{J}_{i,j+1}$ has the same expected duration[24] and that all transitions have an exponential distribution, implying that the transition probability is $p = 1 - \exp\left(-4/\overline{\Gamma}\right)$.[25] The complete mathematical description of J and its calibration are contained in Appendix A.4.

6.2 SCC with Stochastic Climate Tipping

We choose parameter values that are roughly the average of the range of opinions in the literature.[26] More precisely, the climate tipping benchmark case assumes $\lambda = 0.0035$, $\overline{\mathfrak{D}}_\infty = 0.05$, $q = 0.2$, and $\overline{\Gamma} = 50$. The choice of $\lambda = 0.0035$ implies that the conditional annual probability of tipping increases by 0.35 percent for a warming of 1 °C. The Epstein–Zin preference parameters in the climate tipping benchmark case are again $\psi = 1.5$ and $\gamma = 10$.

First, we use Figure 4 to show two sample paths of damage to output, $1 - \Omega(T_{\mathrm{AT},t}, J_t)$, in percentage terms (the solid lines) and their corresponding paths for SCC (the dashed lines). The two lines in the left panel represent one sample realization of a tipping process that tips in 2146 with the largest long-run damage level, and the two lines in the right panel represent one sample realization of a tipping process that tips in 2102 with the smallest long-run damage level. The sample paths of damage to output clearly show the four transient post-tipping stages (jumps).

The realized duration of the whole transient post-tipping process is 38 years for the sample path

[24] Technically, DSICE could easily handle cases where the the duration of each stage depends on the entire state space.

[25] Experimentation indicated that five post-tipping stages is adequate to approximate processes with more states.

[26] A discussion of that literature is contained in Appendix A.4.

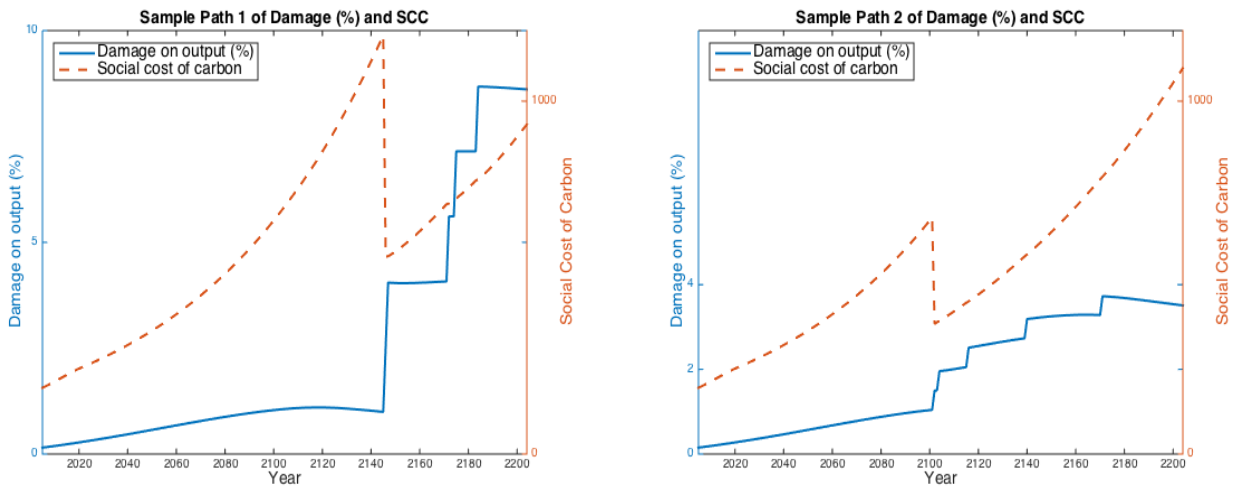

Figure 4: Sample paths of damage on output and social cost of carbon for the stochastic climate tipping benchmark

in the left panel (from 2146 to 2184) and 69 years for the sample path in the right panel (from 2102 to 2171), while the expected duration of the post-tipping process, $\overline{\Gamma}$, is 50 years. Moreover, the realized final damage level $\mathfrak{D}_\infty$ is 7.74 percent for the sample path in the left panel and 2.26 percent for the sample path in the right panel, while its expectation is 5 percent. Therefore, the jumps in SCC paths just represent the fact that the decision with regard to emissions mitigation strongly depends on whether the tipping process has been triggered or not. This is due to the fact that once the tipping process is triggered, we cannot prevent or delay the sequential damages that occur during its multiple post-tipping stages and thus there is no strong incentive to change mitigation policy, such as strong discontinuous changes in the SCC.

It is obvious that adding a tipping element to DICE will increase SCC. The question is how much SCC is increased given the magnitude of the tipping point damages. Figure 5 helps addressing that issue. It is based on computing the BAU case with and without tipping. The red line shows the relative increase in damage when we add the tipping element. The increase is negligible until 2050, and peaks in the late 2100s at 40 percent. Reducing carbon emissions in 2005 would delay any tipping and shift the red curve to the right, but most of those changes would occur after 2100. The black line shows the impact of tipping on the marginal social cost of carbon, and shows that the marginal damage of carbon rises by 60% in 2005. Even though the tipping element only moderately increases damages 150 years from now, the impact on SCC today is substantial.

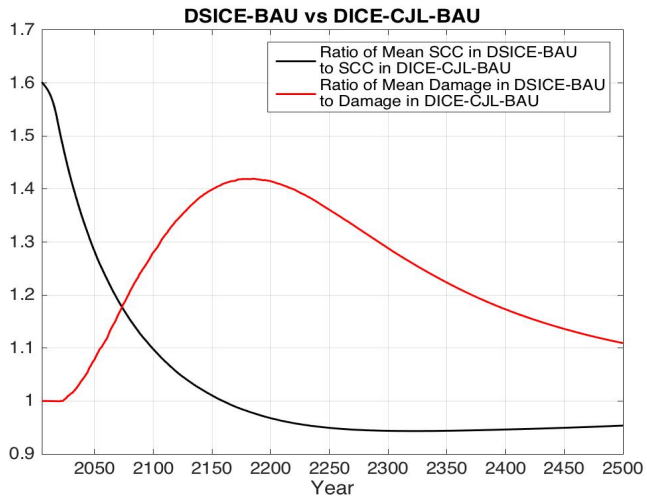

Figure 5: Comparison of DSICE and DICE-CJL with $\psi = 1.5$ under Business-As-Usual (BAU)

We next consider the implications of a climate tipping point for the dynamics of SCC, the carbon tax, the emissions control rate, and the two most important climate states (atmospheric carbon concentration and surface temperature), respectively. Figure 6 shows the results of 10,000 simulation paths over the first 200 years for these variables.

Adding climate tipping risk is expected to result in a more intense mitigation compared to that of a deterministic model. In fact, as panel C in Figure 6 shows, the optimal emissions control rate follows a pattern related to that of the abatement expenditures in the previous figure. Throughout this century it is optimal to more than double mitigation efforts in response to the threat of a tipping point in the climate. Emissions reductions imply a strict reduction of atmospheric carbon concentrations (panel D) compared to the result from the deterministic model. The resulting path of surface temperature (panel E) corresponds to the temperature paths from the lowest of the most recent emissions scenarios used by the IPCC (2013), implying a peak temperature increase before 2100 of around 2°C and a decline afterward.

A striking result in panel A of Figure 6 is the 2005 SCC of \$188/tC, a large increase over the \$94/tC resulting from DICE-CJL (where $\psi = 1.5$) in the absence of any tipping risk. To underline the significance of this major increase in SCC, recall our rather conservative assumptions on the nature of the tipping point's processes: we assume an expected duration of the tipping process of 50 years, expected post-tipping damage of 5%, and a mean squared-variance ratio of 0.2. As can be seen from the blue dashed line in panel A of Figure 6, these assumptions indicate that there

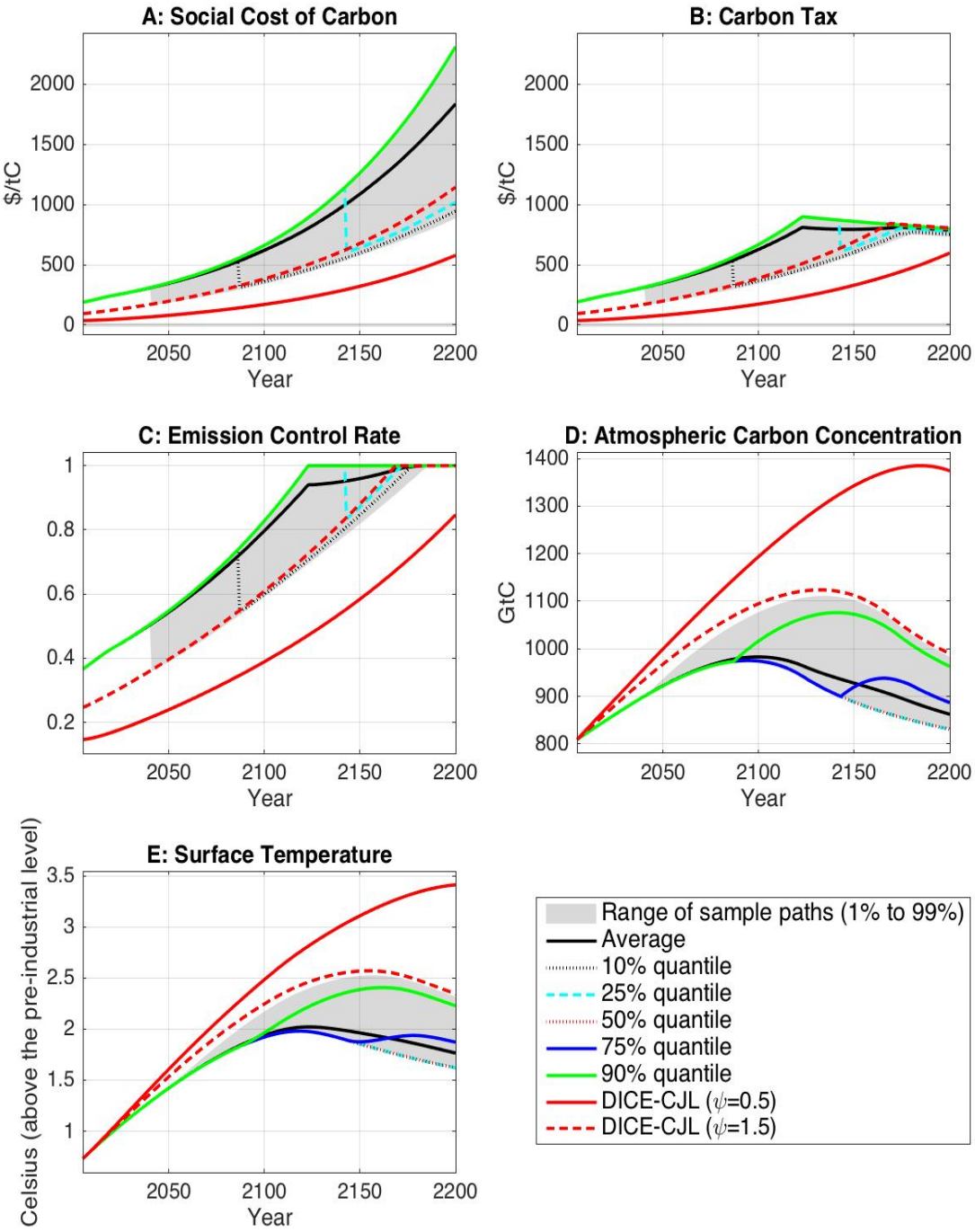

Figure 6: Simulation results for the climate tipping benchmark—climate system and policies

is a 75 percent probability that a tipping process will not be triggered before 2150. Yet today's optimal SCC is \$188/tC, twice the value obtained from the same model when the tipping point is ignored. This strongly suggests that analyses of climate policy that simply ignore the potential of abrupt changes to the climate system—as does the current United States government study (IWG 2010)—are significantly underestimating SCC.

Related to the analysis in the previous section, we also note here that by the year 2125 some of our 10,000 simulated paths will produce a carbon tax that is less than SCC, indicating that the benefits of mitigation may be exhausted. In fact, it appears that, with a slightly higher than 75 percent probability, mitigation policies will reach the limits of their effectiveness by 2125 and alternative carbon management options might prove useful.

6.3 Discounting Damages

To understand the impact of a tipping element on SCC, we compute the discount rate of damages from carbon emissions, denoted by ρ_Δ, which is implicitly used to value the marginal damage; that is, we solve the following equation:

$$\mathrm{SCC}_0 = \sum_{t=0}^{\infty} (1+\rho_\Delta)^{-t} \Delta_t$$

for the unknown ρ_Δ , where SCC_0 is given by the DSICE solution and Δ_t is the expected extra damage to consumption at time t caused by one extra unit of carbon emission in 2005. We do this for DICE-CJL with $\psi = 1.5$, which has no tipping element, and the benchmark tipping case described above (with $\psi = 1.5$ and $\gamma = 10$), but we use their BAU versions to cancel out the impact of mitigation on damages by following the method of IWG (2010) for computing the discount rate of damages. For comparison, we also compute the internal rate of return on capital investment, which is the rate used to discount the additional consumption caused by one extra unit of capital in 2005.

Table 7 displays ρ_Δ and the internal rate of return on capital investment for both cases. We first see that marginal damages in DICE-CJL BAU are discounted at 3.7 percent and capital investment at 2.7 percent, a difference that is natural given that BAU sets mitigation to zero. These results are also similar to other deterministic models. In contrast, we see that ρ_Δ is only 2.4 percent in the

Table 7: Discount rate of damages and internal rate of return on capital investment

	DSICE BAU with the benchmark tipping setting	DSICE BAU with the benchmark tipping setting but $\Omega_T(T_{\mathrm{AT}}) \equiv 1$	DICE-CJL BAU with $\psi = 1.5$
Discount rate of damages	2.4%	2.4%	3.7%
Internal rate of return on capital investment	2.6%	2.7%	2.7%

tipping benchmark DSICE BAU case. But the internal rates of return on capital investment are close to each other between DICE-CJL BAU and DSICE BAU. This tells us that our higher SCC is not simply implied by higher potential damage from climate tipping events; otherwise the discount rate of damage would be the same as with DICE-CJL. This lower discount rate of damages from carbon emissions indicates that part of the justification of the higher SCC is demand for insurance. To illustrate this justification more clearly, Table 7 also reports ρ_Δ and the internal rate of return on capital investment for the case of DSICE BAU with the benchmark tipping setting but $\Omega_T(T_{\mathrm{AT}}) = 1$ (i.e., damages are only from tipping events). Both rates are, respectively, close to those in the case of DSICE BAU with the benchmark tipping setting and the default damage factor for temperature. This implies that the lower discount rate of damages in DSICE BAU with the benchmark tipping setting is not arising from the potentially nonlinear effect of additional damages from tipping events. The intuition is that the damages from a permanent shock to productivity have little covariance with short-run fluctuations in consumption. Therefore, damages from tipping events should be discounted less than damages from temperature (which are proportional to output).

6.4 Sensitivity Analysis for the Climate Tipping Process

We next examine how the 2005 SCC is affected by parameter uncertainty in the climate tipping process by recomputing SCC over a range of parameter choices that reflect scientific opinions. We examine the six-dimensional collection of parameter values defined by the tensor product of the following finite sets:

$$\lambda \in \{0.0025, 0.0035, 0.0045\}, \qquad \overline{\mathfrak{D}}_\infty \in \{0.025, 0.05, 0.10\}, \qquad q \in \{0, 0.2, 0.4\},$$

$$\psi \in \{0.5, 1.25, 1.5, 1.75, 2.0\}, \qquad \gamma \in \{2, 5, 10, 15\}, \qquad \overline{\Gamma} \in \{5, 50, 200\}.$$

We compute SCC in 2005 for all the cases and Table 8 presents the 2005 SCC for some of the representative cases. For example, when $\lambda = 0.0025$, $\overline{\Gamma} = 5$, $\overline{\mathfrak{D}}_\infty = 0.025$, and $q = 0$ (i.e., the first row in Table 8) the 2005 SCC is \$132/tC for $\psi = 1.5$ and $\gamma = 10$. The value of SCC with the climate tipping process is always greater than in the deterministic case in which the climate tipping process is ignored. This is expected since the tipping element increases possible future damage.

Table 8 also shows that SCC is larger for a higher inter-temporal elasticity of substitution (IES, i.e., ψ), and a higher value of the risk-aversion parameter γ. Furthermore, we observe that the initial-time SCC increases with higher (present-discounted) expected damage from the climate tipping process, which can be caused by a higher hazard rate parameter (λ), a shorter expected duration of the tipping process ($\overline{\Gamma}$), a higher mean long-run damage level ($\overline{\mathfrak{D}}_\infty$), or a higher mean squared-variance ratio of the expected damage level (q). As mentioned earlier, our specification of a climate tipping point in an economic growth model is unique by the standards of how climate scientists view the nature of climate tipping points (e.g., Lenton and Ciscar 2013).

Hazard rate λ	Mean duration of tipping process $\overline{\Gamma}$	Mean long-run damage level $\overline{\mathfrak{D}}_\infty$	Mean squared -variance ratio q	Social cost of carbon (SCC)					
				$\psi = 0.5$		$\psi = 1.5$		$\psi = 2$	
				$\gamma = 2$	$\gamma = 10$	$\gamma = 2$	$\gamma = 10$	$\gamma = 2$	$\gamma = 10$
0.0025	5	0.025	0	61	61	128	132	160	164
			0.4	61	61	129	135	161	169
		0.100	0	74	87	275	349	348	412
			0.4	75	107	285	413	357	458
	200	0.025	0	59	59	111	112	136	138
			0.4	59	59	111	112	136	139
		0.100	0	62	64	174	195	237	276
			0.4	63	65	177	224	241	318
0.0045	5	0.025	0	63	64	148	151	188	192
			0.4	63	64	149	155	189	200
		0.100	0	91	113	365	418	422	461
			0.4	94	144	375	462	429	498
	200	0.025	0	59	59	120	121	149	151
			0.4	59	59	120	122	150	153
		0.100	0	66	69	227	259	318	356
			0.4	67	71	233	304	324	394

Table 8: Initial social cost of carbon (\$/tC) with stochastic climate tipping

Table 8 shows that with climate risk only, a higher IES (i.e., ψ) implies a higher SCC. Given the positive productivity growth in our cases with climate risk only, a higher IES implies a lower desire to smooth per capita consumption, leading to a higher investment on mitigation for higher

future consumptions. This effect occurs because the discount rate of damages from tipping events is less than the internal rate of return on capital investment, as shown in Table 7.

Table 8 also shows that a higher risk-aversion will imply a higher SCC too. Since our tipping probability $p_{\text{tip},t}$ can be reduced by slowing down the temperature increase through mitigation, and the variance of output at time t is proportional to $p_{\text{tip},t}(1-p_{\text{tip},t})$, we see that the variance of output can be diminished by reducing $p_{\text{tip},t}$ (note that $p_{\text{tip},t}$ is always smaller than 0.5 in our cases), and so can the variance of consumption be reduced. Thus, a higher risk-aversion implies a greater value for mitigation, and therefore also a higher SCC for the cases with climate risk only.

Table 8 shows that when $\overline{\mathfrak{D}}_\infty$, the mean long-run damage level, is small, SCC is insensitive to risk-aversion γ, or to the mean squared-variance ratio q (see rows with $\overline{\mathfrak{D}}_\infty = 0.025$), as a small $\overline{\mathfrak{D}}_\infty$ implies a much lower variance of the uncertain, final, long-run damage level, which is equal to $q\overline{\mathfrak{D}}_\infty^2$. For example, when $\lambda = 0.025$, $\overline{\Gamma} = 5$, $\overline{\mathfrak{D}}_\infty = 0.025$, and $\psi = 0.5$, SCC is \$61/tC for all $\gamma \in \{2, 10\}$ and $q \in \{0, 0.4\}$. However, when $\overline{\mathfrak{D}}_\infty$ is large (e.g., $\overline{\mathfrak{D}}_\infty = 0.1$), SCC increases significantly when γ increases from 2 to 10, since a large $\overline{\mathfrak{D}}_\infty$ implies a much larger variance of the uncertain, final, long-run damage level. This reflects the nonlinear effect of climate tipping damage on SCC. In models assuming separable preferences, it is only the mean of the uncertain damage, and not its variance, that affects SCC.

In addition, from the results of all the cases we find that one common pattern exists: the 2005 SCC is linear in q. Figure 7 shows the numbers for SCC for various values of γ and q, when $\psi = 1.5$, $\lambda = 0.0035$, $\overline{\Gamma} = 50$, and $\overline{\mathfrak{D}}_\infty = 0.05$.[27] Other cases have the same qualitative pattern, so we omit them here. In Figure 7, the green line, the blue line, the black line, and the red line represent the cases of $\gamma = 2$, 5, 10, and 15, respectively. We see that all these lines are straight, and that a higher γ implies a greater slope, meaning that it is more sensitive to the variance of the uncertain damage level.

[27] In Figure 7, the horizontal axis is the variance of the uncertain damage level at the final absorbing stage–namely $q\overline{\mathfrak{D}}_\infty^2$, and it is scaled by 10,000.

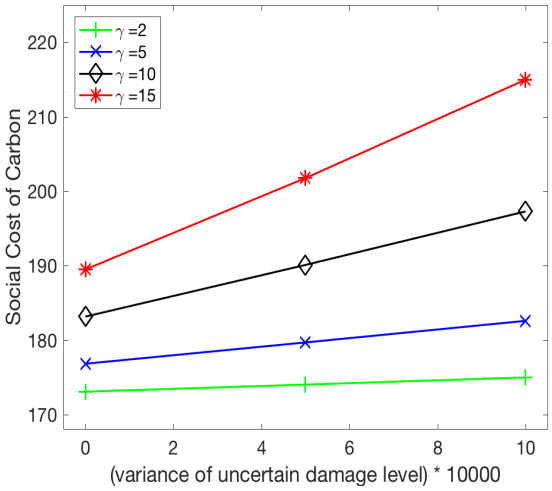

Figure 7: Sensitivity of the social cost of carbon to the risk-aversion parameter and uncertainty regarding post-tipping damage ($\psi = 1.5$, $\lambda = 0.0035$, $\overline{\Gamma} = 50$, and $\overline{\mathfrak{D}}_\infty = 0.05$).

This is not surprising since it fits into the logic of the basic consumption-based capital asset pricing model (Lucas 1978), which tells us that the price of risk is related to its covariance with the aggregate endowment. Since the magnitude of the damage is proportional to output, the damage is strongly related to output; in fact, in this case climate damage is the only stochastic element of output conditional on the tipping event. Therefore, the correlation is unity and SCC has a price of risk component that is linear in the variance of the uncertain damage level.

Note that SCC increases as the variance of the uncertain damage level increases. One interpretation of variance is that it represents our ignorance of the consequences of an unfolding tipping process. With that interpretation, the horizontal difference represents the decline in SCC that would result if we carried out more scientific research and reduced the uncertainty regarding the post-tipping damage level. This observation shows that DSICE could be used to identify the value of reducing uncertainty and to indicate which kinds of scientific studies would be the most valuable to pursue. We leave this point for further research.

6.5 Other Approaches to Damage Functions with Tipping Events

A few studies attempt to integrate stochastic tipping into their deterministic models (e.g., Lempert et al. 2000, Nordhaus 2009, Ackerman et al. 2010 and Ackerman and Stanton 2012). They do this by altering the damage function in an ad hoc manner, such as increasing the convexity of

the damage function. Similarly, Mastrandrea and Schneider (2001) increase the exponent of the DICE-type damage function with the purpose of quantifying the effects of an abrupt non-linear climatic change: the shutdown of the thermohaline circulation. The reasoning for increasing the exponent is to deal with abrupt non-linear changes by creating a more non-linear damage function. Lontzek et al. (2015) discusses this approach, showing with the DSICE model that this approach fails to capture the implications of stochastic tipping points. It is furthermore, incompatible with the real-world task of decision making under uncertainty. Therefore, DSICE shows that there is no need for such ad hoc approximations.

Several IAMs have carried out analyses of a climate catastrophe event that directly affects utility. Early studies include Gjerde et al. (1999) and Castelnuovo et al. (2003). These studies assume that an uncertain level of global temperature triggers an irreversible drop in utility and find that large reductions in emissions are optimal. Nevertheless, these studies apply a certainty-equivalent approach to uncertainty and do not address stochasticity. Most recently, Cai et al. (2015) use the DSICE framework to include the possibility of a catastrophic event with non-market impacts, such as impacts on the ecosystem. Most cost−benefit analyses rarely take account of environmental tipping points leading to abrupt and irreversible impacts on market and nonmarket goods and services, including those provided by the climate and by ecosystems. Cai et al. (2015) shows that including environmental tipping point impacts in a stochastic dynamic integrated assessment model profoundly alters cost−benefit assessment of global climate policy. The risk of a tipping point, even if it only has nonmarket impacts, could substantially increase the present optimal carbon tax.

7 The Social Cost of Carbon with Stochastic Growth and Climate Tipping

The previous two sections have examined the impacts on SCC from stochastic growth and stochastic climate tipping, both in isolation. The real world system includes both uncertainties and this section presents the results of DSICE in the presence of long-run risk in both economic growth and in the climate tipping process. The optimal policy will now have to balance the need to delay the triggering of the tipping point process with the accumulation of additional capital in the face of stochastic growth, and with the desire to smooth our consumption patterns. We study a *stochastic growth and*

climate tipping benchmark case of parameter specification and carry out a sensitivity analysis.

7.1 The Stochastic Growth and Climate Tipping Benchmark

We use $\psi = 1.5$, $\gamma = 10$, $\lambda = 0.0035$, $\overline{\mathfrak{D}}_\infty = 0.05$, $q = 0.2$, and $\overline{\Gamma} = 50$ for the stochastic growth and climate tipping benchmark. Figure 8 shows the results of 10,000 simulation paths over the first 100 years with regard to the dynamics of SCC, the carbon tax, and the ratio of SCC to gross world output, respectively. Other variables such as capital, consumption and its growth, atmospheric carbon concentration, and surface temperature have pictures visually similar to the corresponding pictures in Figures 2 and 6, so we omit them. We use the same line and color types as in previous figures.

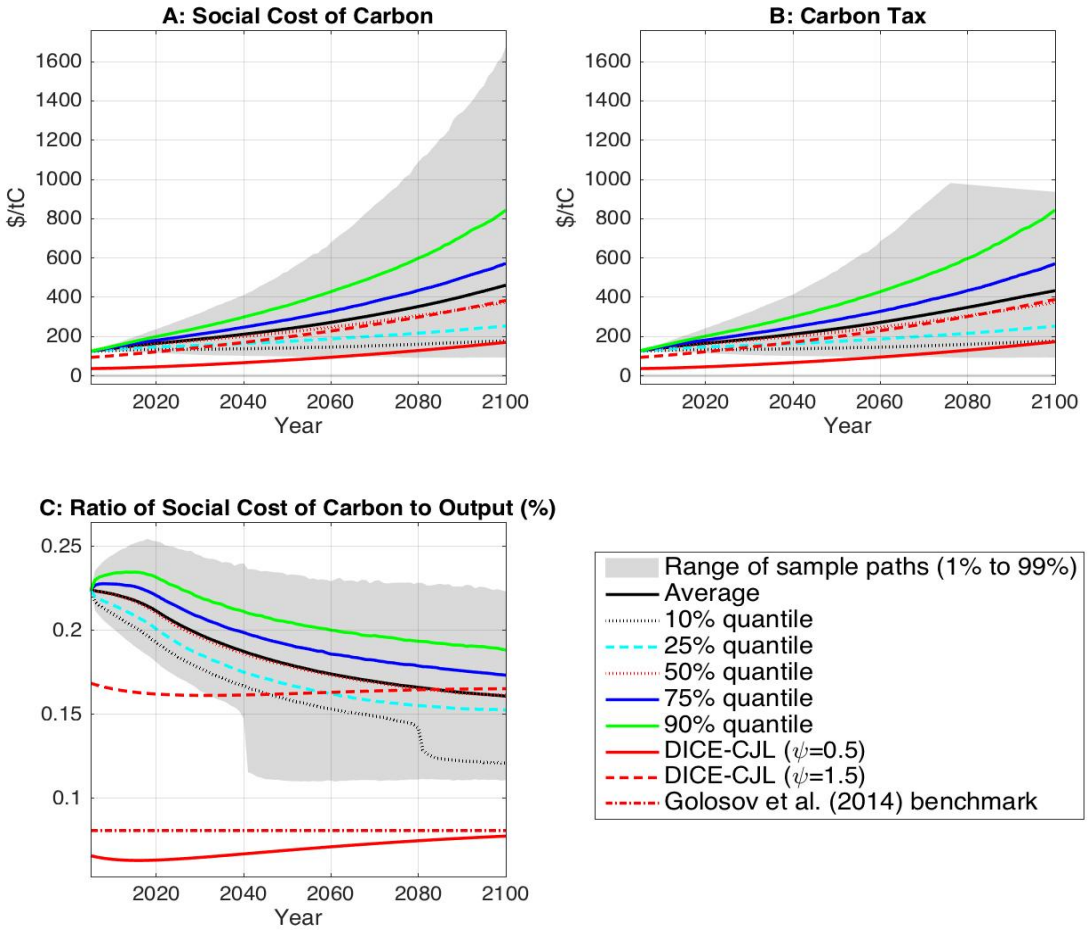

Figure 8: Simulation results for the stochastic growth and climate tipping benchmark

We first study SCC. Its initial-time level is \$124/tC and at 2100 the average (or expected) SCC is \$461/tC. Thus, the path of the expected SCC is situated between its paths obtained from our analyses of each risk component in isolation. In 2005, the SCC of \$124/tC is even exactly the average of the numbers obtained from the two cases from the previous sections (\$61/tC and \$188/tC). Compared to a deterministic model, which ignores both risk components and has $\psi = 0.5$, we find that the 2005 SCC increases by a factor of more than three, and that with 90 percent probability SCC will be significantly higher throughout this century.

The presence of both stochastic growth and climate tipping risk also increases the variance of the future SCC relative to the separate stochastic growth and climate tipping benchmarks. For example, SCC in 2100 ranges from \$100/tC (the 1% quantile) to \$1,700/tC (the 99% quantile). The carbon tax, which we present in panel B, is also more likely to hit its upper bound after 2072 than in either of the single-risk benchmarks. The combination of these risks implies that there is a probability of about 7.5 percent that mitigation policies will have reached the limit of their effectiveness by 2100.

We also revisit the analysis of the SCC_t/Y_t ratio of Section 5.1, here in the case of stochastic growth and climate tipping. Panel C in Figure 8 shows that the SCC_t/Y_t ratio is quite stochastic. First, we note that compared to the deterministic version of the model, SCC_t/Y_t is about three times larger in 2005 while at 2100 it is expected to be about twice as large. Second, we find that the expected SCC_t/Y_t is decreasing over time by about 50 percent, and is thus not constant. Third, and most importantly, we find that the ratio of SCC to gross world output is not close to any simple path, but is rather a stochastic process varying over the interval [0.00108, 0.00215] at year 2100 and over an even larger interval for the second half of this century. When contrasting our results with the constant SCC_t/Y_t ratio postulated by Golosov et al. (2014), we find that the inclusion of a long-run risk and a tipping process with Epstein–Zin preference makes SCC_t/Y_t significantly uncertain, and leads to substantially different qualitative and quantitative results.

Table 9 shows SCC in 2005, mean SCC in 2100, and the 90 percent quantile SCC in 2100 for our three benchmark examples. We see that the 2005 SCC and the mean SCC in 2100 of the stochastic growth and tipping benchmark are, respectively, in between those of the stochastic growth benchmark and the climate tipping benchmark, but that the 90 percent quantile SCC in 2100 of the stochastic growth and tipping benchmark is larger than the corresponding ones of both

	DICE-CJL ($\psi = 0.5$)	DICE-CJL ($\psi = 1.5$)	DSICE ($\psi = 1.5$, $\gamma = 10$)		
			Stochastic Growth	Climate Tipping	Stochastic Growth with Tipping
2005 SCC	37	94	61	188	124
Mean SCC in 2100	180	389	357	620	461
S.d. of SCC in 2100	0	0	247	105	316
90% quantile SCC in 2100	180	389	667	662	844

Table 9: SCC ($/tC) for three benchmark examples

the stochastic growth benchmark and the climate tipping benchmark.

Furthermore, when comparing our results to DICE-CJL with $\psi = 0.5$, we find that the interaction between multiple uncertainties, such as long-run risk and a tipping process, could be nontrivial: for example, while either long-run risk or a tipping process leads to a higher 2005 SCC than does the deterministic DICE-CJL (with $\psi = 0.5$), their combination does not imply a further increase in the 2005 SCC when compared to cases with only one type of uncertainty. Thus, it is necessary to carry out a systematic analysis rather than a stylized one. For the case of DICE-CJL with $\psi = 1.5$ we have already argued in Section 5.2 that, as there is no risk, the positive future growth in productivity and a higher ψ (here: 1.5) leads to larger incentives to avoid climate damage increases. The SCC will be larger in that case when compared to the case of stochastic growth only. Nevertheless, the combination of stochastic growth and climate tipping will produce SCC substantially larger than in the deterministic case.

7.2 Sensitivity of the Stochastic Growth and Climate Tipping Benchmark

We compute the sensitivity of the stochastic growth and climate tipping benchmark case to several parameters. Table 10 lists the 2005 SCC for selected combinations of the parameter values for the sensitivity analysis. These parameters are the hazard rate λ, the mean duration time of the tipping process $\overline{\mathcal{D}}$, the mean long-run damage level $\overline{\mathfrak{D}}_\infty$, the mean squared-variance ratio q, the elasticity of inter-temporal substitution ψ, and the risk-aversion parameter γ.

Hazard rate λ	Mean duration of tipping process $\overline{\Gamma}$	Mean long-run damage level $\overline{\mathfrak{D}}_\infty$	Mean squared -variance ratio q	Social cost of carbon (SCC)					
				$\psi = 0.5$		$\psi = 1.5$		$\psi = 2$	
				$\gamma = 2$	$\gamma = 10$	$\gamma = 2$	$\gamma = 10$	$\gamma = 2$	$\gamma = 10$
0.0025	200	0.025	0	43	70	102	76	120	78
0.0035	5	0.05	0	63	133	177	144	217	149
		0.10	0.4	93	337	313	373	400	384
	50	0.05	0	55	111	156	124	192	128
			0.2	55	114	157	124	193	132
		0.10	0	73	187	247	212	317	220
0.0045	5	0.10	0	102	294	356	314	455	324
			0.2	104	339	365	372	466	383
			0.4	106	394	374	438	477	446
	50	0.05	0	59	122	171	136	212	141

Table 10: Initial social cost of carbon ($/tC) under stochastic growth and climate tipping

We find that some qualitative properties found in previous examples with only climate risk still hold—that is, a higher hazard rate parameter λ, a shorter expected duration $\overline{\Gamma}$, a higher mean damage level $\overline{\mathfrak{D}}_\infty$, or a larger mean squared-variance ratio q will lead to a higher SCC, although their quantitative values differ substantially. For example, by comparing the cases of $\lambda = 0.0035$ and $\lambda = 0.0045$ from Table 10 with the other values, $\overline{\mathfrak{D}}_\infty = 10\%$, $\overline{\Gamma} = 50$, and $q = 0$, we see that the case with $\lambda = 0.0045$ has a larger SCC than the case with $\lambda = 0.0035$. Moreover, the range of the initial SCC is wide, from \$43/tC (the case with $\lambda = 0.0025$, $\overline{\Gamma} = 200$, $\overline{\mathfrak{D}}_\infty = 0.025$, $q = 0$, $\psi = 0.5$, and $\gamma = 2$) to \$477/tC (the case with $\lambda = 0.0045$, $\overline{\Gamma} = 5$, $\overline{\mathfrak{D}}_\infty = 0.10$, $q = 0.4$, $\psi = 2$, and $\gamma = 2$).

Furthermore, by comparing the columns with $\gamma = 2$, or comparing the columns with $\gamma = 10$, we see that a higher IES always implies a higher SCC for the cases with $\gamma \leq 10$. This is consistent with our findings in earlier examples with stochastic growth only or climate risk only. However, the qualitative properties of the risk-aversion parameter γ are now nontrivial: Table 10 shows that the effect of a higher γ on SCC can be positive or negative. This non-triviality comes from the combination of economic risk and climate tipping risk: with economic risk only, if ψ is small (i.e., $\psi < 0.7$), then a higher risk-aversion implies a higher SCC; but with climate tipping risk only, a higher risk-aversion always implies a higher SCC for any IES. For the cases with $\psi = 0.5$ in Table 10, a higher risk-aversion implies a higher SCC, but in most cases in Table 10 with $\psi = 1.5$ and all cases with $\psi = 2$, a higher γ will result in a smaller SCC, implying that for lower levels of ψ the effect of higher γ is more likely to be positive. This complicated pattern originates from the

interaction between the nontrivial pattern of the impact of ψ and γ on SCC in the stochastic growth cases and the pattern in the climate tipping cases.

8 Summary and Conclusion

This study has presented DSICE, a computational framework for the stochastic integrated assessment of issues related to the joint evolution of the economy and the climate. We analyzed the optimal level and dynamic properties of SCC and the associated optimal carbon tax in the face of stochastic and irreversible climate change and its interaction with economic factors including growth uncertainty and preferences regarding risk. We did this in a manner that allows us to compare our results to the deterministic model in Nordhaus (2008), an influential and well-known integrated assessment model. The specific examples in this study show three basic points.

First, we use Epstein–Zin preferences in order to specify tastes that are more compatible with the evidence on risk-aversion and the inter-temporal elasticity of substitution. This also allows us to separate the impact of risk-aversion from the impact of inter-temporal substitution. The impact of risk-aversion depended on the nature of the risk, with SCC rising as risk-aversion increases for the tipping case, but with more ambiguous implications for the case of stochastic growth.

Second, the incorporation of long-run risk shows that SCC is itself a stochastic process with considerable uncertainty. Climate change issues are not just about the expected state of the climate in the future, but also about avoiding disasters. Climate change policy has to recognize the uncertainty about future SCC, and be prepared to consider policies, such as geoengineering and carbon capture, that are currently considered to have costs that would never be justified by deterministic models which essentially focus on the expected SCC. An examination of parameter uncertainty also shows that the range of plausible SCC values is much larger than implied by other integrated assessment analyses. We found this great uncertainty for single parameterizations; when we also include uncertainty regarding economic parameters, that uncertainty is even greater.

Third, climate scientists have recently argued that tipping elements in the climate system contribute to uncertainty regarding future climate conditions. We incorporate the concept of tipping elements into DSICE, and find that the threat of a tipping element leads to significant and immediate increases in SCC, where the increase in SCC is disproportionately large relative to the damage

caused by tipping. This is true even for moderate assumptions regarding the likelihood and impacts of climate tipping events. SCC can be very high, even without assuming catastrophic climate change events, but rather by merely assuming plausibly parameterized examples of uncertain and irreversible climate change. An internal rate of return analysis showed that one should use a smaller discount rate when valuing damage from tipping events, a natural result since the damage caused by tipping is less correlated to total consumption than is damage to output.

Finally, we have also shown that it is possible to solve empirically plausible nine-dimensional models of the climate and the economy that include (i) productivity shocks of the kind studied in macroeconomics, (ii) dynamically non-separable preferences consistent with observed prices of risk, and (iii) stochastic tipping elements in the climate system. Our examination of examples including both the usual Nordhaus-style damage function for output and damages arising from tipping events show that these separate sources of damages must be examined together. The usual approach is to study them separately or to create ad hoc approximations. This study shows that analyses that avoid such shortcuts are both feasible and necessary if we are to obtain reliable answers. This study has ignored many features that might be important, but this weakness is unavoidable when modeling a complex system, particularly one which is a merger of two enormously complex systems. This study has shown that it is possible to look at models of much greater complexity than before, and, in the case of SCC, incorporating that complexity significantly alters the implications for economic policy. The fact that we solve a nine-dimensional dynamic programming problem with long-run risk over a span of 600 years implies that it is straightforward to use those nine dimensions to model other topics, such as geoengineering, robust optimization, and learning.

References

[1] Ackerman, F., Stanton, E.A. (2012). Climate risks and carbon prices: Revising the social cost of carbon. Economics: The Open-Access, Open-Assessment E-Journal, 6 (2012-10): 1–25.

[2] Ackerman, F., Stanton, E. A. & Bueno, R. (2010). Fat tails, exponents, extreme uncertainty: Simulating catastrophe in DICE. Ecological Economics, 69, 16571665.

[3] Ackerman, F., E.A. Stanton, and R. Bueno (2013). Epstein–Zin utility in DICE: is risk-aversion irrelevant to climate policy? Environ Resource Econ 56, 73–84.

[4] Anderson, E. Brock, W., Hansen, L.P. and Sanstad A.H (2014). Robust Analytical and Computational Explorations of Coupled Economic-Climate Models with Carbon-Climate Response. RDCEP Working paper No. 13-05

[5] Anthoff, D., and Tol, R.S.J. (2014), The Climate Framework for Uncertainty, Negotiation and Distribution (FUND), Technical Description, Version 3.9

[6] Azar and Sterner (1996). Discounting and distributional considerations in the context of global warming, Ecological Economics, 1996

[7] Babiker, M., Gurgel, A., Paltsev, S. and Reilly, J. (2009) Forward-Looking Versus Recursive-Dynamic Modeling in Climate Policy Analysis: A Comparison. Economic Modelling, 26(6): 1341-1354

[8] Bansal, R., D. Kiku, and A. Yaron (2012). An empirical evaluation of the long-run risks model for asset prices. Critical Finance Review, Issue 1, 183–221.

[9] Bansal, R., and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. The Journal of Finance, 59(4), 1481–509.

[10] Bansal, R., and M. Ochoa (2011). Welfare costs of long-run temperature shifts, Duke University Working Paper.

[11] Barrage, L. (2014). Sensitivity analysis for Golosov, Hassler, Krusell, and Tsyvinski (2014): 'Optimal taxes on fossil fuel in general equilibrium' supplemental material, Econometrica, 82(1): 41-88.

[12] Barro, R.J. (2009). Rare disasters, asset prices, and welfare costs. American Economic Review, 99(1): 243–264.

[13] Beeler, J., and J.Y. Campbell (2011). The long-run risks model and aggregate asset prices: an empirical assessment. Critical Finance Review 1(1): 141–182.

[14] Bellman, R. (1957). Dynamic Programming. Princeton University Press.

[15] Brock, W.A., Durlauf, S.D., West, K.D., (2007). Model uncertainty and policy evaluation: Some theory and empirics. Journal of Econometrics 136 (2), 629–664.

[16] Cai, Y. (2010). Dynamic Programming and Its Application in Economics and Finance. PhD thesis, Stanford University.

[17] Cai, Y., and K.L. Judd (2010). Stable and efficient computational methods for dynamic programming. Journal of the European Economic Association, 8(2-3): 626–634.

[18] Cai, Y., K.L. Judd, and T.S. Lontzek (2012a). Open science is necessary. Nature Climate Change, 2(5): 299.

[19] Cai, Y., Judd, K.L, Leton, T.M., Lontzek, T.S., Narita, D. (2015). Risk to ecosystem services could significantly affect the cost-benefit assessments of climate change policies. The Proceedings of the National Academy of Sciences (PNAS)

[20] Cai, Y., K.L. Judd, and T.S. Lontzek (2012b). Continuous-time methods for integrated assessment models. NBER working paper 18365.

[21] Cai, Y., K.L. Judd, G. Thain, and S. Wright (2015). Solving dynamic programming problems on computational grid. Computational Economics, 45(2): 261–284.

[22] Cai, Y., T.M. Lenton, and T.S. Lontzek (2016). Risk of multiple interacting tipping points should encourage rapid CO2 emission reduction. Nature Climate Change 6, 530–525.

[23] Castelnuovo, E., Moretto, M. & Vergalli, S. (2003). Global warming, uncertainty and endogenous technical change. Environ. Model. Assess. 8, 291301.

[24] Cogley, T., De Paoli, B., Matthes, C., Nikolov, K., & Yates, T. (2011). A Bayesian approach to optimal monetary policy with parameter and model uncertainty. Journal of Economic Dynamics and Control, 35(12), 2186-2212.

[25] Constantinides, G.M., and A. Ghosh (2011). Asset pricing tests with long-run risks in consumption growth. The Review of Asset Pricing Studies, 1(1), 96–136.

[26] Crost, B., and Traeger, C. (2014). Optimal CO2 mitigation under damage risk valuation. Nature Climate Change, June 15, 2014.

[27] Dell, M., Jones, B.F., and Olken, B.A. (2014). What Do We Learn from the Weather? The New Climate-Economy Literature. Journal of Economic Literature, Sept. 2014, 52(3), 740—798

[28] Deschenes, O. and Greenstone, M. (2011) Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the U.S. American Economic Journal: Applied Economics, 3(4): 152-185.

[29] Dietz, S., and N. Stern (2015). Endogenous growth, convexity of damages and climate risk: how Nordhaus' framework supports deep cuts in carbon emissions. The Economic Journal, 125, 574–620.

[30] Epstein, L.G., and S.E. Zin (1989). Substitution, risk-aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica, 57(4), 937–969.

[31] Epstein, L.G., E. Farhi, and T. Strzalecki (2014). How Much Would You Pay to Resolve Long-Run Risk? American Economic Review 2014, 104(9): 2680–2697.

[32] Gjerde, J., Grepperud, S. & Kverndokk, S. (1999). Optimal climate policy under the possibility of a catastrophe. Resour. Energy Econ. 21, 289317.

[33] Golosov, M., J. Hassler, P. Krusell, and A. Tsyvinski (2014). Optimal taxes on fossil fuel in general equilibrium. Econometrica, 82(1): 41–88.

[34] Hansen, L. P., J. Heaton, and N. Li (2008). Consumption Strikes Back? Measuring Long Run Risk. Journal of Political Economy, 116, 260–302.

[35] Heutel G, Moreno-Cruz J, Shayegh S. (2016) Climate tipping points and solar geoengineering. J. Econ. Behave. Organ. 132:19–45.

[36] Hope, C. (2011). The PAGE09 integrated assessment model: a technical description. Cambridge Judge Business School Working Paper 4/2011. Accessed June 23, 2014: `http://www.jbs.cam.ac.uk/fileadmin/user_upload/research/workingpapers/wp1104.pdf`

[37] Hwang IC, Reynes F, Tol RSJ. (2017). The effect of learning on climate policy under fat-tailed uncertainty. Resour. Energy Econ. 48:1–18

[38] IWG (2010). Social Cost of Carbon for Regulatory Impact Analysis under Executive Order 12866. United States Government. http://www.whitehouse.gov/sites/default/files/omb/inforeg/for-agencies/Social-Cost-of-Carbon-for-RIA.pdf

[39] IPCC (2007). Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK.

[40] IPCC (2013). Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[41] IPCC (2014). Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[42] Jensen, S., and Traeger, C. (2014). Optimal climate change mitigation under long-term growth uncertainty: stochastic integrated assessment and analytic findings. European Economic Review 69, 104–125.

[43] Joughin, I., B.E. Smith, and B. Medley (2014). Marine ice sheet collapse potentially under way for the Thwaites Glacier Basin, West Antarctica. Science 344, 735–738.

[44] Judd, K.L. (1998). Numerical Methods in Economics. The MIT Press.

[45] Keller, K., B. Bolker, and D.F. Bradford (2004). Uncertain climate thresholds and optimal economic growth. Journal of Environmental Economics and Management 48, 723–741.

[46] Kelly DL, Kolstad CD. (1999). Bayesian learning, growth, and pollution. J. Econ. Dyn. Control 23:491–518.

[47] Kelly DL, Tan Z. (2015). Learning and climate feedbacks: optimal climate insurance and fat tails. J. Environ. Econ. Manag. 72:98–122.

[48] Kopits, E., A. Marten, and A. Wolverton (2014). Incorporating 'catastrophic' climate change into policy analysis. Climate Policy 14(5): 637–664.

[49] Kriegler, E., J.W. Hall, H. Held, R. Dawson, and H.J. Schellnhuber (2009). Imprecise probability assessment of tipping points in the climate system. PNAS 106(13): 5041–5046.

[50] Lemoine, D.M., and C. Traeger (2014). Watch Your Step: Optimal Policy in a Tipping Climate. American Economic Journal: Economic Policy, 6(1), 137–166.

[51] Lempert, R.J, M. E. Schlesinger, S. C. Bankes and N. G Andronova (2000).The Impacts of Climate Variability on Near-Term Policy Choices and the Value of Information, 45(1), 129-161.

[52] Lenton, T.M. (2010). Earth System Tipping Points. Working Paper. http://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0564-112.pdf/$file/EE-0564-112.pdf

[53] Lenton, T.M., and J-C. Ciscar (2013). Integrating tipping points into climate impact assessments. Climatic Change 117, 585–597.

[54] Lenton, T.M., H. Held, E. Kriegler, J. Hall, W. Lucht, S. Rahmstorf, and H.J. Schellnhuber (2008). Tipping elements in the Earth's climate system. PNAS 105, 1786–1793.

[55] Ljungqvist, L. and Sargent, T. J. 2004. Recursive Macroeconomic Theory, 2nd Edition, MIT Press Books, The MIT Press.

[56] Lucas, R. E. (1978). Asset pricing in an exchange economy. Econometrica 46, 1429–1445.

[57] Manne, A., and R. Richels (2005). MERGE: an integrated assessment model for global climate change. Energy and Environment (175–189) edited by. Loulou, R., Waaub, J-P. and Zaccour, G.

[58] Mastrandrea, M. D. & Schneider, S. H. (2001) Integrated assessment of abrupt climatic changes. Clim. Policy 1, 433449.

[59] Meinshausen, M., S.C.B. Raper, and T.M.L. Wigley (2011). Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6: Part I – Model Description and Calibration. Atmospheric Chemistry and Physics 11, 1417–1456.

[60] Nordhaus, W.D. (1992). An Optimal Transition Path for Controlling Greenhouse Gases. Science 258, 1315–1319.

[61] Nordhaus, W.D. (2007). A review of the Stern Review on the economics of climate change. Journal of Economic Literature 45, 686–702.

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

[63] Nordhaus,W. D. (2009). An Analysis of the Dismal Theorem. Cowles Foundation for Research in Economics, Yale University.

[64] Nordhaus, W.D., and Z. Yang (1996). A regional dynamic general-equilibrium model of alternative climate-change strategies. American Economic Review, 86, 741–765.

[65] National Research Council (2013). Abrupt Impacts of Climate Change: Anticipating Surprises. Washington, DC: The National Academies Press, 2013.

[66] Oberkampf, W.L., and C.J. Roy (2010). Verification and Validation in Scientific Computing. Cambridge University Press.

[67] Pindyck, R.S (2011). Fat Tails, Thin Tails, and Climate Change Policy. Review of Environmental Economics and Policy, Volume 5, Issue 2, Pages 258–274.

[68] Pindyck, R.S. (2013). Climate change policy: what do the models tell us? Journal of Economic Literature, 51(3): 860–872.

[69] Pindyck, R.S., and N. Wang (2013). The economic and policy consequences of catastrophes. American Economic Journal: Economic Policy, 5(4): 306–339.

[70] Robock, A., A. Marquardt, B. Kravitz, and G. Stenchikov (2009), Benefits, risks, and costs of stratospheric geoengineering, Geophys. Res. Lett., 36, L19703.

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

[72] Roe, G.H., and Baker, M.B. (2007). Why is Climate Sensitivity So Unpredictable? Science, 318, 629—632.

[73] Rust, J. (2008). Dynamic Programming. In: Durlauf, S.N., Blume L.E. (Eds.), New Palgrave Dictionary of Economics. Palgrave Macmillan, second edition.

[74] Scheffer, M., Carpenter, S., Foley, J.A., Folke, C. and Walker, B. (2001). Catastrophic shifts in ecosystems. Nature 413, 591-596.

[75] Schneider, S.H. (1989). Global warming: scientific reality of political hype. In: Global Warming: Hearings Before the Subcommittee on Energy and Power of the Committee on Energy and Commerce House of Representatives One Hundred First Congress, Serial No. 101-31, pp. 53–66.

[76] Schneider, S.H., and S.L. Thompson (1981). Atmospheric CO_2 and Climate: Importance of the Transient Response. Journal of Geophysical Research, 86(C4), 3135–3147.

[77] Schorfheide, F., D. Song, and A. Yaron (2014). Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach. NBER Working Paper 20303.

[78] Schultz, G. P. (2015). A Reagan approach to climate change. The Washington Post, (Opinions) March 13, 2015.

[79] Smith, J.B., S.H. Schneider, M. Oppenheimer, G.W. Yohe, W. Hare, M.D. Mastrandrea, A. Patwardhan, I. Burton, J. Corfee-Morlot, C.H.D. Magadza, H.M. Füssel, A.B. Pittock, A. Rahman, A. Suarez, and J.P. Van Ypersele. (2009). Assessing dangerous climate change through an update of the Intergovernmental Panel on Climate Change (IPCC) 'Reasons for concern'. PNAS 106(11), 4133–4137.

[80] Steffen, W, K. Richardson, J. Rockstrom, S. E. Cornell, I. Fetzer, E. M. Bennett, R. Biggs, S. R. Carpenter, W. de Vries, C. A. de Wit, C. Folke, D. Gerten, J. Heinke, G. M. Mace, L. M. Persson, V. Ramanathan, B. Reyers, S. Sorlin (2015). Planetary boundaries: Guiding human development on a changing planet, Science, Volume 347.

[81] Stern, N. H. (2007). The economics of climate change: the Stern Review. Cambridge, UK: Cambridge University Press.

[82] Stern, N.H. and Taylor, C. (2007). Climate Change: Risk, Ethics and the Stern Review. Science, Volume 31.7

[83] Stokey, N. L., and Lucas, R. E., Jr. with Prescott, E.C. (1989). Recursive Methods in Economic Dynamics. Cambridge, Mass.: Harvard University Press.

[84] Tauchen, G. (1986). Finite state Markov-chain approximations to univariate and vector autoregressions, Economic Letters, 20, 177–181.

[85] Vissing-Jørgensen, A., and O.P. Attanasio (2003). Stock-market participation, inter-temporal substitution, and risk-aversion. The American Economic Review 93(2): 383–391.

[86] Weitzman, M.L. (2009). On modeling and interpreting the economics of catastrophic climate change. Review of Economics and Statistics, 91, 1–19.

[87] Wuebbles, D.J. (2016). Setting the stage for risk management: severe weather under a changing climate. Risk Analysis of Natural Hazards, Springer International Publishing, 61–80.

[88] Zickfeld K., A. Levermann, M.G. Morgan, T. Kuhlbrodt, S. Rahmstorf, and D.W. Keith (2007). Expert judgements on the response of the Atlantic meridional overturning circulation to climate change. Climate Change 82, 235–265.

Appendices

A.1 Definition of Parameters and Variables

In Equations (10), (11), (1), and (6), we introduce L_t (world population), σ_t (carbon intensity of output), $\theta_{1,t}$ (mitigation cost coefficient), $E_{\text{Land},t}$ (annual carbon emissions from biological processes), and $F_{\text{EX},t}$ (exogenous radiative forcing). In the following, we specify these variables more precisely. Here, we use annual time steps, which correspond to the decadal time steps in Nordhaus (2008):

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

$$\sigma_t = \sigma_0 \exp\left(-0.0073(1 - e^{-0.003t})/0.003\right) \tag{A.2}$$

$$\theta_{1,t} = \frac{1.17\sigma_t\left(1 + e^{-0.005t}\right)}{2\theta_2} \tag{A.3}$$

$$E_{\text{Land},t} = 1.1e^{-0.01t} \tag{A.4}$$

$$F_{\text{EX},t} = \begin{cases} -0.06 + 0.0036t, & \text{if } t \leq 100 \\ 0.3, & \text{otherwise.} \end{cases} \tag{A.5}$$

Tables A.1–A.6 list the values and/or definitions of all critical parameters, variables, and symbols used in the main text. We use Nordhaus (2008) for economic parameters common to DSICE and DICE. Table A.1 lists the critical parameters and variables in the economic system in the absence of climate impact and mitigation.

$t \in \{0, 1, ..., 600\}$	time in years (t represents year $t + 2005$)
$\psi \in [0.5, 2]$	inter-temporal elasticity of substitution (default: 1.5)
$\gamma \in [2, 15]$	risk-aversion parameter (default: 10)
$\beta = 0.985$	utility discount factor
$\alpha = 0.3$	capital share
α_1	2005 growth rate of the productivity trend (default: 0.0092)
α_2	rate of decline in growth rate of the productivity (default: 0.001)
$\delta = 0.1$	annual capital depreciation rate
A_t	productivity trend at time t, $A_0 = 0.0272$
L_t	population at time t
K_t	capital at time t (in \$ trillions), $K_0 = 137$
C_t	consumption at time t
I_t	investment at time t
Y_t	gross world output at time t

Table A.1: Parameters, variables, and symbols for the economic system of DSICE

Table A.2 lists the critical parameters and variables about climate impact, emission and mitigation in the economic system.

Ω_T	damage factor function from temperature increase at time t
$\pi_1 = 0$	climate damage factor parameter
$\pi_2 = 0.0028388$	climate damage factor parameter
$\theta_2 = 2.8$	mitigation cost parameter
$\sigma_0 = 0.13418$	carbon intensity in 2005
σ_t	carbon intensity at time t
$\theta_{1,t}$	adjusted cost for backstop at time t
Ψ_t	mitigation expenditure at time t
μ_t	emission control rate at time t
E_t	carbon emissions (billions tons) at time t
$E_{\text{Ind},t}$	industrial carbon emissions (billions tons) in year t
$E_{\text{Land},t}$	carbon emissions (billions tons) from land use in year t
SCC_t	social cost of carbon at time t

Table A.2: Parameters, variables, and symbols for the economic system of DSICE

Table A.3 lists the critical parameters and variables for the long-run risk in productivity.

$\widetilde{A}_t = \zeta_t A_t$	stochastic productivity at time t
ζ_t	stochastic productivity shock at time t, $\zeta_0 = 1$
χ_t	persistence of productivity shock at time t, $\chi_0 = 0$
$\varrho = 0.035$	stochastic productivity process parameter
$r = 0.775$	stochastic productivity process parameter
$\varsigma = 0.008$	stochastic productivity process parameter
g_ζ	transition function of ζ_t
g_χ	transition function of χ_t
$\omega_{\zeta,t}$	i.i.d. shocks in transition of ζ_t
$\omega_{\chi,t}$	i.i.d. shocks in transition of χ_t

Table A.3: Parameters, variables, and symbols for the stochastic growth specification

Table A.4 lists the critical parameters and variables for the carbon cycle.

$M_{\mathrm{AT},t}$	carbon concentration in atmosphere (billion tons) at time t, $M_{\mathrm{AT},0} = 808.9$
$M_{\mathrm{UO},t}$	carbon concentration in upper ocean (billion tons) at time t, $M_{\mathrm{UO},0} = 1255$
$M_{\mathrm{LO},t}$	carbon concentration in lower ocean (billions tons) at time t, $M_{\mathrm{LO},0} = 18365$
$\mathbf{M}_t = (M_{\mathrm{AT},t}, M_{\mathrm{UO},t}, M_{\mathrm{LO},t})^\top$	carbon concentration vector at time t
Φ_{M}	transition matrix of carbon cycle
$\phi_{12} = 0.019$	rate of carbon diffusion from atmosphere to upper ocean
$\phi_{23} = 0.0054$	rate of carbon diffusion from upper ocean to lower ocean
$\phi_{21} = 0.01$	rate of carbon diffusion from upper ocean to atmosphere
$\phi_{32} = 0.00034$	rate of carbon diffusion from lower ocean to upper ocean

Table A.4: Parameters and variables in the carbon system

Table A.5 lists the critical parameters and variables for the temperature system.

$T_{\mathrm{AT},t}$	global average surface temperature increase above the year 1900 temperature level (in °C), $T_{\mathrm{AT},0} = 0.7307$
$T_{\mathrm{OC},t}$	global average ocean temperature increase above the year 1900 temperature level (in °C), $T_{\mathrm{OC},0} = 0.0068$
$\mathbf{T}_t = (T_{\mathrm{AT},t}, T_{\mathrm{OC},t})^\top$	temperature vector at time t
Φ_{T}	transition matrix of temperature system
$\xi_1 = 0.037$	temperature transition parameter
$\xi_2 = 0.047$	rate of $T_{\mathrm{AT},t}$ decrease due to infrared radiation to space
$\varphi_{12} = 0.010$	rate of heat diffusion from atmosphere to ocean
$\varphi_{21} = 0.0048$	rate of heat diffusion from ocean to atmosphere
$\mathcal{F}_t$	radiative forcing at time t
$F_{\mathrm{EX},t}$	exogenous radiative forcing in year t
$\eta = 3.8$	radiative forcing parameter
$M^*_{\mathrm{AT}} = 596.4$	preindustrial atmospheric carbon concentration

Table A.5: Parameters and variables in the temperature system

Table A.6 lists the critical variables for the tipping element part of the climate system:

J_t	tipping state at time t
$\mathcal{J}_0$	initial state of J_t (i.e., pre-tipping state)
$\mathcal{M}_i$	post-tipping Markov chain $i \in \{1, 2, 3\}$
$\mathcal{J}_{i,j}$	possible values of the tipping state variable in $\mathcal{M}_i$ for $j \in \{1, 2, 3, 4, 5\}$
J_∞	uncertain long-run tipping state given time $t = 0$ information
$\mathfrak{D}_{i,j} = D(\mathcal{J}_{i,j})$	damage level at the tipping state $\mathcal{J}_{i,j}$
$\mathfrak{D}_\infty = D(J_\infty)$	uncertain long-run tipping damage level given time $t = 0$ information
$p_{\text{tip},t}$	tipping probability (i.e., probability that the tipping state leaves its initial state, $\mathcal{J}_0$) at time t
p	transition probability of transient states in the post-tipping process
$\omega_{J,t}$	i.i.d. shock in transition of J_t
$g_J(\omega_{J,t})$	transition function of J_t
$\underline{T_{\text{AT}}} = 1$	surface temperature parameter used in tipping probability function
$\lambda \in [0.0025, 0.0045]$	annual hazard rate parameter
$\overline{\mathfrak{D}}_\infty \in [0.025, 0.10]$	mean tipping damage level in the long run
$q \in [0, 0.4]$	mean-square-variance ratio for the uncertain tipping damage level
$\overline{\Gamma} \in [5, 200]$	expected duration (years) of the whole tipping process

Table A.6: Parameters and variables for the climate tipping element

A.2 Calibration and Discretization of the Productivity Process

We construct a time-dependent, finite-state Markov chain for the productivity process (ζ_t, χ_t), which depends on three parameters: ϱ, r, and ς. Our calibration target is to choose these three parameters so that we match the conditional and unconditional moments of endogenous consumption growth with the statistics from observed annual market data. The data on consumption growth are from 1930 to 2008. For each choice of ϱ, r andς, and the sizes of the Markov chains, n_ζ and n_χ, we solve DSICE assuming damage from climate is zero (i.e., $J_t = \mathcal{J}_0$ and $\pi_1 = \pi_2 = 0$) because there is little climate damage in the period with observed consumption data. We also use our default case preference parameters, $\psi = 1.5$ and $\gamma = 10$. Once we obtained a solution, we computed 10,000 simulations of the consumption process over the first century and computed the conditional and unconditional moments of the per capita consumption growth paths.

We now present the details. Our construction of the Markov chain is adapted from the following processes:[28]

$$\log\left(\hat{\zeta}_{t+1}\right) = \log\left(\hat{\zeta}_t\right) + \hat{\chi}_t + \varrho\hat{\omega}_{\zeta,t},$$

[28]To avoid confusion, we use $\widehat{\zeta}_t$ and $\widehat{\chi}_t$ to represent the continuous random variables at each time t, while ζ_t and χ_t represent the Markov chains.

$$\hat{\chi}_{t+1} = r\hat{\chi}_t + \varsigma\hat{\omega}_{\chi,t},$$

where $\hat{\omega}_{\zeta,t}$, $\hat{\omega}_{\chi,t} \sim i.i.d.\mathcal{N}(0,1)$. We know that $\widehat{\chi}_t$ is a normal random variable with mean 0. Denote $\Upsilon_t \equiv \text{Var}\left\{\widehat{\chi}_t\right\}$, we have $\Upsilon_{t+1} = r^2\Upsilon_t + \varsigma^2$, then from recursive iteration we get

$$\Upsilon_t = \frac{\varsigma^2\left(1 - r^{2t}\right)}{1 - r^2} \tag{A.6}$$

for $t > 0$. We also know that $\log\left(\widehat{\zeta}_t\right)$ is a normal random variable with mean 0, and

$$\log\left(\widehat{\zeta}_t\right) = \sum_{s=1}^{t-1}\widehat{\chi}_t + \varrho\sum_{s=0}^{t-1}\hat{\omega}_{\zeta,s}.$$

Denote $\Delta_t \equiv \text{Var}\left\{\log\left(\widehat{\zeta}_t\right)\right\}$. Since $\mathbb{E}\left\{\widehat{\chi}_t\widehat{\chi}_s\right\} = r^{|t-s|}\Upsilon_{\min\{t,s\}}$, we have

$$\Delta_t = \mathbb{E}\left\{\left(\log\left(\widehat{\zeta}_t\right)\right)^2\right\} = \left(\sum_{s=1}^{t-1}\Upsilon_s + 2\sum_{\tau=2}^{t-1}\sum_{s=1}^{\tau-1}r^{\tau-s}\Upsilon_s\right) + \varrho^2 t, \tag{A.7}$$

for $t > 0$.

Now we use the the values of Υ_t in (A.6) and Δ_t in (A.7) to define the values of our Markov chains (ζ_t, χ_t). We choose the values of χ_t as $\{\chi_{t,j} : j = 1, ..., n_\chi\}$ where $\chi_{t,j}$ are equally spaced in $\left[-3\sqrt{\Upsilon_t}, 3\sqrt{\Upsilon_t}\right]$ (the probability that the continuously distributed random variable $\widehat{\chi}_t$ is in $\left[-3\sqrt{\Upsilon_t}, 3\sqrt{\Upsilon_t}\right]$ is 99.7 percent), at each time t. Similarly, we choose the values of ζ_t as $\{\zeta_{t,i} : i = 1, ..., n_\zeta\}$ so that $\log(\zeta_{t,i})$ are equally spaced in $\left[-3\sqrt{\Delta_t}, 3\sqrt{\Delta_t}\right]$ (the probability that the continuously distributed random variable $\log\left(\widehat{\zeta}_t\right)$ is in $\left[-3\sqrt{\Delta_t}, 3\sqrt{\Delta_t}\right]$ is 99.7 percent).

Next we set the transition probability matrices for our Markov chains (ζ_t, χ_t). For the paired values $\{(\zeta_{t,i}, \chi_{t,j}) : i = 1, ..., n_\zeta, j = 1, ..., n_\chi\}$ ($\zeta_{0,i} \equiv 1$ and $\chi_{0,j} \equiv 0$), we apply the method in Tauchen (1986) to define the transition probability from $\chi_{t,j}$ to $\chi_{t+1,j'}$:

$$\begin{aligned}\Pr\left\{\chi_{t+1,j'} \mid \chi_{t,j}\right\} &= \Phi\left(\frac{1}{\varsigma}\left(\frac{\chi_{t+1,j'} + \chi_{t+1,j'+1}}{2} - r\chi_{t,j}\right)\right) \\ &\quad -\Phi\left(\frac{1}{\varsigma}\left(\frac{\chi_{t+1,j'-1} + \chi_{t+1,j'}}{2} - r\chi_{t,j}\right)\right),\end{aligned}$$

for $j' = 2, ..., n_\chi - 1$, and

$$\Pr\left\{\chi_{t+1,1} \mid \chi_{t,j}\right\} = \Phi\left(\frac{1}{\varsigma}\left(\frac{\chi_{t+1,1} + \chi_{t+1,2}}{2} - r\chi_{t,j}\right)\right),$$

$$\Pr\left\{\chi_{t+1,n} \mid \chi_{t,j}\right\} = 1 - \Phi\left(\frac{1}{\varsigma}\left(\frac{\chi_{t+1,n-1} + \chi_{t+1,n}}{2} - r\chi_{t,j}\right)\right),$$

where $\Phi(\cdot)$ is the cumulative normal distribution function, for any $j = 1, ..., n_\chi$. Similarly, the transition probability from $(\zeta_{t,i}, \chi_{t,j})$ to $\zeta_{t+1,i'}$ is defined as

$$\begin{aligned}\Pr\left\{\zeta_{t+1,i'} \mid (\zeta_{t,i}, \chi_{t,j})\right\} &= \Phi\left(\frac{1}{\varrho}\left(\frac{\log\left(\zeta_{t+1,i'}\right) + \log\left(\zeta_{t+1,i'+1}\right)}{2} - \left(\log\left(\zeta_{t,i}\right) + \chi_{t,j}\right)\right)\right) \\ &\quad -\Phi\left(\frac{1}{\varrho}\left(\frac{\log\left(\zeta_{t+1,i'}\right) + \log\left(\zeta_{t+1,i'+1}\right)}{2} - \left(\log\left(\zeta_{t,i}\right) + \chi_{t,j}\right)\right)\right),\end{aligned}$$

for $i' = 2, ..., n_\zeta - 1$, and

$$\Pr\left\{\zeta_{t+1,1} \mid (\zeta_{t,i}, \chi_{t,j})\right\} = \Phi\left(\frac{1}{\varrho}\left(\frac{\log\left(\zeta_{t+1,1}\right) + \log\left(\zeta_{t+1,2}\right)}{2} - \left(\log\left(\zeta_{t,i}\right) + \chi_{t,j}\right)\right)\right),$$

$$\Pr\left\{\zeta_{t+1,n} \mid (\zeta_{t,i}, \chi_{t,j})\right\} = 1 - \Phi\left(\frac{1}{\varrho}\left(\frac{\log\left(\zeta_{t+1,n-1}\right) + \log\left(\zeta_{t+1,n}\right)}{2} - \left(\log\left(\zeta_{t,i}\right) + \chi_{t,j}\right)\right)\right),$$

for any $i = 1, ..., n_\zeta, j = 1, ..., n_\chi$.

We examine sizes of the Markov chains, and find that $n_\zeta = 91$ and $n_\chi = 19$ are enough for a good approximation (i.e., distributions of the solutions of our stochastic growth benchmark example in the first 100 years are almost invariant to a higher n_ζ or n_χ) with the calibrated values at $\varrho = 0.035$, $r = 0.775$, and $\varsigma = 0.008$. Moreover, with $n_\zeta = 91$ and $n_\chi = 19$, our transition probability matrix is not nearly degenerated at all times in our examples.[29]

We compare the performance of DSICE to match the statistics of market data with the perfor-

[29] A nearly degenerated probability matrix means that at one state, the maximum of its nonzero transition probabilities is more than 99.9%.

mance of the study by Jensen and Traeger (2014) who specify a model for labor productivity growth. Because the logarithm of the total factor productivity is equal to $(1-\alpha)$ times the logarithm of the labor productivity, where $\alpha = 0.3$ is the output elasticity of capital, the process implied in Jensen and Traeger (2014) is equivalent to the following total factor productivity process:

$$\log\left(\tilde{\zeta}_{t+1}\right) = \log\left(\tilde{\zeta}_t\right) + \tilde{\chi}_t + 0.0133\tilde{\omega}_{\zeta,t},$$

$$\tilde{\chi}_{t+1} = 0.5\tilde{\chi}_t + 0.0133\tilde{\omega}_{\chi,t}.$$

Table A.7 provides a comparison of the statistics of the observed annual consumption data with the simulation statistics of DSICE (with $\varrho = 0.035$, $r = 0.775$, and $\varsigma = 0.008$), and with the simulation statistics of DSICE when using the Jensen and Traeger (2014) parameters ($\varrho = 0.0133$, $r = 0.5$, and $\varsigma = 0.0133$). For each simulation path of the per capita consumption growth, g_c, we compute its mean $\mathbb{E}(g_c)$, standard deviation $\sigma(g_c)$ and order-j autocorrelations (i.e., the correlation between $g_{c,t+j}$ and $g_{c,t}$). Using 10,000 simulation paths of DSICE, we compute the medians and confidence intervals of $\mathbb{E}(g_c)$, $\sigma(g_c)$, and order-j autocorrelations. Nordhaus (2008) uses an initial-time growth rate of 0.92 percent (and decreasing afterward) for the productivity trend, which leads to 1.3 percent for per capita consumption growth for this century in DICE-2007 and DICE-CJL with $\psi = 0.5$. We follows Nordhaus (2008) and use the same productivity trend, so our median of $\mathbb{E}(g_c)$ is also 1.3 percent; smaller than the market estimate of the per capita consumption growth in the last century. However, our confidence interval still contains the historical estimate. The medians of $\sigma(g_c)$ and the order-1 autocorrelation are 0.023 and 0.43, close to the corresponding numbers from the annual market data, and the order-2 autocorrelation of the data is also inside its 90 percent corresponding confidence interval from our simulations.

However, the process in Jensen and Traeger (2014) has a much smaller median standard deviation, only 0.013, and even its confidence interval $[0.010, 0.016]$ does not contain the market estimate of the standard deviation, 0.022 . Moreover, the 90 percent confidence interval of the order-2 autocorrelation of the simulation from the Jensen–Traeger process, $[0.27, 0.70]$, does not contain the order-2 autocorrelation of the data.

We also performed a lag-1 autoregression analysis for each consumption growth path $g_{c,t}$, that

is, we assume

$$g_{c,t+1} - \overline{g}_{c,t+1} = \Lambda(g_{c,t} - \overline{g}_{c,t}) + \epsilon_t, \tag{A.8}$$

where $\overline{g}_{c,t}$ is the mean of g_t from the 10,000 simulation points at time t, and Λ is the autoregression coefficient. For each simulation path, we use the first 100 years to get an estimate of Λ, and the standard deviation of the lag-1 autoregression residuals, denoted $\sigma(\epsilon)$, which is also known as the one-period-ahead conditional standard deviation. In total, we obtain 10,000 estimates of Λ and $\sigma(\epsilon)$.

From Table A.7, we see that the autoregression statistics with our calibrated parameter values also perform better in matching the data than the statistics resulting from the parameter choices in Jensen and Traeger (2014), particularly $\sigma(\epsilon)$ as its value from observed data, 0.0179, is outside of the confidence interval from the Jensen–Traeger process, but is still within the 90 percent confidence interval from DSICE.

	Observed data	DSICE			Jensen and Traeger (2014)		
Variable	Estimate	Median	5%	95%	Median	5%	95%
$\mathbb{E}(g_c)$	0.019	0.013	0.002	0.025	0.013	0.006	0.020
$\sigma(g_c)$	0.022	0.023	0.019	0.028	0.013	0.010	0.016
order-1 autocorrelation	0.48	0.43	0.19	0.64	0.55	0.30	0.73
order-2 autocorrelation	0.17	0.37	0.13	0.59	0.51	0.27	0.70
Λ	0.46	0.48	0.24	0.68	0.59	0.36	0.76
$\sigma(\epsilon)$	0.0179	0.0203	0.0177	0.023	0.011	0.009	0.012

Table A.7: Statistics of per capita consumption growth

A.3 Calibration of the Deterministic Climate System

The climate system parameters of Nordhaus (2008) were chosen to represent a 10-year time period discretization and match the business-as-usual path (i.e., no emissions control) in Nordhaus (2008). Analyses of economic uncertainty and productivity shocks typically use much shorter time periods. Our choice of one-year time periods for the economic system requires us to choose climate system parameters, ϕ_{12}, ϕ_{21}, ϕ_{23}, ϕ_{32}, ξ_1, ξ_2, φ_{12}, and φ_{21}, consistent with one-year periods. As in Nordhaus (2008), $\phi_{21} = \phi_{12}\tilde{M}^*_{\mathrm{AT}}/\tilde{M}^*_{\mathrm{UO}}$ and $\phi_{32} = \phi_{23}\tilde{M}^*_{\mathrm{UO}}/\tilde{M}^*_{\mathrm{LO}}$, where $\tilde{M}^*_{\mathrm{AT}} = 587.5$, $\tilde{M}^*_{\mathrm{UO}} = 1144$, and $\tilde{M}^*_{\mathrm{LO}} = 18340$ are the preindustrial equilibrium carbon concentrations in the atmosphere, upper ocean, and lower ocean, respectively; moreover, we also set $\xi_2 = \xi_1\eta/\xi_3$, where $\xi_3 = 3$ is

the equilibrium climate sensitivity and $\eta = 3.8$ is the radiative forcing parameter, as in Nordhaus (2008). Thus, there are five parameters left for the calibration: ϕ_{12}, ϕ_{23}, ξ_1, φ_{12}, and φ_{21}.

We choose parameters to match the first 500 years of the business-as-usual paths. At first, we compute the business-as-usual data on the dynamics of the three-dimensional carbon cycle and two-dimensional temperature from the model of Nordhaus (2008), denoted as $\mathbf{M}_{\mathrm{N}}(t) = (M_{\mathrm{N,AT}}(t), M_{\mathrm{N,UO}}(t), M_{\mathrm{N,LO}}(t)$ and $\mathbf{T}_{\mathrm{N}}(t) = (T_{\mathrm{N,AT}}(t), T_{\mathrm{N,OC}}(t))$, respectively, for only the values of t at the decadal time nodes, $t_0 = 0$, $t_1 = 10$, $t_2 = 20$, ..., and $t_{50} = 500$. These serve as our data. We take the rate of emissions in the Nordhaus business-as-usual path at decadal time, and use linear interpolation to produce data on the business-as-usual emissions path, denoted $E_{\mathrm{N}}(t)$, for each year t. For each choice of the parameters, we solve the climate system equations with emissions set at $E_{\mathrm{N}}(t)$ and find the implied climate system paths, $\mathbf{M}(t) = (M_{\mathrm{AT}}(t), M_{\mathrm{UO}}(t), M_{\mathrm{LO}}(t))$, and $\mathbf{T}(t) = (T_{\mathrm{AT}}(t), T_{\mathrm{OC}}(t))$. We choose values for the parameters ϕ_{12}, ϕ_{23}, ξ_1, φ_{12}, and φ_{21} to minimize the $\mathcal{L}^2$ norm of the difference between the decadal time series data, $\mathbf{M}_{\mathrm{N}}(t_i)$ and $\mathbf{T}_{\mathrm{N}}(t_i)$, and our annual paths at those decadal times, $\mathbf{M}(t_i)$ and $\mathbf{T}(t_i)$; more precisely, we solve

$$
\begin{aligned}
\min_{\phi_{12},\phi_{23},\xi_1,\varphi_{12},\varphi_{21}} \quad & \sum_{i=1}^{50}\left\{\left(\frac{M_{\mathrm{AT}}(t_i) - M_{\mathrm{N,AT}}(t_i)}{M_{\mathrm{N,AT}}(t_i)}\right)^2 + \left(\frac{M_{\mathrm{UO}}(t_i) - M_{\mathrm{N,UO}}(t_i)}{M_{\mathrm{N,UO}}(t_i)}\right)^2 + \right. \\
& \left.\left(\frac{M_{\mathrm{LO}}(t_i) - M_{\mathrm{N,LO}}(t_i)}{M_{\mathrm{N,LO}}(t_i)}\right)^2 + \left(\frac{T_{\mathrm{AT}}(t_i) - T_{\mathrm{N,AT}}(t_i)}{T_{\mathrm{N,AT}}(t_i)}\right)^2 + \left(\frac{T_{\mathrm{OC}}(t_i) - T_{\mathrm{N,OC}}(t_i)}{T_{\mathrm{N,OC}}(t_i)}\right)^2\right\}, \\
\text{s.t.} \quad & \mathbf{M}(t+1) = \mathbf{\Phi}_{\mathrm{M}}\mathbf{M}(t) + (E_{\mathrm{N}}(t), 0, 0)^{\top}, \\
& \mathbf{T}(t+1) = \mathbf{\Phi}_{\mathrm{T}}\mathbf{T}(t) + (\xi_1 \mathcal{F}_t(M_{\mathrm{AT}}(t)), 0)^{\top}, \\
& t = 0, 1, 2, ..., 500.
\end{aligned}
$$

The solution is $\phi_{12} = 0.019$, $\phi_{23} = 0.0054$, $\xi_1 = 0.037$, $\varphi_{12} = 0.01$, and $\varphi_{21} = 0.0048$, which then implies $\phi_{21} = \phi_{12}\tilde{M}^*_{\mathrm{AT}}/\tilde{M}^*_{\mathrm{UO}} = 0.01$, $\phi_{32} = \phi_{23}\tilde{M}^*_{\mathrm{UO}}/\tilde{M}^*_{\mathrm{LO}} = 0.00034$, and $\xi_2 = \xi_1\eta/\xi_3 = 0.047$.

Figure A.1 shows the five business-as-usual paths from the model of Nordhaus (2008) with decadal time steps in marks, circles, and pluses, and the corresponding fitted business-as-usual paths from our annualized deterministic model at the calibrated parameter values. The match is very good in the first 100 years, a horizon often used for calibration in the climate literature. The errors in the fit are in the later centuries, where they are unlikely to significantly affect the results

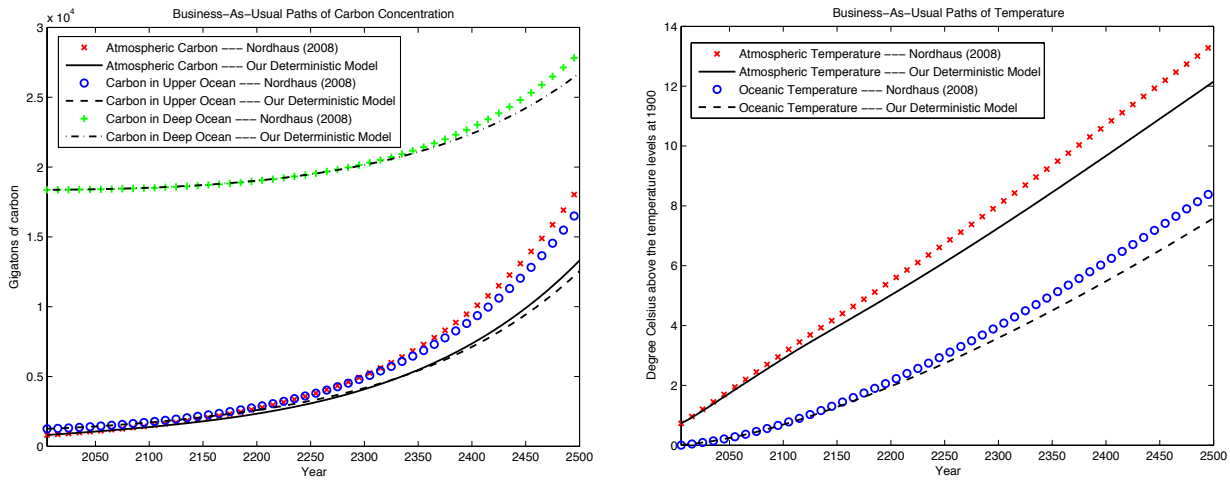

Figure A.1: Business-as-usual paths for calibration

for the first century, the focus of this study.

A.4 Calibration of the Stochastic Climate Tipping Process

One example of tipping elements would be the glaciers in the west Antarctic. Suppose that we believe this system of glaciers is stable at preindustrial temperatures but worry about the impact of warming. Increases in atmospheric temperatures could accelerate melting in the glaciers. Perhaps a small increase in temperature and melting will not upset their long-run stability, but higher temperatures may increase melting by so much that the glaciers' structure is disturbed and melting is accelerated and will continue even if temperatures were to go back to preindustrial levels.

The climate literature has identified several major tipping elements in the climate system. An example of a major tipping element is irreversible melting of the Greenland ice sheet. Recent projections of future global warming suggest that it is more likely than not that the Greenland ice sheet tipping point will be triggered within this century (IPCC 2014). Joughin et al. (2014) argue that marine west Antarctic ice sheet collapse is already under way. Melting of the Greenland ice sheet (which contains the equivalent of about seven meters of global sea level) could lead to a global sea level rise of up to 0.5–1 meter per century (Lenton et al. 2008). Other examples of climate tipping elements include the weakening or shutdown of the Atlantic thermohaline circulation or the dieback of the Amazon rainforest.

We first calibrate the probability of triggering the tipping point event. Given Equation (18),

we need to calibrate λ, the hazard rate parameter, and $\underline{T_{\text{AT}}}$, the lower bound of atmospheric temperature for which $p_{\text{tip},0} = 0$. Ideally, climate scientists would specify a value for the hazard rate parameter λ. Unfortunately, no studies exist that would provide specific conditional probabilities of triggering a tipping point event for any given location in the state space. According to Kriegler et al. (2009) there is a substantial lack of knowledge about the underlying physical processes of climate tipping elements. The most reliable source of information is reviews of expert opinion on the cumulative trigger probability at some distant time given some degree of global warming, for example 3 °C in 2100 (see Kriegler et al. 2009; Zickfeld et al. 2007; and Lenton 2010). The subjective probability assessments differ significantly between experts, albeit the range of uncertainty about the likelihood of triggering a tipping point is much lower for some climate elements (Kriegler et al. 2009). However, the absence of precise knowledge about the physical system is not essentially a problem for our analysis: decisions today must be based on current beliefs about the climate system and will reflect the imprecise nature of those beliefs. Therefore, we make use of the assessed cumulative trigger probabilities of tipping and infer trigger hazard rates from them. Doing so, we assume tacitly that the subjective expert opinions are the same as those held by the social planner. Lenton (2010) studies several climate elements and provides the cumulative probability of triggering a tipping point of each element. We choose Lenton (2010) as the source of data because he summarizes the findings from Kriegler et al. (2008) and other expert elicitation studies. Since our interest is in modeling a representative tipping element we use the average of these numbers to approximate the probability of triggering a representative climate tipping element.[30]

The first two rows of Table A.8 depict the cumulative probabilities of triggering a climate tipping event until 2100 for different levels of global warming until 2100. We have inferred these numbers from Lenton (2010).[31] These numbers imply, for example, that given a 4 °C increase in global warming in the year 2100, the likelihood of a tipping point event being triggered until 2100 is 50 percent. Our goal is to use these numbers to compute an annual conditional trigger probability. We do this by first specifying the general relationship between a hazard rate h_t and the contemporaneous level of global warming above the 2000 temperature in a simple linear form by $h_t = \lambda \tilde{T}_{\text{AT},t}$, where λ

[30]The tipping elements used for the calibration of the benchmark tipping element are: Arctic summer sea ice, Greenland ice sheet, Amazon rain forests, west Antarctic ice sheet, Boreal forests, Atlantic thermohaline circulation, El Niño—Southern Oscillation, and West African monsoon.

[31]We thank Timothy Lenton for helping us identify the probability numbers for each tipping element and their average, based on Lenton (2010).

is a hazard rate parameter and $\tilde{T}_{\mathrm{AT},t}$ is the nominal change in temperature above the year 2000 level (t represents year $t+2000$ here). We then assume that experts in general express their subjective probabilities while bearing in mind that temperature linearly evolves until the specified level at the end of this century.[32]

We then integrate the hazard rate h_t for up for 100 years and obtain the cumulative hazard H_{100}. The probability of a tipping point having *not* occurred during this century is then $\exp(-H_{100})$. It follows that P, the cumulative probability of a tipping point having been triggered until 2100, is $P = 1 - \exp(-H_{100})$. With the former equation we can solve for the hazard rate parameter λ:

$$\lambda = -\frac{\log(1-P)}{50\left(\tilde{T}_{\mathrm{AT},100} - \tilde{T}_{\mathrm{AT},0}\right)}. \tag{A.9}$$

$\tilde{T}_{\mathrm{AT},100} - \tilde{T}_{\mathrm{AT},0}$	1 °C	2 °C	3 °C	4 °C	5 °C	6 °C
P (cumulative probability until 2100)	12.5%	25%	37.5%	50%	62.5%	75%
λ (inferred hazard rate parameter)	0.00267	0.00288	0.00313	0.00347	0.00392	0.00462

Table A.8: Cumulative probabilities of triggering a representative tipping element (until 2100) for different levels of global warming and the inferred hazard rate parameter

We use the above expression for the hazard rate parameter and plug in for each column of Table A.8 the levels of $\tilde{T}_{\mathrm{AT},100} - \tilde{T}_{\mathrm{AT},0}$ and P. We report the computed hazard rate parameters in the bottom row of Table A.8. For our previous example of a 4 °C increase in global warming and a cumulative trigger probability of 50 percent, our calibration implies a hazard rate parameter of roughly 0.0035, which is approximately the central value in our calibrated range. Practically, $\lambda = 0.0035$ implies that, for example, a two-degree increase in temperature (above the year 2000 temperature) until year 2090 would yield a conditional probability of triggering the tipping point event in year 2090 of about 0.7 percent. We choose $\lambda = 0.0035$ as the default parameter setting for our benchmark case and set up a range for the sensitivity analysis—$\lambda \in [0.0025, 0.0045]$. Since we choose not to model any specific tipping element explicitly, we set up the range with respect to how different hazard rate specifications relate to the optimal SCC. Finally, we set $\underline{T_{\mathrm{AT}}} = 1$ °C, which we

[32]The linearity assumption comes in handy in deriving a simple formula for the hazard rate parameter, while at the same time not being incompatible with the recent temperature scenarios used by the Intergovernmental Panel on Climate Change, which we have obtained from Meinshausen et al. (2011).

calibrate as today's global warming level using a climate emulator (Meinshausen et al. 2011).

Next, we turn to the calibration of $\overline{\Gamma}$, the expected duration of the tipping process. Climate scientists note that tipping elements in the climate system exhibit heterogeneity with respect to the assumed duration of the post-tipping process. Lenton et al. (2008) consider current climate models and paleo-data, and argue that the post-tipping process may last less than ten years or up to several centuries. Table A.9 presents the expected transition scale for the selection of the most prominent tipping elements drawn from Lenton et al. (2008), where we consider the lower estimates in Lenton et al. (2008) based on the limits of physical plausibility or on past experience.

Tipping Element	Greenland ice sheet	Arctic summer sea ice	Amazon rainforest	Boreal forests	Atlantic thermohaline circulation	West Antarctic ice sheet	West African monsoon	El Niño—Southern Oscillation
Expected transition scale (years)	>300	10	50	50	100	>300	10	100

Table A.9: Duration of expected transition scales for different climate tipping elements

Following the numbers in Table A.9 we choose $\overline{\Gamma} = 50$ (years) for our benchmark tipping process. In addition, in our sensitivity analyses we also consider $\overline{\Gamma} = 5$ (years) as the lower bound, $\overline{\Gamma} = 100$ (years), and $\overline{\Gamma} = 200$ (years) as the upper bound.[33] As mentioned in the main text, we assume that each of our five stages occurs stochastically and the expected transition time at each stage of the post-tipping process is $\overline{\Gamma}/4$ years.

In this paper, all of our examples of tipping elements assume that there are three possible tipping damage levels in the long run if $q > 0$, and that we do not know the long-run damage level until tipping begins. Since the variance of the long-run tipping damage level is $q\overline{\mathfrak{D}}_\infty^2$, we assume that the values of the three possible tipping damage levels in the long run are $\left(1 + (i-2)\sqrt{1.5q}\right)\overline{\mathfrak{D}}_\infty$, for $i = 1, 2, 3$, with equal probability. For our benchmark case with $\overline{\mathfrak{D}}_\infty = 5\%$ and $q = 0.2$, the three possible long-run damage levels are 2.26%, 5%, and 7.74%, each with 1/3 probability. In the extreme cases of $\overline{\mathfrak{D}}_\infty = 2.5\%$ (or $\overline{\mathfrak{D}}_\infty = 10\%$) and $q = 0.4$ in our sensitivity analysis, the three possible long-run damage levels are 0.56%, 2.5%, and 4.44% (or 2.25%, 10%, and 17.75%).

We also assume that the cascading damage linearly increases over the five post-tipping stages.

[33] For $\overline{\Gamma} > 200$ we find very little change in the 2005 SCC as compared to $\overline{\Gamma} = 200$.

That is, if the post-tipping Markov chain $\mathcal{M}_i$ is realized, then its corresponding level of long-run tipping damage is $\left(1+(i-2)\sqrt{1.5q}\right)\overline{\mathfrak{D}}_\infty$, and the damage at the j-th post-tipping stage is $\left(1+(i-2)\sqrt{1.5q}\right)\overline{\mathfrak{D}}_\infty \cdot j/5$, for $j=1,2,...,5$. Thus, the values of J_t are $\mathcal{J}_0=0$ and

$$\mathcal{J}_{i,j}=\frac{j}{5}\left(1+(i-2)\sqrt{1.5q}\right)\overline{\mathfrak{D}}_\infty, \tag{A.10}$$

for $i=1,2,3$, and $j=1,...,5$. When $q=0$, it means that there is only one possible post-tipping Markov chain, and there are only six values of J_t: $\mathcal{J}_0=0$ and $\mathcal{J}_{i,j}=\overline{\mathfrak{D}}_\infty \cdot j/5$ for $j=1,2,...,5$ and all i.

Since we assume that we know the long-run damage level once the tipping event happens and the three possible levels are equally distributed, the probability from the initial pre-tipping state $\mathcal{J}_0=0$ to $\mathcal{J}_{i,1}$ at time t is $p_{\text{tip},t}/3$ for $i=1,2,3$. Moreover, in the post-tipping Markov chain $\mathcal{M}_i$, if its current state is $\mathcal{J}_{i,j}$ with $j<5$, meaning that it is at the j-th post-tipping stage, then it could either move to state $\mathcal{J}_{i,j+1}$ with the probability $p=1-\exp\left(-4/\overline{\Gamma}\right)$, or stay with the probability $1-p$. Thus the probability transition matrix for $\mathcal{M}_i$ is

$$\begin{bmatrix} 1-p & p & 0 & 0 & 0 \\ 0 & 1-p & p & 0 & 0 \\ 0 & 0 & 1-p & p & 0 \\ 0 & 0 & 0 & 1-p & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

A.5 Value Function Iteration

Here, we summarize the numerical method we used to solve the dynamic programming problems in this study. The value functions for our problems are continuous, and almost everywhere differentiable in the continuous state variables. In each case, we approximate the value function $V(x,\theta)$ using a functional form $\hat{V}(x,\theta;\mathbf{b})$ with a finite number of parameters, $\mathbf{b}$. In DSICE, the state variable x represents the six-dimensional vector $(K,\mathbf{M},\mathbf{T})$ of continuous states, and θ represents the three-dimensional vector of discrete states, (ζ,χ,J), contained in a finite set Θ. The functional form $\hat{V}$ may be a linear combination of polynomials, a rational function, a neural net, or

some other parameterization specially designed for the problem. After the functional form is set, we focus on finding the vector of parameters, $\mathbf{b}$, such that $\hat{V}(x,\theta;\mathbf{b})$ approximately satisfies the Bellman equation (Bellman 1957). Numerical dynamic programming with value function iteration can solve the Bellman equation approximately (Judd 1998).

A general dynamic programming model is based on the Bellman equation:

$$\begin{aligned} V_t(x,\theta) = \max_{a\in\mathcal{D}(x,\theta,t)} \quad & u_t(x,a) + \beta\mathcal{H}_t\left(V_{t+1}(x_+,\theta_+)\right), \\ \text{s.t.} \quad & x_+ = F(x,\theta,a), \\ & \theta_+ = G(x,\theta,\omega), \end{aligned}$$

where $V_t(x,\theta)$ is the value function at time $t \leq \mathcal{T}$ (the terminal value function $V_{\mathcal{T}}(x,\theta)$ is given), (x_+,θ_+) is the next-stage state, $\mathcal{D}(x,\theta,t)$ is a feasible set of a, ω is a random variable vector, F and G are the transition laws of x and θ, β is a discount factor and $u_t(x,a)$ is the utility function, and $\mathcal{H}_t$ is a functional operator[34] at time t. The following is the algorithm of parametric dynamic programming with value function iteration for finite horizon problems.

Algorithm 1. *Numerical Dynamic Programming with Value Function Iteration for Finite Horizon Problems*

Initialization. *Choose the approximation nodes, $\mathbb{X}_t = \{x_{i,t} : 1 \leq i \leq m_t\}$ for every $t < \mathcal{T}$, and choose a functional form for $\hat{V}(x,\theta;\mathbf{b})$, where $\theta \in \Theta$. Let $\hat{V}(x,\theta;\mathbf{b}_{\mathcal{T}}) \equiv V_{\mathcal{T}}(x,\theta)$. Then for $t = \mathcal{T}-1, \mathcal{T}-2, \ldots, 0$, iterate through steps 1 and 2.*

[34] A typical one is $\mathbb{E}_t$, the expectation operator conditional on time-t information. In DSICE, it is

$$\mathcal{H}_t\left(V_{t+1}\right) = \left[\mathbb{E}_t\left\{V_{t+1}^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\right\}\right]^{\frac{1-\frac{1}{\psi}}{1-\gamma}}.$$

Step 1. *Maximization step. Compute*

$$v_{i,j} = \max_{a \in \mathcal{D}(x_i, \theta_j, t)} u_t(x_i, a) + \beta \mathcal{H}_t \left(\hat{V} \left(x_+, \theta_{j+}; \mathbf{b}_{t+1} \right) \right)$$
$$\text{s.t.} \quad x_+ = F(x_i, \theta_j, a),$$
$$\theta_{j+} = G(x_i, \theta_j, \omega),$$

for each $\theta_j \in \Theta$, $x_i \in \mathbb{X}_t$, $1 \leq i \leq m_t$.

Step 2. *Fitting step. Using an appropriate approximation method, compute the* $\mathbf{b}_t$ *such that* $\hat{V}(x, \theta_j; \mathbf{b}_t)$ *approximates* $(x_i, v_{i,j})$ *data for each* $\theta_j \in \Theta$.

We considered alternative functional forms for $\hat{V}(x, \theta; \mathbf{b})$, including ordinary polynomials and polynomials in logs. We were always able to find some parsimonious functional form for approximating the value function. Detailed discussion of numerical dynamic programming methods can be found in Cai (2010), Cai and Judd (2010), Judd (1998), and Rust (2008).

The terminal value function is computed by assuming that after 2600, the system is deterministic, population and productivity growth ends, all emissions are eliminated, and the consumption–output ratio is fixed at 0.78. Consideration of alternatives showed that changes in the terminal value function at 2600 had no significant impact on any results for the twenty-first century.

A.6 Constructing the State Space Domains

The key challenge for any dynamic programming problem, deterministic or stochastic, is to accurately approximate the value function over the set of states that could be visited by the optimal decision rules. Those domains vary greatly across the various instances of DSICE we solve. Furthermore, in each case, the approximation domain is different at each time t, starting with a narrow domain at $t = 0$ and expanding as t increases. The state variable domain at every time t contains the optimal state at time t in the solution to the deterministic case. The time $t+1$ domain is chosen so that any combination of time t states, time t optimal action, and time t shocks will be a point

inside the time $t+1$ domain . For example, Figure A.2B shows the lower and upper bounds of the capital stock (on $\log_{10}$ scale) for the stochastic growth benchmark case.

The approximation method will differ across problems due to the different ranges of state variables and the different amount of curvature. For example, of the three benchmark cases, the approximation domains for the benchmark with both stochastic growth and climate tipping are the widest because it has the largest volatility, and we use degree-8 complete Chebyshev polynomials for approximation. For the stochastic growth benchmark case, we also use degree-8 complete Chebyshev polynomials. The climate tipping benchmark case has a narrower band of possible paths, allowing us to achieve high accuracy with degree-4 complete Chebyshev polynomials. Some experimentation is required to find good choices for the approximation method, but this was not a major challenge.

A.7 Code Test

In order to have confidence in our code, we compared how our code performed on a few cases where we could compute the true solution using another method. Our tests used two ways of computing solutions to deterministic versions of DSICE—that is, versions where we eliminate all randomness. The two methods were our dynamic programming code and optimal control methods. Since the latter can be computed with high accuracy with standard finite-dimensional solvers, we can compare those solutions with the results of the dynamic programming algorithm.

We performed three tests for solving the deterministic versions of DSICE, each using the same approximation domains and polynomial degrees as those used in one of the three benchmark cases, as our approximation domains of the state variables in the stochastic versions of DSICE contain the optimal states in the deterministic versions. Table A.10 shows the relative $\mathcal{L}^1$ errors, over the first one hundred years, of three of the endogenous state variables (K, M_{AT}, T_{AT}), two control variables (C, μ), and SCC. It also shows the relative error in 2005 for the control variables and SCC. Table A.10 shows that all the relative errors were small. This gives us confidence that our implementation is reliable.

Variable	Approximation degree and domains of the stochastic growth benchmark		Approximation degree and domains of the climate tipping benchmark		Approximation degree and domains of the stochastic growth & climate tipping benchmark	
	first 100 years	2005	first 100 years	2005	first 100 years	2005
K	5.5(−3)	—	2.1(−4)	—	4.0(−3)	—
M_{AT}	1.5(−4)	—	1.3(−5)	—	1.1(−4)	—
T_{AT}	8.7(−5)	—	2.5(−5)	—	8.3(−5)	—
C	2.0(−3)	4.6(−5)	2.4(−5)	2.6(−5)	1.6(−3)	4.0(−5)
μ	2.4(−3)	1.3(−4)	4.4(−4)	1.7(−4)	2.0(−3)	1.3(−4)
SCC	9.2(−3)	5.4(−3)	4.1(−3)	7.2(−4)	1.0(−2)	6.2(−3)

Table A.10: Relative errors for the solution of the deterministic case obtained by our numerical dynamic programming algorithm with approximation domains and polynomial degrees used by the three benchmark cases. Note: $a(-n)$ means $a \times 10^{-n}$.

A.8 Dynamics of the Economic System in the Stochastic Growth Benchmark Case

Figure A.2 shows the dynamics of gross world output (Y_t), total capital (K_t), and the per capita consumption growth rate. Both gross world output and the total capital stock grow exponentially, which is not surprising since productivity grows exponentially. The uncertainty of both is substantial; for example, the 10 percent and 90 percent quantiles of the capital stock in 2100 are $270 trillion and $1,500 trillion respectively. The growth of per capita consumption in Figure A.2C ranges from −4.4 percent to 7.1 percent. The mean and median of the dynamic stochastic equilibrium paths are close to the deterministic solution, but the variation around those paths is very large.

The wide range of possible outcomes for the total capital stock displayed in Figure A.2B implied that our approximation of the value function had to be defined over an even larger range. The blue dashed lines in Figure A.2B represent, for each time t, the minimum and maximum level of the capital stock (on $\log_{10}$ scale) over which the value function is computed. The ratio of maximum to minimum capital stock exceeded 175 in 2100, and continues to grow thereafter. If we had used significantly smaller ranges, then the Bellman equation for states close to the maximum or minimum capital stock would try to evaluate the value function beyond the domain on which it is approximated. These ranges are the natural outcome of the productivity process. The large range for capital creates difficulties for simpler approximation methods, such as orthogonal polynomials,

but nonlinear changes in variables allowed us to find suitable function forms for approximating the value function.

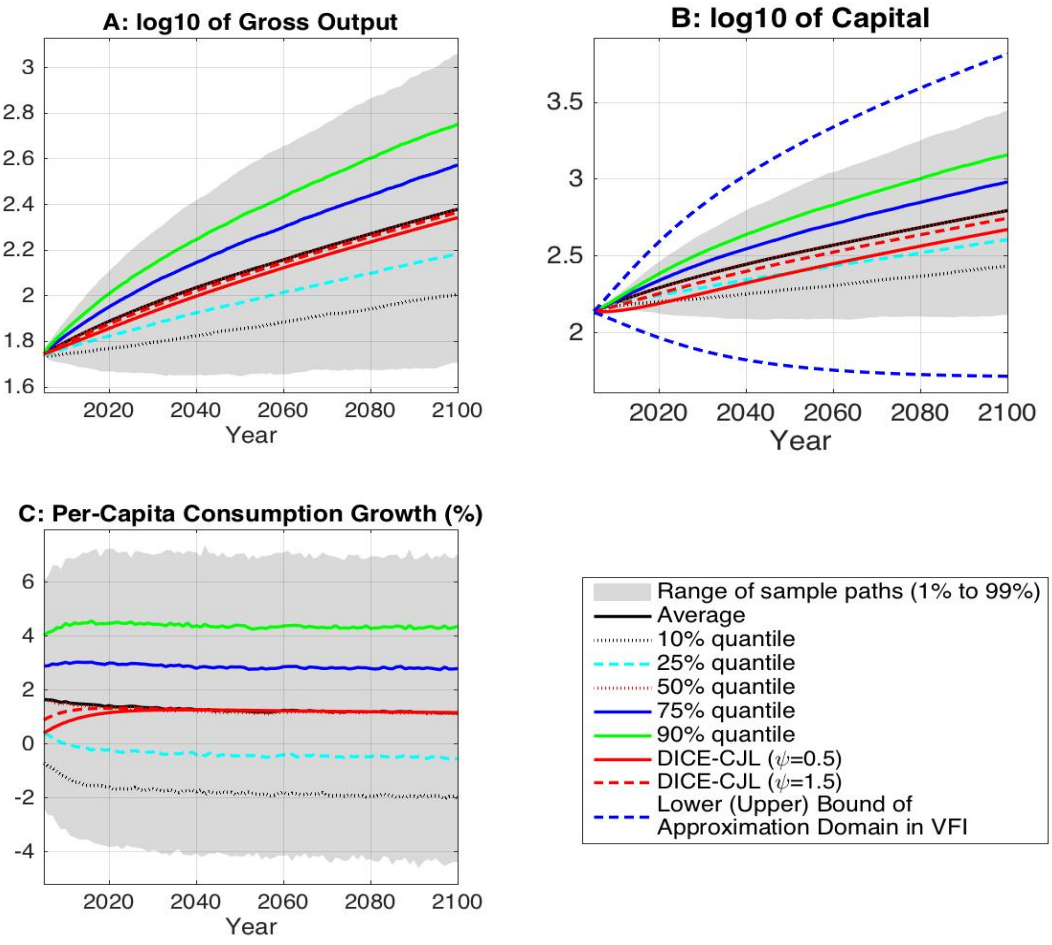

Figure A.2: Simulation results of the stochastic growth benchmark—economics

A.9 Stochastic Growth with Infinite Risk-aversion

For DSICE with stochastic growth only, the Bellman equation is

$$V_t(\mathbf{S}) = \max_{C,\mu} \left\{ u(C_t, L_t) + \beta \left[\mathbb{E}_t \left\{ \left(V_{t+1}(\mathbf{S}_+) \right)^{\frac{1-\gamma}{1-1/\psi}} \right\} \right]^{\frac{1-1/\psi}{1-\gamma}} \right\},$$

where $\mathbf{S} = (\mathbf{x}, \zeta, \chi)$ with (ζ, χ) representing the discretized long-tun risk of economic growth defined in Appendix A.2. Let $\underline{\zeta_t}$ and $\underline{\chi_t}$ be the lower bounds of (ζ, χ) at time t—that is, $\underline{\zeta_t} = \exp\left(-3\sqrt{\Delta_t}\right)$

and $\underline{\chi_t} = -3\sqrt{\Upsilon_t}$ from Appendix A.2. When $\gamma = \infty$, the Bellman equation transforms into one of solving the class of max-min problems:

$$\begin{aligned} V_t(\mathbf{S}) &= \max_{C,\mu} \{u(C_t, L_t) + \beta \min \{V_{t+1}(\mathbf{S}_+) : \mathbb{P}((\zeta_+, \chi_+) \mid (\zeta, \chi)) > 0\}\} \\ &= \max_{C,\mu} \Bigg\{ u(C_t, L_t) + \beta V_{t+1} \Bigg(\mathbf{x}_+, \min \{\zeta_+ : \mathbb{P}(\zeta_+ \mid (\zeta, \chi)) > 0\}, \\ &\qquad\qquad \min \{\chi_+ : \mathbb{P}(\chi_+ \mid (\zeta, \chi)) > 0\} \Bigg) \Bigg\} \\ &= \max_{C,\mu} \left\{ u(C_t, L_t) + \beta V_{t+1} \left(\mathbf{x}_+, \underline{\zeta_{t+1}}, \underline{\chi_{t+1}} \right) \right\}, \end{aligned}$$

where the second equality holds because the value function V_{t+1} is increasing over ζ and χ, and the third equality holds because the support of our conditional probabilities includes the lower bounds in order to avoid the unbounded nature of the (log-)normal distribution. Thus, the problem is equivalent to solving the deterministic optimal control problem

$$\max \sum_{t=0}^{T} \beta^t u(C_t, L_t)$$

with the production function

$$f(K, L, \underline{A_t}) = \underline{A_t} K^{\alpha} L^{1-\alpha},$$

where $\underline{A_t} = A_t \underline{\zeta_t}$ is the smallest productivity at time t under the discretized long-run risk process.

A.10 Impact of IES and Productivity Growth in Deterministic Growth Cases

A higher IES means a lower desire for consumption smoothness over time, implying less consumption and higher total investment (a sum of capital investment I_t and mitigation expenditure Ψ_t) in the current period if both average short-term economic growth and average long-term economic growth are positive, or more consumption and less total investment in the current period if both average short-term economic growth and average long-term economic growth are negative, where investment in mitigation is reflected by SCC. A higher productivity growth implies that less investment today can yield the same level of future consumption. Consequently, a higher productivity growth implies more consumption and lower total investment in the current period. Since capital investment

makes up a major portion of total investment (around 99.7%), whether total investment increases or decreases capital investment will have the same direction, which is opposite to the direction of consumption. We validated these economic intuitions in Tables A.11 and A.12, which report consumption (C_0) and capital investment (I_0) in 2005, respectively, for a variety of IES (ψ) and initial productivity growth (α_1), for the deterministic growth case.[35] Note that when $\alpha_1 = -0.01$, a higher IES implies more consumption and less capital investment in 2005, because both average short-term economic growth and average long-term economic growth are negative; when $\alpha_1 \geq 0$, a higher IES implies less consumption and higher capital investment in 2005, because both average short-term economic growth and average long-term economic growth are positive; when $\alpha_1 = -0.002$, a higher IES still implies less consumption and higher capital investment in 2005, because average short-term economic growth is positive although average long-term economic growth is negative, as capital investment pays more weights in short-term welfare due to the 10% depreciation rate on capital.

IES (ψ)	Initial productivity growth (α_1)					
	-0.01	-0.002	0	0.002	0.005	0.0092 (default)
0.5	36.8	39.2	39.8	40.3	41.1	42.1
0.7	37.5	39.2	39.6	40.0	40.6	41.3
0.9	37.8	39.1	39.4	39.7	40.2	40.8
1	37.9	39.1	39.3	39.6	40.0	40.6
1.1	38.0	39.0	39.3	39.5	39.9	40.4
1.5	38.2	38.9	39.0	39.2	39.4	39.7
2	38.3	38.7	38.8	38.9	39.0	39.2

Table A.11: Initial consumption (trillions of USD) for parameters of IES (ψ) and 2005 productivity growth (α_1) for the deterministic growth case

[35] When $\psi = 1$, the utility $u(C_t, L_t)$ becomes $\log(C_t/L_t)L_t$.

IES (ψ)	Initial productivity growth (α_1)					
	-0.01	-0.002	0	0.002	0.005	0.0092 (default)
0.5	18.6	16.3	15.8	15.2	14.5	13.5
0.7	18.1	16.4	16.0	15.6	15.0	14.2
0.9	17.8	16.5	16.1	15.8	15.4	14.8
1	17.6	16.5	16.2	15.9	15.5	15.0
1.1	17.6	16.5	16.3	16.0	15.7	15.2
1.5	17.3	16.7	16.5	16.4	16.1	15.8
2	17.2	16.9	16.8	16.7	16.5	16.3

Table A.12: Initial capital investment (trillions of USD) for parameters of IES (ψ) and 2005 productivity growth (α_1) for the deterministic growth case

However, the direction of mitigation investment (a small portion of total investment, around 0.3 percent only) will depend on a combination of IES and productivity growth. Table A.13 reports the 2005 SCC for the deterministic growth case with various IES and 2005 productivity growth rates. Our analysis with the deterministic version of the model confirms the findings in Azar and Sterner (1996)—that economic growth has a substantial effect on SCC.

IES (ψ)	Initial productivity growth α_1					
	-0.01	-0.002	0	0.002	0.005	0.0092 (default)
0.5	175	73	63	55	46	37
0.7	85	67	64	60	56	51
0.9	64	64	64	64	64	64
1.0	58	63	64	66	67	70
1.1	54	62	65	67	70	75
1.5	46	60	65	70	80	94
2.0	41	59	66	73	87	111

Table A.13: 2005 SCC for parameters of ψ and α_1 for the deterministic growth case

At first we see that when the productivity growth rate is positive, SCC is increasing over IES, but when the 2005 productivity growth rate is negative (implying the whole path of productivity growth is negative from our definition of productivity A_t), SCC is decreasing over IES. This happens because negative productivity growth corresponds to negative average long-term consumption growth g_c (although average short-term growth could be positive), and from Ramsey's growth theory we know that the internal rate of return (interest rate) is proportional to g_c/ψ, so a higher IES (ψ) implies a smaller internal rate of return when $g_c > 0$, but a larger internal rate of return when $g_c < 0$; moreover, SCC is the present value of future damage from one additional ton of carbon emitted, so

a larger internal rate of return implies a smaller SCC.

We also see that when $\psi < 0.9$, SCC is decreasing over 2005 productivity growth, but when $\psi > 0.9$, SCC is increasing. This happens because an increase in productivity growth leads to an increase in long-term consumption growth g_c and then an increase in the internal rate of return, but higher productivity growth also implies more carbon emissions (as carbon emissions E_t is linearly dependent on productivity). Since SCC is the present value of future damage from one additional ton of carbon emitted, the direction of SCC will depend on which factor—higher damage or a higher internal rate of return—plays a more important role. From Ramsey's growth theory, a smaller ψ implies a faster speed of the increase in the internal rate of return; so when ψ is small (here $\psi < 0.9$), the increase in the internal rate of return will overwhelm the increase in damage, implying a smaller SCC for higher 2005 productivity growth. But when ψ is large (here $\psi > 0.9$), the speed of the increase in the internal rate of return is overwhelmed by the increase in damage, implying a larger SCC for higher 2005 productivity growth, and when $\psi = 0.9$ the increase in the internal rate of return is completely balanced by the increase in damage, so it has the constant SCC, \$64/tC, for all 2005 productivity growth rates.