# A Deep Learning Analysis of Climate Change, Innovation, and Uncertainty*

Michael Barnett
Arizona State University

William Brock
University of Wisconsin

Lars Peter Hansen
University of Chicago

Ruimeng Hu
University of California Santa Barbara

Joseph Huang
University of Pennsylvania

**Abstract**

We study optimal carbon-neutral transition policy in response to climate change and climate-economic model uncertainty. To solve this model, we develop a neural-network-based computational algorithm that incorporates pseudo-states variables to derive accurate global solutions for dynamic models which are relatively high-dimensional, and where endogenous nonlinearities from model uncertainty aversion and jump processes are first-order considerations, as in our setting. Our social planner places substantial value on investment in R&D and green capital. However, the significant "social pay-off" of green technological innovation means only the robustly optimal R&D investment choice is significantly augmented by aversion to the various climate-model uncertainty channels.

*We are grateful for valuable suggestions and comments from Fahiz Baba-Yara, Victor Duarte, Harrison Hong, Felix Kubler, Maxime Sauzet, Simon Scheidegger, Jose Scheinkman, Michael Sockin, and Rebecca Willett. We also thank participants at the 2024 Midwest Finance Association Annual Meeting, the 2024 Blue Collar Working Group on Robustness and Uncertainty at the University of Chicago, the 2024 Society for Financial Studies Cavalcade Conference, the 2024 Western Finance Association Annual Meeting, the 2024 Stanford Institute for Theoretical Economics (SITE) Climate Finance and Banking Workshop, the 19th Annual Olin Finance Conference at WashU, the 2023 Conference of the Society for the Advancement of Economic Theory, the MFR Program-Bocconi University Conference on Decision-Making Under Uncertainty in Climate and Macroeconomics, and the "Assessing the Economic and Environmental Consequences of Climate Change: Incorporating Uncertainty and Quantifying Its Importance" Workshop at the UChicago Institute on Mathematical and Statistical Innovation (IMSI), which is supported by the National Science Foundation (Grant No. DMS-1929348).

# 1 Introduction

The characterization of optimal quantities, pricing outcomes, and social welfare implications under specified optimality criteria are the core focus of theoretical economic and financial market modeling. The outcomes from these modeling frameworks are often the emphasis for academic analysis and exploration, but such results can also be immensely consequential for government decision-making and policy action. As the complexity of modeling frameworks and policy questions increase, the limits of existing methodologies to provide sufficiently robust and accurate numerical solutions become ever more apparent. In this paper, we characterize the optimal carbon-neutral transition policy in response to climate change and climate-economic model uncertainty using a dynamic, general equilibrium model with multiple capital stock types. To solve this model, we construct a deep learning algorithm designed to solve economic and financial models that incorporate model features that have proven to be notably problematic for existing computational methods. In particular, we focus on continuous-time, dynamic general equilibrium models with high-dimensional state spaces governed by jump-diffusion processes that are subject to model uncertainty concerns, which are the type of model features that are central to our climate-economics setting.

The potential consequences of climate change are becoming increasingly apparent, with the the Sixth Assessment Report (AR6) of the Intergovernmental Panel on Climate Change (IPCC) stating that "[i]t is unequivocal that human influence has warmed the atmosphere, ocean and land" and that anthropogenic climate change "has caused widespread adverse impacts and related losses and damages to nature and people, beyond natural climate variability."[1] Given these concerns, policymakers and organizations are increasingly focusing on the necessity and feasibility of a transition to a carbon-neutral economy. Reports from the OECD, the IEA, the White House, McKinsey & Co., and Princeton University's High Meadows Environmental Institute[2], among many others, emphasize the need to heavily invest in cleaner production methods to reduce current carbon emissions, as well as the need to devote significant resources to R&D for developing new green technology to prevent future carbon emissions. These conclusions are based largely on scenario analysis aimed at achieving climate policy goals such as the 1.5 $C^{\circ}$ GMT temperature anomaly ceiling proposed by the 2015 Paris Climate Agreement. The implementation of a socially efficient carbon-neutral transition, taking into account the various frictions, risks, uncertainties, and endogenous feedbacks and responses related to climate change and climate policy action, is the key economic question that we address in this paper.

[1] See Masson-Delmotte et al. (2021a) and Pörtner et al. (2022).

[2] See OECD (2019), Bouckaert et al. (2021), White House (2021), McKinsey Global Initiative (2022), and Jenkins et al. (2021).

In this paper, we develop and solve a dynamic general equilibrium framework with three types of capital: "dirty" capital that generates carbon emissions when used in production, "green" capital that produces no emissions but is initially less productive than dirty capital, and "knowledge" capital that increases the likelihood of a technology shock that augments the productivity of green capital. By focusing on capital stocks, our model incorporates dynamic features that play an important role in the transition to carbon-neutrality. First, the model framework leads to an emissions pathway that is "sticky" or persistent because dirty capital depreciates gradually and is costly to disinvest. Second, convex adjustment costs related to investment capital mean that the accumulation of new green capital can take significant time. Finally, the arrival of improved green technology depends on the stock of knowledge capital in the economy, which requires R&D investment across time and substitutes resources away from consumption, as well as from other types of investment.

Two important considerations in our analysis are the potential for thresholds or jumps in the climate damage and green innovation processes, as well as the substantial uncertainty about the central mechanisms in our model. We assume that the severity of economic damages from climate change is revealed via a stochastic jump process that is dependent upon the global mean temperature anomaly, and that the realization of a green technological innovation that enhances the productivity of the green capital occurs via a jump process that depends upon the level of knowledge capital accumulated through R&D investment. The uncertainty in our pertains to carbon-climate dynamics, or the mapping from carbon emissions into atmospheric carbon into temperature changes; the severity of the economic damage functions, or the negative impact on output due to changes in atmospheric temperature; and the technological innovation process, or the likelihood of a technology shock that augments the productivity of green capital via R&D investment. While each of these model components allows for risk in the form of stochastic realizations, we are interested in exploring alternative forms of uncertainty. In particular, we focus on the possibility that a given model is misspecified in a meaningful and unknown way. We explicitly incorporate these uncertainty considerations into the social planner's decision problem by applying tools from dynamic decision theory. As a result, our analysis highlights the endogenous interdependencies and feedback effects that substantially impact optimal policy choices and social valuations coming from uncertainty aversion, with particular concerns related to shifts in climate damage severity and green technology.

Our general equilibrium framework with multiple endogenous state variables requires solving relatively high-dimensional PDEs with significant non-linearities due to the model uncertainty concerns and stochastic jump processes related to technological change and climate damages. As a result, our analysis requires computational methods that provide global

solutions to accurately characterize the endogenous optimal policies and the social valuations of interest. Standard finite difference methods commonly used in economics and finance are not well equipped for such problems. We, therefore, develop and implement an algorithm using deep neural network methods, augmenting the deep Galerkin method-policy improvement algorithm setting (DGM-PIA) by incorporating key model parameters as pseudo-state variables. While using parameters as pseudo-state variables has been done in other settings[3], to the best of our knowledge, we are the first to incorporate this feature into the DGM-PIA setting to improve the algorithm's solution accuracy when confronting nonlinearities from jump processes and model uncertainty aversion. The use of pseudo-states is critical for allowing the algorithm to accurately learn the monotonicity and other features related to the nonlinearity in our framework arising from model uncertainty aversion and stochastic jumps. As a result, the algorithm is able to handle multiple non-stationary, endogenous state variables with considerable non-linearity in a continuous-time, infinite horizon setting. The algorithm provides a toolbox that expands the ability of researchers to address questions not only in climate-economics, but more broadly in economics, finance, dynamic decision theory, and stochastic optimal control, where numerical solutions are necessary due to the "curse of dimensionality" and the complexities associated with heterogeneity, uncertainty, and nonlinearity in the model framework.

## 1.1 Climate Economics Literature

Our paper builds on and contributes to a number of important areas of research across economics, finance, geoscience, and applied mathematics. The implications of anthropogenic emissions have been a central focus of geoscientists for many decades, beginning with the seminal work of Arrhenius (1896). Recent work has focused on characterizing the dynamic relationship between carbon emissions, atmospheric carbon concentration, and temperature change via pulse experiments (Joos et al. 2013, Geoffroy et al. 2013, Eby et al. 2009), deriving simplified emulators or approximations of these complex relationships for the use of policymakers (Matthews et al. 2009, Pierrehumbert 2014, MacDougall et al. 2017), as well as quantifying the dynamic and stochastic features of carbon-climate dynamics (Ricke and Caldeira 2014, Palmer and Stevens 2019). These components are critical inputs into our model framework and uncertainty analysis.

The climate economics literature has given substantial focus to estimating the economic consequences of climate change and deriving a Social Cost of Carbon (SCC), or the cost to social welfare of emitting an additional amount of carbon emissions into the atmosphere.

[3]See Norets (2012), Duarte (2018), and Friedl et al. (2023) for examples of using pseudo-states in neural network frameworks related to structural model estimation and model sensitivity analysis.

Economists have estimated the economic costs generated by observed climate change for numerous economic sectors, regions, and dimensions of the economy (Dell et al. 2012, Hsiang et al. 2017, Colacito et al. 2019, Allen et al. 2019, Carleton and Greenstone 2021). Theoretical modeling and integrated assessment model analysis to derive SCC valuations have considered for numerous economic mechanisms (Golosov et al. 2014, Acemoglu et al. 2016, Nordhaus 2018, Cai et al. 2017, Cai and Lontzek 2019) and proposed climate damage function approximations and frameworks (Lenton et al. 2008, Weitzman 2012, Cai et al. 2015, Drijfhout et al. 2015, Ritchie et al. 2021).

Recent work has begun to examine important dimensions related to interconnected climate change and economic model uncertainty (Olson et al. 2012, Lemoine and Traeger 2014, Hassler et al. 2018, Nordhaus 2018, Dietz and Venmans 2019, Barnett et al. 2020, Rudik 2020, Barnett et al. 2022, Barnett 2023). Importantly, many of these studies of uncertainty have exploited the powerful toolset developed in dynamic decision theory (Anderson et al. 2003, Maccheroni et al. 2006, Hansen and Sargent 2007, Klibanoff et al. 2009, Hansen and Miao 2018, Barnett et al. 2020, Cerreia-Vioglio et al. 2021), allowing researchers to account for model uncertainty explicitly in the decision-maker's problem within the model.

This paper also links to an important area of macroeconomic research. This includes the foundational work on endogenous economic growth (Brock and Mirman 1972, Baumol 1986, Lucas Jr 1988, Romer 1990), as well as recent analysis of the social and private benefits of innovation (Bloom et al. 2019, Lucking et al. 2019), and considerations for the transition to a green economy (Acemoglu et al. 2012, 2016, Jaakkola and Van der Ploeg 2019). In addition, the connection between macroeconomics and asset pricing in production-based asset pricing (Brock 1982, Cochrane 1991, Jermann 1998), as well as the importance links to economic growth and innovation (Papanikolaou 2011, Kogan and Papanikolaou 2014, Kung and Schmid 2015) have important implications for the social valuations we derive in our climate-economics-innovation linked framework.

## 1.2 Deep Learning Literature

In recent years, deep learning algorithms, built on the neural network's remarkable ability to represent and approximate high-dimensional functions and efficient gradient descent optimizers, have been very successful in many areas, ranging from computer vision and speech recognition to scientific computing (see, e.g. Carleo and Troyer (2017), Gao and Duan (2017), Han et al. (2018), Zhang et al. (2018), Han et al. (2019)). The climate change problem considered in this paper aims to study how choices for investment in clean and dirty capital, and R&D for technological innovation, are determined when facing uncertainties, which, math-

ematically, is a stochastic control problem. Along this direction, deep learning algorithms are roughly categorized into: direct parameterization, partial differential equations (PDEs), and forward-backward stochastic differential equations (FBSDEs) approach.

In this first category, the seminal work in high-dimension was proposed by Han and E (2016), which solves a global-in-time minimization problem in accordance with the utility of the control problem. Later, Bachouch et al. (2021) extended to the local-in-time approach combined with dynamic programming techniques. Han and Hu (2021) solve the control problem with aftereffects, modeled by stochastic differential delayed equations, using the same spirit of Han and E (2016) but with advanced network architectures, and Carmona and Laurière (2022) extended to mean-field control problems.

In the PDE approach, the stochastic control problem will first be reformulated into a Hamilton-Jacobi-Bellman (HJB) PDE using the dynamic programming principle, and then solved by deep learning algorithms. For generic PDEs, Sirignano and Spiliopoulos (2018) proposed the deep Galerkin method (DGM), and Raissi et al. (2019) proposed the physics-informed neural networks (PINNs), roughly at the same time. Both approximate solutions to PDEs by training neural networks to minimize the residuals coming from the initial conditions, the boundary conditions, and the PDE operators. Saporito and Zhang (2021) followed this idea and solved path-dependent PDEs using recurrent neural networks. Later, Al-Aradi et al. (2022) and Duarte et al. (2023) extended the DGM to deal with HJB equations in their unsimplified primal form and solved for the value function and the optimal control simultaneously by characterizing both as deep neural networks.

For semi-linear parabolic PDEs, which correspond to stochastic control problems with uncontrolled volatility, E et al. (2017) and Han et al. (2018) proposed a deep BSDE solver which, to the best of our knowledge, is the first work in this area and has inspired much follow-up work. They recast the solution of a semi-linear parabolic PDE into a global-in-time optimization problem based on the associated BSDEs via the non-linear Feynman Kac formula and variational form. Huré et al. (2020) dealt with the same associated BSDE but obtained the resolution by solving backward in time using a sequence of small optimization problems and termed it as deep learning backward dynamic programming (DBDP). Another work focusing on semi-linear PDE is the deep splitting method proposed by Beck et al. (2021), where the partial differential operators are split into the linear part and the nonlinear part, with the nonlinear part propagating first by freezing the solution and the linear part then being taken care of by an approximate Feynman-Kac representation. When the control appears in the state dynamics' diffusion coefficient, this leads to a fully nonlinear PDE. In this direction, in the same spirit of E et al. (2017) and Han et al. (2018), Beck et al. (2019) proposed a DL algorithm by solving the corresponding second-order BSDE, and Pham et al.

(2021) extended the DBDP idea in Huré et al. (2020) by further approximating the Hessian matrix using auto differentiation of first derivatives' neural networks.

In the FBSDE approach, the control problem will first be reformulated into a coupled forward-backward system using the stochastic Pontryagin maximum principle. The system is, in general, hard to solve numerically, let alone in high dimensions, due to its coupled nature and opposing directions: one with an initial condition running forward in time and one with a terminal condition running backward in time. Han and Long (2020) solved the fully-coupled FBSDE by extending the deep BSDE solver with convergence analysis subject to neural networks' universal approximation, and Ji et al. (2020) further extended the results with the Picard iteration method.

Recently, attention has also been drawn to stochastic control problems with jumps. The problem is more involved as the PDE becomes partial integro-differential, thus non-local; and the FBSDE system now contains a Lévy process. For such problems, Boussange et al. (2022) solved the associated non-local PDE by extending the deep splitting method in Beck et al. (2021) and multilevel Picard approximation method in Hutzenthaler et al. (2019); Castro (2022) solved the associated forward-backward system by extending the DBDP method in Huré et al. (2020); and Gnoatto et al. (2022) solved the same system in the same spirit of the deep BSDE solver, proposed in E et al. (2017) and Han et al. (2018). These settings all focus on the semi-linear problem, where the control only appears in the drift coefficient. Very recently, Lu et al. (2024) proposed an actor-critic reinforcement learning algorithm to address the fully nonlinear control problem with jumps.

When using neural networks to parameterize the quantity of interest, usually the value function or the control policy, people may want to incorporate domain knowledge, such as guaranteed monotonicity or convexity of the learned function with respect to some of the input variables. There have been fruitful studies on how to build network structures in order to fulfill such requirements, see for instance, Sill (1997), Gupta et al. (2016), Wehenkel and Louppe (2019) and Runje and Shankaranarayana (2022).

# 2 A Multi-Capital Climate-Economics Model

We now outline the example economy used to study optimal climate policy related to a carbon-neutral transition. The model incorporates components related to climate change dynamics, economic damages from climate change, economic production technologies, technological innovation through R&D investment, and preferences that include model uncertainty aversion. We outline each of these model pieces in what follows, and then derive the central theoretical results before presenting the numerical algorithm developed to derive our

model solution and the numerical solutions to the model using our computational algorithm.

## 2.1 Climate Dynamics

We begin with relatively simple climate dynamics based on the simplified approximations of long-run climate change outcomes developed by Matthews et al. (2009) and validated by Friedlingstein (2015). Specifically, recent geoscientific analysis has shown that the output of large-scale atmosphere-ocean general circulation models (AOGCM) can be approximated by a proportional relationship between temperature anomaly and cumulative carbon emissions:

$$\text{temperature anomaly} \quad \approx \underline{\text{climate sensitivity}}\ (\theta) \ \times \ \text{cumulative emissions.}$$

For our analysis, we use a stochastic variant of this relationship given as follows:

$$dY_t = \mathcal{E}_t \left( \bar{\theta} dt + \varsigma dW_t^y \right),$$

where $Y_t$ is the global mean temperature anomaly, in degrees Celsius, relative to the preindustrial level, $\mathcal{E}_t$ is global carbon emissions, $W_t$ is a vector Brownian motion process with filtration $\mathcal{F}_t$ and scalar components $W_t^y$, $W_t^d$, $W_t^g$ and $W_t^r$, $\varsigma$ is the volatility loading for temperature, and $\bar{\theta}$ is the average of the possible climate sensitivity proportionality constants $\theta(m)$ from a set of possible values denoted by $m = 1, \ldots, M$, with equal-weighted likelihood given to each model. Importantly, this long-run approximation implies a permanent impact of carbon emissions[4]. In addition, concerns about model uncertainty naturally arise in this setting given the variation in estimated $\theta(m)$ values, known as the transient climate response to cumulative emissions (TCRE) parameter, found across large-scale AOGCMs[5].

In addition, the inclusion of Brownian shocks serves a few purposes. First, it is consistent with the seminal work of Hasselmann (1976)[6], which started the field of stochastic climate modeling to characterize the fast/slow time scale separation of climate variability, and the more recent work of Palmer and Stevens (2019) as a means of improving predictability and approximating variation coming from chaotic models of climate dynamics. Second, Brownian shocks can be seen as a crude way of representing the unmodelled "complex interplay of multiple physical and biogeochemical processes" underlying the TCRE proportionality result

[4] A reasonable approximation in our view given the estimated lifespan of atmospheric carbon emissions found by Pierrehumbert (2014) and others.

[5] Matthews et al. (2009), Friedlingstein (2015), and Masson-Delmotte et al. (2021b) all provide insightful representations and details related to this climate sensitivity uncertainty result.

[6] Franzke et al. (2022) surveys Hasselman's work, as well as the work of Syukuro Manabe and Giorgi Parisi, who were jointly awarded the 2021 Nobel Prize in Physics for their seminal work on modeling complex physical systems, including direct application to modeling climate dynamics.

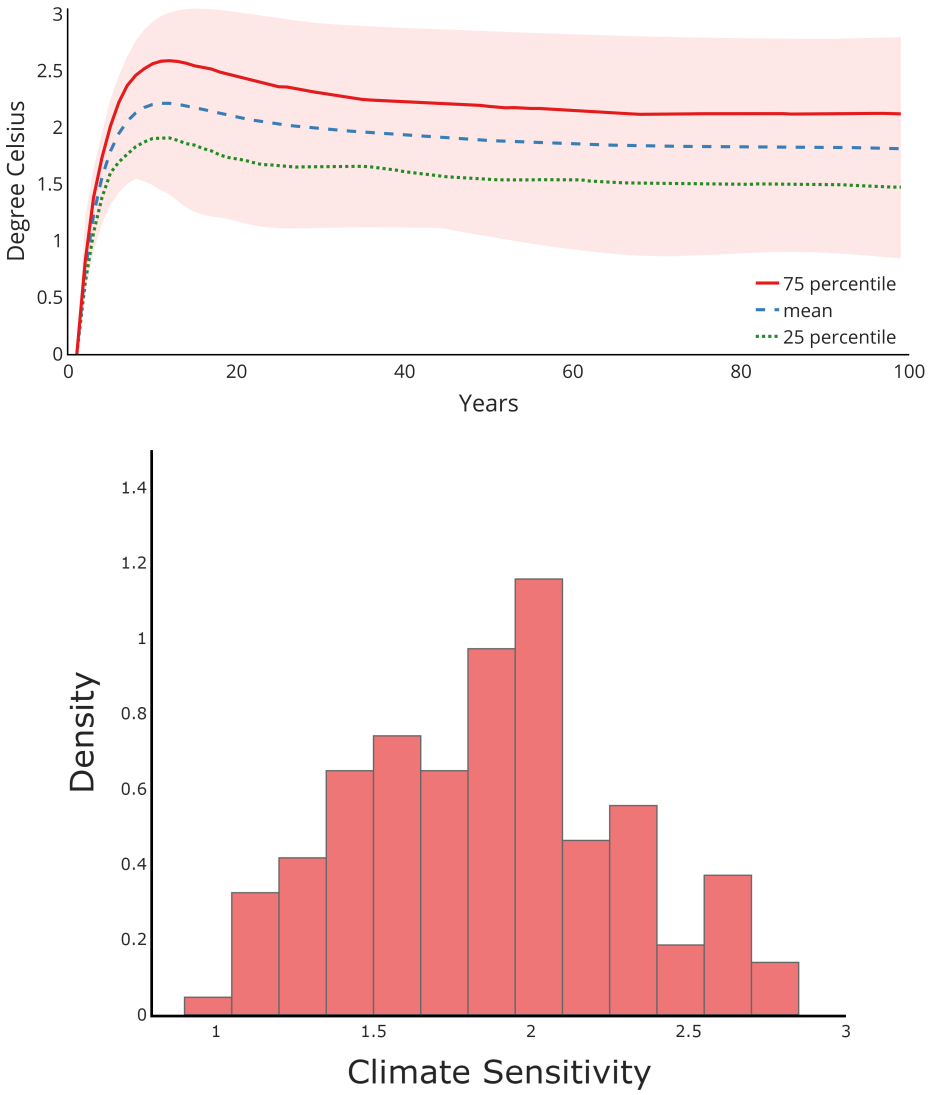



Figure 1: Implied TCRE coefficients across time and models, capturing carbon and temperature uncertainty. The top figure shows the percentiles for temperature responses to emission pulses for all carbon and temperature model combinations. The bottom figure is a histogram for the exponentially weighted average responses of temperature to an emissions pulse from 144 different models using a rate $\delta = .01$.

that is "not amenable to a simple physical explanation" as recently shown by Gillett (2023). Third, the use of Brownian shocks means that the true climate model is "disguised" in the statistical sense that a small set of observations do not immediately reveal the true model, allowing for a meaningful uncertainty analysis.

**Remark 2.1.** *While the proportionality relationship used here is based on a relationship of carbon emissions to temperature change that is born out at a time horizon of 10 years and beyond, Dietz and Venmans (2019), Barnett et al. (2022), Barnett (2023) and others have*

*shown that this relationship is well suited for frequencies as short as one year. We follow Barnett et al. (2022) and Barnett et al. (2023) and use pulse experiment results of Joos et al. (2013) and Geoffroy et al. (2013) to build the set of climate sensitivity models for our analysis.*[7] *Figure 1 shows the dynamic pathways and the implied TCRE parameters for the 144 models produced from this exercise. We can see from the pathways that the temperature response peaks around 10 years, and flattens out thereafter. The histogram implied TCRE parameters, also shown in Figure 1, are quite dispersed, with values ranging from around 1 to just below 3, with the average value being* 1.86.

## 2.2 Climate Damages

Our model of economic damages from climate change, or climate damages, follows a piecewise log quadratic specification where a potential change in the curvature of the damage function is triggered by a jump process realization, just as in Barnett et al. (2023). The climate damages, denoted $N_t$, are assumed to impact capital, output, investments, and consumption in a proportional manner. There is a stochastic jump process that determines dynamically which value of the damage function curvature $\lambda_3(\ell)$ will be realized and when that potential change will be realized, denoted by a temperature anomaly value $\hat{Y}$. Specifically, the jump process has $L$ absorbing states, where each state $\ell$ corresponds to a value of $\lambda_3(\ell)$. Before the realization of the jump, each $\lambda_3(\ell)$ value has a an equal-weighted likelihood of occurring. When the jump is realized, the value of $\lambda_3(\ell)$ is revealed and the true damage function curvature becomes known to the social planner in the model. The likelihood of a damage function jump is determined by a jump intensity $\mathcal{J}_n^\ell(y)$ that is increasing in the endogenous temperature anomaly state variable $y$ and is localized to the range of temperature values $[\underline{y}, \overline{y}]$. There is uncertainty associated both with the probability distribution on possible $\lambda_3(\ell)$ model realizations, as well as the probability of when a damage jump will occur in the $[\underline{y}, \overline{y}]$. The climate damage function $N_t$ is given in logarithmic form

$$\log N(y) = \begin{cases} \lambda_1 y + \frac{\lambda_2}{2} y^2 & 0 \leqslant y \leqslant \hat{y} \\ \lambda_1 y + \frac{\lambda_2}{2}\hat{y}^2 + \frac{\lambda_2}{2}(y - \hat{y} + \bar{y})^2 + \frac{\lambda_3(\ell)}{2}(y - \hat{y})^2 - \frac{\lambda_2}{2}\lambda_2(\bar{y})^2 & y \geqslant \hat{y}. \end{cases}$$

Note that this specification depends on the temperature value when the jump occurs, $\hat{Y}$,

[7] Joos et al. (2013) examines carbon dynamics variation and uncertainty based on responses of atmospheric carbon concentration to emission pulses for several alternative Earth System models. Geoffroy et al. (2013) provides temperature dynamics variation and uncertainty based on approximate dynamics relating the log of atmospheric carbon to future temperature, following the Arrhenius equation. We take 9 atmospheric carbon pulse responses as inputs into 16 temperature dynamics approximations to derive our set of carbon-climate TCRE models.

such that the earlier it occurs the more severe the damages from climate change are.

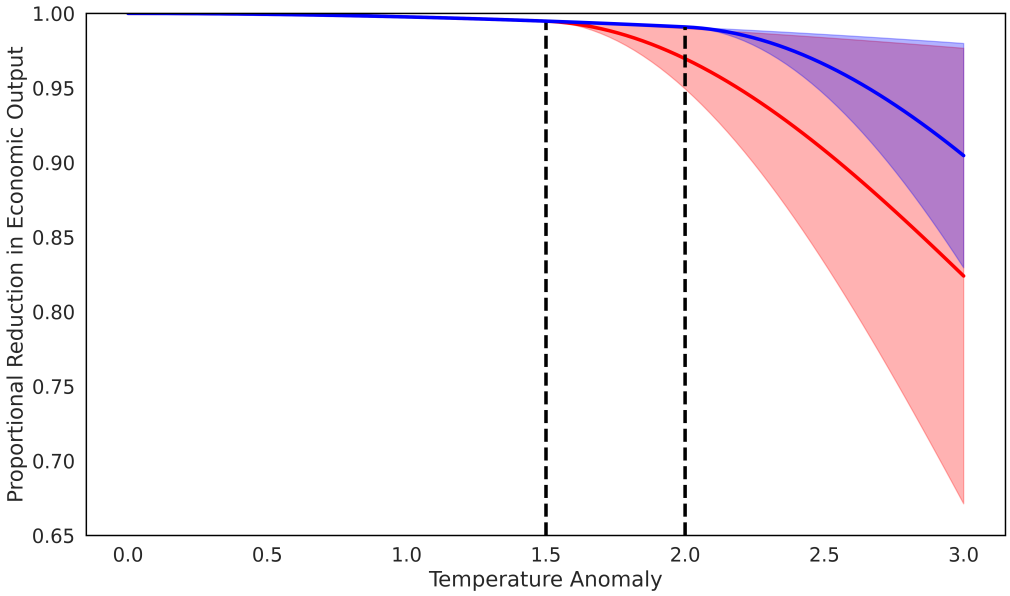



Figure 2: Range of possible damage functions for two cases with different jump thresholds. The solid lines show the average values, and the shaded regions give the range of possible values for $\exp(-n)$, which measures the proportional reduction of the productive capacity of the economy. The blue line and region show the damage function curvature when the jump occurs at $Y_t = \bar{y} = 2.0$. The red line and region show the damage function curvature when the jump occurs at $Y_t = \underline{y} = 1.5$. The black dashed lines indicate the values of $Y_t$ for the upper and lower jump thresholds for the temperature anomalies.

For our main setting, we assume $[\underline{y}, \bar{y}] = [1.5, 2.5]$, consistent with the critical temperature thresholds used by the IPCC and the 2015 Paris Climate Agreement. Figure 2 shows the implied climate damages across all $\lambda_3(\ell)$ values for the jump realization occurring at the temperature anomaly values of $\tilde{y} = \underline{y} = 1.5$ and $\tilde{y} = \bar{y} = 2$. The red shaded region shows the range of possible damage outcomes across the possible values of $\lambda_3(\ell)$, and the black line gives the mean value. While there is significant discussion about the likelihood of tipping points or thresholds at the global scale captured by discrete jumps or shifts in climate damages[8], our setting is instead a smooth shift where the planner realizes whether the damage function is more convex than previously known. In this sense, our "tipping points" should be seen as an endogenous informational event that reveals the potential severity of

[8]See Brook et al. (2013) and Levitan (2013) for examples in the climate literature discussing the plausibility of global scale tipping points. Drijfhout et al. (2015) and Armstrong McKay et al. (2022) examine regional tipping points and their potential regional and global consequences.

economic damages from additional future climate change. This structure is distinct from much of the existing literature which there is either static information or permanently unresolved uncertainty about climate damages. We therefore view our framework as a valuable tool for characterizing the possibility and uncertainty of potentially severe, nonlinear climate damage outcomes that are highly relevant for the optimal policy decisions of a social planner confronting climate change.

## 2.3 Production

For the economic component of the model, we assume there are two sectors producing perfectly substitutable consumption goods using their own AK technologies[9]:

$$F_j(K_t^j) \quad = \quad A_j K_t^j, \quad j \in \{d, g\}$$

Each sector has its own capital stock that evolves with logarithmic adjustment costs and Brownian shocks as follows:

$$dK_t^j = K_t^j[\alpha_j + \Gamma_j \log(1 + \theta_j i_t^j)]dt + \sigma_j K_t^j dW_t^j, \quad j \in \{d, g\}$$

For computational tractability, we redefine the state variables characterizing the capital stocks in the model and use $\log K = \log(K^d + K^g)$ and $Z = \frac{K^g}{K^d + K^g}$. By Ito's lemma, we can characterize the evolution of our two new state variables. Log total capital evolves as follows

$$\begin{aligned} d\log K_t = (1 - Z_t)[\alpha_d + \Gamma_d \log(1 + \theta_d i_t^d)]dt + Z_t[\alpha_g + \Gamma_g \log(1 + \theta_g i_t^g)]dt \\ - \frac{1}{2}|\sigma_d(1 - Z_t)|^2 dt - \frac{1}{2}|\sigma_g Z_t|^2 dt + (1 - Z_t)\sigma_d dW_t^d + Z_t \sigma_g dW_t^g, \end{aligned}$$

while the green-capital-to-total-capital ratio evolution is given by

$$\begin{aligned} dZ_t = Z_t(1 - Z_t)\{[\alpha_g + \Gamma_g \log(1 + \theta_g i_t^g)] - [\alpha_d + \Gamma_d \log(1 + \theta_d i_t^d)]\}dt \\ + Z_t(1 - Z_t)\{(1 - Z_t)|\sigma_d|^2 - Z_t|\sigma_g|^2\}dt + Z_t(1 - Z_t)\sigma_g dW_t^g - Z_t(1 - Z_t)\sigma_d dW_t^d. \end{aligned}$$

There are two key differences across the different types of production technologies in

[9]While the AK model has in part been rejected by Bernanke and Gürkaynak (2001) and others based on cross-equation restrictions relating the output-capital ratio to the sensitivity of growth to the saving rate, they also note that "the key prediction of the model that the saving rate (rate of capital accumulation) is important for explaining the growth as well as the level of per capita output seems to hold considerable validity." This latter implication is the relevant feature for our analysis. We therefore see the AK model as a reasonable approximation for our purposes, and leave refinements of the model using alternative production technologies for future research.

the economy. First, Sector d is the only sector that generates emissions, so that $E_t$ in our temperature evolution equation is given by

$$\mathcal{E}_t = \eta A_d K_t^d,$$

where $\eta$ is the emissions intensity parameter of Sector d production. Second, Sector d is initially more productive than Sector g in that $A_d > A_g$. While Sector d is initially more productive than Sector g ($A_d > A_g$), we assume that there is the potential for "green" technology innovation that would augment Sector g productivity. The arrival rate of the jump in Sector g productivity is an increasing function of the aggregate knowledge capital stock in the economy $R$. The evolution of the aggregate knowledge stock is given by

$$dR_t = -\zeta R_t dt + \psi_0 \left(I_t^r\right)^{\psi_1} \left(R_t\right)^{1-\psi_1} dt + R_t \sigma_r dW_t^r, \tag{1}$$

where $I^r$ is the level of R&D investment made to increase the knowledge capital stock, $\zeta$ is the decay rate of the knowledge stock, $\psi_0$ and $\psi_1$ capture the effectiveness of R&D investment, and $\sigma_r$ is the volatility associated with the knowledge capital stock which emulates exogenous stochastic information shocks about the future possibility of a technological innovation.

We allow for uncertainty about the realization of the technology shock. We assume a discrete set of possible realizations for the technology jump: and intermediate technology shift $A_g(\ell')$ and a terminal technology shift $A_g(\ell'')$. The technological innovation that determines the change in green capital productivity is realized through a stochastic jump process. The value $A_g''$ is an absorbing "breakthrough" state, while $A_g'$ is a transient intermediate state from which we can still realize a jump to the absorbing $A_g''$ "breakthrough" state. The probability of realizing $A_g(\ell'')$ from the pre-technology jump state if a technology jump occurs is $pi$, and thus the probability of realizing $A_g(\ell')$ from the pre-technology jump state if a technology jump occurs is $1 - \pi$. The probability of realizing $A_g(\ell'')$ from the intermediate-technology jump state if a technology jump occurs is one. The likelihood of a technological innovation depends upon the jump intensities $\mathcal{J}_g^{\ell'}(R)$ and $\mathcal{J}_g^{\ell''}(R)$, which are assumed to be increasing in the endogenous knowledge stock state variable $R$. We assume that $A_g < A_d$, $A_g(\ell') = A_d$, and $A_g(\ell') < A_g(\ell'')$. Uncertainty is about about the probability distribution and arrival rate of potential technological change realization in the $A_g$ state, only the arrival rate of a potential technological change realization in the $A_g'$ state, and there is no uncertainty about technological change in the $A_g''$ state. Thus, with uncertainty about the probability $\pi$ and the arrival rates of potential technology realizations, our structure allows for a broadly conceived uncertainty analysis of a relatively novel specification of "green" technological innovation.

### 2.4 Preferences

Finally, we assume that flow utility in the model is a log function over damaged aggregate consumption, where exponential-quadratic damages multiplicatively scale consumption. The utility function is therefore given by

$$U(\tilde{C}) = U(C/N) = \delta \log(C) - \delta \log N(Y),$$

where $\delta$ is the subjective discount rate, the two types of consumption goods are perfectly substitutable. Investment decisions are made by optimally dividing the output net of consumption across investment into the three types of capital in the economy (dirty production capital, green production capital, and knowledge capital), so that the market clearing final output constraint is given by the relationship

$$C = A_d K^d - i^d K^d + A_g K^g - i^g K^g - i^r (K^g + K^d),$$

where $i^d = I^d/K^d$ is the dirty investment-to-dirty capital ratio, $i^g = I^g/K^g$ is the green investment-to-green capital ratio, and $i^r = I^r/(K^d + K^g)$ is the R&D investment-to-total capital ratio.

## 3 Model Solution

With the model framework laid out, we can now turn to solving the Hamilton-Jacobi-Bellman equations that characterize the solution to the Social Planner's problem in our model. The solution to our model is a recursive Markovian equilibrium that solves the Hamilton-Jacobi-Bellman equation characterizing the planner's social welfare function, or value function. Therefore, the equilibrium solution is determined by optimal investment choices:

$$\{i_t^{g*}, i_t^{d*}, i_t^{r*}\}$$

that maximize the planner's discounted, lifetime expected utility as functions of the states

$$\{\log K_t, Z_t, Y_t, \log R_t\}.$$

These optimal controls must satisfy the evolution equations for the state variables, as well as the market clearing conditions given in the model set-up above and the first order conditions from the HJB equations we outline below. To arrive at the full model solution, we must derive solutions sequentially working from the post-technology jump, post-damage-function-jump

state back to the pre-technology jump, pre-damage-function-jump state, accounting for the different possible combinations of intermediate states, which are the pre-technology jump, post-damage-function-jump state and the post-technology jump, pre-damage-function-jump state. Given the complexity and scope of the various HJB equations we must solve, we omit these details from the main text and direct the reader to Appendices A – C for complete model and HJB equation details.

# 4 A Deep Learning Algorithm for Complex Problems

A solution to our model can only be found by computationally solving a relatively high-dimensional, nonlinear Hamilton-Jacobi-Bellman equation. However, standard finite-difference methods often used in economics and finance are not able to produce accurate solutions for such settings. Thus, a key contribution of this paper is to provide an algorithm that provides accurate global solutions for our model setting.

We begin by outlining the Hamilton-Jacobi-Bellman equation used to characterize the types of dynamic, recursive optimization problems we are interested in solving. We then detail the implications that jump processes and model uncertainty aversion have for our baseline HJB equation. With these pieces in place, we discuss the implications of high-dimensionality in the state space, further motivating the numerical algorithm needed. Finally, we explain the algorithm we construct to solve these types of models, accounting not only for the nonlinearities, but also the high dimensionality. In particular, we show how including "pseudo"-state variables significantly improves the accuracy of the global numerical solutions in such complex settings.

## 4.1 Standard HJB equation

For simplicity and completeness, we begin by outlining the set-up for a standard dynamic, recursive optimization problem. We consider a social planner's problem where the objective is to maximize lifetime, discounted expected utility in continuous time over an infinite horizon. Following the relatively standard heuristic derivation, omitted for conciseness, the social planner's problem can be written as an autonomous partial differential equation (PDE) known as a Hamilton-Jacobi-Bellman (HJB) equation, which takes the following form:

$$0 = -\delta V(x) + \sup_{\alpha} \left\{ U(x,\alpha) + \frac{\partial V}{\partial x'}(x)\mu(x,\alpha) + \frac{1}{2}\text{trace}\left[\sigma(x,\alpha)' \frac{\partial^2 V}{\partial x \partial x'}(x)\sigma(x,\alpha)\right]\right\}.$$

where $V(x)$ represents the value function capturing the optimized social welfare of the

social planner, $x$ and $\alpha$ are realizations of a vector of relevant state variables $X_t$, whose evolution is given by stochastic diffusion process with drift $\mu(x, \alpha)$ and volatility $\sigma(x, \alpha)$, and control choices $A_t$, $\delta$ is the subjective discount rate, $U(X_t, A_t)$ is the flow utility function. Various computational algorithms have been devised to derive viscosity solutions to such relatively standard PDEs when analytical solutions are not available. Our focus is on going beyond such settings as we outline below.

## 4.2 Nonlinearity from Jumps and Model Uncertainty

We are interested in augmenting the basic setting outlined above in ways that are central for climate-economics and climate-finance model frameworks, as well as other stochastic optimal control settings in economics, finance, and applied mathematics. The additional model components we focus on are jump processes and model uncertainty aversion, which introduce endogenous non-linearities that often rule out analytic solutions and many existing numerical methods. We outline the impact that these model features have below.

### 4.2.1 Stochastic Jump Processes

We begin with the introduction of stochastic jumps for the state vector $X_t$. We assume in our setting that there are $\ell$ discrete realizations of the jump process from a set $\mathcal{L}$, with a jump intensity function $\mathcal{J}^\ell(x) = \pi^\ell \mathcal{J}(x)$. Note $\pi^\ell$ is the prior probability of jump realization $\ell$ occurring if a jump takes place and $\mathcal{J}(x)$ is the state-dependent jump arrival rate. Allowing for jumps in our setting leads to the addition of the following term to our HJB equation

$$\sum_{\ell=1}^{\mathcal{L}} \mathcal{J}^\ell(x) \left[ V^\ell(x) - V(x) \right].$$

where $V^\ell(x)$ denotes the post-jump value function for jump realization $\ell$. Depending on the function form of the arrival rate and jump realizations, the jumps can add an additional endogenous, convexly nonlinear element to our HJB equation that complicates the ability to derive numerical solutions to our problem.

### 4.2.2 Brownian Misspecification

Next, we outline the impact of Brownian model misspecification. Following from Girsanov theory, diffusion misspecification comes in the form of a drift distortion related to a positive martingale with expectation equal to one which parameterizes changes in the probability measure. With this, we use robust preferences of the form explored by Hansen and Sargent

(2011) which augments the planners preferences and optimization problem to include a minimization choice based on a quadratic penalty term over $h$ coming from a local measure of relative entropy or Kullback-Leibler divergence to capture "model uncertainty aversion". The problem augments our baseline HJB equation as follows

$$\min_{h} \frac{\partial V}{\partial x'}(x)\left[\mu(x,\alpha)+\sigma(x,\alpha)h\right]+\frac{1}{2}\text{trace}\left[\sigma(x,\alpha)'\frac{\partial^2 V}{\partial x \partial x'}(x)\sigma(x,\alpha)\right]+\frac{\xi}{2}h'h$$

where $\xi$ is a scalar parameter capturing "model uncertainty aversion" and will be determined as in a robust Bayesian setting by considering implied ex-post probabilistic distortions generated by the choice of $\xi$. Optimizing over $h$ gives the optimal control

$$h^* = -\frac{1}{\xi}\sigma(x,\alpha)'\frac{\partial V}{\partial x}(x).$$

As with jumps, Brownian misspecification can add an additional endogenous nonlinear element to our HJB equation, depending on assumed functional forms, which also complicates the ability to derive numerical solutions to our problem.

### 4.2.3 Jump Misspecification

Finally, we highlight the impact of jump model misspecification. To incorporate this type of misspecification, we again follow the robust preferences structure of Hansen and Sargent (2011). We therefore introduce non-negative functions $g^\ell$ which alter the jump intensity to be of the form $\mathcal{J}^\ell(x)g^\ell(x)$. The planner's preferences and optimization problem are again augmented to include a minimization choice based on a local measure of relative entropy or Kullback-Leibler divergence over the $g^\ell(x)$ to capture "model uncertainty aversion" of this form. The problem augments our baseline HJB equation as follows

$$\min_{g^\ell \geqslant 0} \quad \sum_{\ell=1}^{L}\mathcal{J}^\ell(x)g^\ell(x)\left[V^\ell(x)-V(x)\right]+\xi\sum_{\ell=1}^{L}\mathcal{J}^\ell(x)\left[1-g^\ell(x)+g^\ell(x)\log g^\ell(x)\right]$$

where $\xi$ is our aforementioned a scalar parameter capturing "model uncertainty aversion". Optimizing over the $g^\ell(x)$ gives the optimal control

$$g^{\ell *}(x) = \exp\left(-\frac{1}{\xi}\left[V^\ell(x)-V(x)\right]\right).$$

Again, as in our previous two modifications, jump misspecification can add an additional endogenous nonlinear element to our HJB equation, depending on assumed functional forms.

Moreover, this nonlinear interacts with the existing nonlinear adjustment coming from the introduction of jumps even without misspecification concerns. Together, this additional component further complicates the ability to derive numerical solutions to our problem.

## 4.3 Augmented HJB equation

Our climate-economics setting incorporates each of these elements simultaneous, with each playing important parts of the considerations for economic damages from climate change, technological innovation and the carbon-neutral transition, and the model uncertainty associated with these elements, as well as the broader climate-economics dynamics underlying the modeling framework. Taken together, and plugging in the model uncertainty-associated minimizers, our baseline HJB equation is augmented to take the following form

$$0 = -\delta V(x) + \sup_{\alpha} \left\{ U(x,\alpha) + \frac{\partial V}{\partial x'}(x)\mu(x,\alpha) + \frac{1}{2}\text{trace}\left[\sigma(x,\alpha)'\frac{\partial^2 V}{\partial x \partial x'}(x)\sigma(x,\alpha)\right] \right.$$
$$\left. -\frac{1}{2\xi}\frac{\partial V}{\partial x'}(x)\sigma(x,\alpha)\sigma(x,\alpha)'\frac{\partial V}{\partial x}(x) \right\} + \xi \sum_{\ell}^{L} \mathcal{J}^{\ell}(x)\left[1 - \exp\left(-\frac{1}{\xi}\left[V^{\ell}(x) - V(x)\right]\right)\right]$$

where we have included jumps and model misspecification of bother forms, having plugged in the optimized distortions $h$ and $g^{\ell}(x)$. The challenge we face is constructing an algorithm that can accurately capture the nonlinearities resulting from each of these model features, as well as their interactions, in a dynamic climate-economics model setting.

## 4.4 High Dimensionality

The final component we must consider for our computational algorithm is confronting the "curse of dimensionality", or that the complexity of deriving numerical solutions exponentially increases in the number of dimensions. The integrated structure of theoretical climate-economics and finance models requires a high dimensional state space in order to rigorously and accurately account for relevant model features. As noted by Han et al. (2018), numerical solutions using standard finite difference or finite element methods are almost always unavailable for problems where the number of dimensions is greater than or equal to four[10]. Incorporating model uncertainty and jumps, which introduces endogenous non-linearities as noted before, can further exacerbate the computational burden in high dimensions.

[10]There are various alternatives approaches that have been explored for solving problems with more than 4 state variables, of which deep learning algorithms are a recent one that appears to have substantial promise.

## 4.5 Implementing Neural Nets for Numerical Solutions

We therefore propose a novel implementation of deep neural networks for deriving the numerical solutions to HJB equations by focusing on an algorithm designed not only to deal with high-dimensionality, but also the nonlinearities and complexities associated with jumps and model uncertainty as outlined before. Only very recently have deep learning and neural network methods been applied to solving dynamic economic models. Because of the ability of these methods to provide global solutions for problems where other computational solutions methods become infeasible or simply fail, such as with high dimensional state spaces or significant model nonlinearities, deep learning solution methods are seen as a potentially "game-changing" toolset. By approximating the representative agent's value function and optimal controls with deep neural networks, the "curse of dimensionality" can be ameliorated due to the representation of functions in a compositional form, rather than by the standard additive form resulting from finite difference and element methods. Applications in macroeconomics (Fernandez-Villaverde et al. 2020, Maliar et al. 2021, Azinovic et al. 2022, Payne et al. 2024) and finance (Duarte et al. 2023, Sauzet 2021) implementing these types of solutions methods have begun to demonstrate the potential value and importance for studying key problems in the literature more broadly.

We now outline our algorithm, and provide pseudo-code and further details about implementation in the appendix. We build upon the deep Galerkin method-policy improvement algorithms (DGM-PIA) proposed in Al-Aradi et al. (2022) and Duarte et al. (2023) to solve the aforementioned HJB equations. We first illustrate the algorithm on a generic HJB equation for $V(\boldsymbol{x})$:

$$-\delta V(\boldsymbol{x}) + \sup_{\boldsymbol{\alpha}\in\mathcal{A}}\{\mathcal{L}^{\boldsymbol{\alpha}}V(\boldsymbol{x}) + f(\boldsymbol{x},\boldsymbol{\alpha})\} = 0,$$

where $\boldsymbol{x}$ and $\boldsymbol{\alpha}$ denote the state and control variables, $\mathcal{A}$ is the control space, the differential operator $\mathcal{L}^{\boldsymbol{\alpha}}$ is the infinitesimal generator of the controlled state process $\boldsymbol{X}^{\boldsymbol{\alpha}}$, $f$ is the utility function and $\delta$ is the discount factor. DGM-PIA solves for the value function $V$ and the optimal control $\boldsymbol{\alpha}$ simultaneously by parameterizing both as deep neural networks $V^{\theta}$ and $\boldsymbol{\alpha}^{\varphi}$. Then, the networks are trained by taking alternating stochastic gradient descent steps for the two functions. Let $\boldsymbol{\alpha}^{\varphi_0}$ (as a function of $\boldsymbol{x}$) be the initial control parameterized by the neural net, at stage $n$, the algorithm contains two steps:

Step 1. Find a solution to the linear PDE

$$-\delta V^{\theta_n}(\boldsymbol{x}) + \mathcal{L}^{\boldsymbol{\alpha}^{\varphi_n}}V^{\theta_n}(\boldsymbol{x}) + f(\boldsymbol{x},\boldsymbol{\alpha}^{\varphi_n}(\boldsymbol{x})) = 0,$$

for the fixed control $\boldsymbol{\alpha}^{\varphi_n}$, by updating $\theta_n$ via minimizing

$$L_V(\theta) = \|-\delta V^{\theta}(\boldsymbol{x}) + \mathcal{L}^{\boldsymbol{\alpha}^{\varphi_n}} V^{\theta}(\boldsymbol{x}) + f(\boldsymbol{x}, \boldsymbol{\alpha}^{\varphi_n}(\boldsymbol{x}))\|^2.$$

Step 2. Update the policy corresponding to

$$\boldsymbol{\alpha}^{\varphi_{n+1}}(\boldsymbol{x}) \in \arg\max_{\boldsymbol{\alpha}\in\mathcal{A}} \{\mathcal{L}^{\boldsymbol{\alpha}} V^{\theta_n}(\boldsymbol{x}) + f(\boldsymbol{x}, \boldsymbol{\alpha})\},$$

for the fixed value function $V^{\theta_n}$, by update $\varphi_{n+1}$ via minimizing

$$L_{\boldsymbol{\alpha}}(\varphi) = -\int_{\Omega} \left[\mathcal{L}^{\boldsymbol{\alpha}^{\varphi}} V^{\theta_n}(\boldsymbol{x}) + f(\boldsymbol{x}, \boldsymbol{\alpha}^{\varphi}(\boldsymbol{x}))\right] \mathrm{d}\nu(\boldsymbol{x}),$$

where $\nu(\boldsymbol{x})$ is a probability measure on the domain $\Omega$ of $\boldsymbol{x}$ characterizing the different regions' relative importance.

### 4.5.1 Augmenting the Algorithm with "Pseudo"-State Variables

To effectively account for the nonlinearities resulting from model uncertainty and jumps, we extend our algorithm by incorporating "pseudo"-state variables that are used as inputs for the parameterized neural network. Specifically, we alter the stochastic gradient descent step in the DGM-PIA algorithm to randomly sample points from the pseudo-state space, defined in our setting by values of the model uncertainty aversion parameter $\xi$ and the potential realizations of the jump processes denoted by $\ell \in L$, as well as the standard state space. We are therefore searching for solutions over a range of $\xi$ and $\ell$ values, even though there is no dynamic evolution associated with these pseudo-states. This enhances the algorithm's ability to characterize the nonlinearities associated with the jump and model uncertainty model mechanisms, allowing us to find accurate global solutions to our problem, in particular, those aligned with economic intuition.

The reason why the use of pseudo-state variables improves the algorithm's accuracy is actually fairly intuitive. By sampling from the pseudo-state space, we provide the neural networks access to relevant information over which to train that would not be available if we were to train individual networks for each model solution across different sets of parameters. Importantly, given our choice of pseudo-state variables, this additional cross-sectional information precisely targets the model components that are the most complex and difficult characterize. Because the algorithm samples from the pseudo-state space, the network tries to learn the functional relationship with respect to the pseudo-state variables, similar to an operator learning framework. A tendency of neural networks is to exhibit implicit regulariza-

tion, meaning they tend to search for simpler functions solutions that fit the data well, such as those with monotonicity, even without imposing such constraints. If the neural networks finds a form of regularity consistent with the data, it will likely prefer this solution to other alternatives, implicitly providing additional structure to the overall solution. With this additional structure over the pseudo-state space, as the the neural network updates the solution for one set of parameter values, it will impact the solution for other parameter values. These updates are generally beneficial because the model solutions we are searching for are usually continuous and smooth in the state and parameter space. As a result, the solution for one set of parameters is likely to be relatively close to the solution for a slight perturbation to the parameter values. This additional information and structure provided by the inclusion of pseudo-states appears to be particularly valuable in allowing the neural networks to more accurately identify and fit the types of complex nonlinearities encompassed by our setting.

Our pseudo-state approach extends the deep Galerkin method, which aims to solve PDEs while maintaining a monotonic relationship with certain model parameters. This fundamentally differs from neural operator learning methods like DeepONet, which learns the functional relationship between model parameters and the solution and can be considered a form of operator learning (Kovachki et al. 2024)[11]. Rather than learning an operator from a limited set of parameters and corresponding PDE solutions, our approach samples a broad range of parameters and solves the corresponding PDEs directly. This approach emphasizes accurately solving the PDEs across a range of parameters while ensuring certain monotonicity properties, thereby distinguishing it from the operator learning framework.

## 4.6 The Numerical Algorithm Contribution in Context

To be concrete, we are unaware of any existing contributions in literature related to deep learning algorithms on solving dynamic optimal control problems with jumps and model uncertainty. Moreover, while there are non-deep learning algorithms that can handle problems of dimension 4 or higher, they are often tailored for specific problems, taking advantage of particular model structures. In contrast, our model solution setting is very general and complex, and thus a contribution of ours is to develop a "toolbox" for highly nonlinear control problems with jumps and model uncertainty.

Though existing work has explored settings with (potentially many) more state variables than ours, a number of settings require assumptions such as state variables that are symmetric, stationary, or exogenously specified, or linearity in the model framework, to maintain tractability with such scale. And while some frameworks can handle non-trivial nonlinear-

[11]Li et al. (2020) and Duarte et al. (2023) provide additional examples of operator learning.

ities or forms of uncertainty analysis, to the best of our knowledge ours is the first using neural network solution method to confront the endogeneous nonlinearities resulting from the application of dynamic decision theory to model uncertainty aversion, combined with stochastic jump processes. Developing a methodology that handle such complexities is essential to analyzing the type of climate-economics model with model uncertainty aversion broadly conceived that we put forward next.[12]

We address these issues in our framework by implementing an extended deep Galerkin method algorithm that considers critical parameters as additional "pseudo" state variables in the network input, an extension of the DGM-PIA setting not previously explored in the literature to our knowledge. Because of this, we are able to derive global numerical solutions for a continuous-time, infinite horizon setting with significant non-linearities, multiple endogenous state variables, and stochastic jumps while incorporating aversion to model uncertainty via the toolset of dynamic decision theory. This is critical to our analysis, as exploring transition dynamics for endogenous optimal policy responses is at the heart of understanding and analyzing outcomes related to climate change, model uncertainty, and the transition to carbon-neutrality. Furthermore, our algorithm can still handle high-dimensional PDEs, attenuating the "curse of dimensionality" in such cases. As such, the resiliency of our algorithm to various modeling complexities, while still being able to scale to higher-dimensional settings, opens the door to exploring models with regional, household, and firm heterogeneities across technologies, economic frictions, policy objectives, and other sources. Thus, our numerical algorithm not only enriches our ability to explore a more extensive set of questions and models in climate-economics and climate finance than before, specifically those that apply dynamic decision theory to address model uncertainty, but also more broadly in economics, finance, and other problems applying dynamic stochastic optimal control.

# 5 Numerical Results

We now present and discuss the results from the model solution derived using the numerical algorithm outlined above. Before getting into the results, we briefly outline a few details for

[12]For example, while an improvised finite difference method using neural networks to solve linear systems may help mitigate the curse of dimensionality, it does not produce satisfactory results for complex models with strong nonlinearities as in our setting. Also, due to the strong nonlinearities in our model, naively parameterizing the value function using a neural network and minimizing the loss associated with the PDE operator (like the standard deep Galerkin method) will not produce accurate enough results. While an extended version of the deep Galerkin method (Al-Aradi et al. 2022, Duarte et al. 2023) is indeed helpful in improving accuracy, it does not effectively preserve the value function's monotonicity with respect to certain parameters which are supported by economic arguments, nor does it sufficiently handle the nonlinearities arising from our application of dynamic decision theory to address concerns about model uncertainty.

completeness regarding model assumptions, functional forms, and parameter values. After presenting and discussing the numerical results, we discuss details about how we validate our neural-network-based solutions.

## 5.1 Functional Forms, Assumptions, and Parameter Values

For tractability we consider the case of independent Brownian shocks, i.e.,

$$\begin{aligned} 0 &= \sigma_d' \sigma_g = \sigma_d' \varsigma = \sigma_d' \sigma_\kappa, \\ 0 &= \sigma_g' \varsigma = \sigma_g' \sigma_\kappa, \\ 0 &= \varsigma' \sigma_\kappa. \end{aligned}$$

We use the following functional forms for our jump arrival rates. The damage jump intensity $\mathcal{J}_d^\ell(y)$ follows the arrival rate from Barnett et al. (2023):

$$\begin{aligned} \mathcal{J}_n^\ell(y) &= \left(\frac{1}{L}\right) \mathcal{J}_n(y), \ \ \ell = 1, ..., L, \\ \mathcal{J}_n(y) &= r_1 \left( \exp\left( \frac{r_2}{2} (y - \underline{y})^2 \right) - 1 \right) \mathbf{1}_{\mathbf{y} \geqslant \underline{\mathbf{y}}} \end{aligned}$$

We assume a technology change jump arrival rate that is affine in the knowledge stock:

$$\mathcal{J}_g^{\ell'}(R) = (1 - \pi) R, \quad \mathcal{J}_g^{\ell''}(R) = \pi R$$

The knowledge stock variable is scaled such that $R$ is in arrival rate units, and is chosen based on expected green policy implementation timelines proposed by various countries. Section D provides further details about this calibration choice.

The probabilities and arrival rates are taken as baseline probability densities, and we allow for robust probability distortions to these baseline specifications.

Finally, we briefly outline the parameter values chosen for our numerical example. Further details provided in the appendix. Our economic parameters largely follow the calibration strategy of Barnett et al. (2022). In particular, $\delta$ matches values of the subjective discount rate used in the macroeconomics and asset pricing literature. The values of $(\Gamma_d, \theta_d)$ and $(\Gamma_g, \theta_g)$ are set to match empirical measures of Tobin's q and economic growth from BEA and World Bank databases, as well as to normalize the ratio of the adjustment cost parameters normalized to one. The values of $\sigma_d$ and $\sigma_g$ match the time series of GDP from the World Bank database. The values of $A_d$ and $A_g$ generate output consistent with World Bank World GDP values based on the values of initial capital stocks. We construct a set of intermediate

and terminal technological innovation values of green productivity, $\{A_g(\ell'), A_g(\ell'')\}$, using the empirical estimates of innovation steps and gaps from Acemoglu et al. (2016).

The parameters related to the knowledge stock were determined as follows. While the knowledge stock depreciation parameter $\zeta$ and R&D investment cost parameter $\psi_1$ are chosen in part for simplicity and computational tractability, together with the choice of the R&D investment cost parameter $\psi_0$, our values are set be consistent with estimates for the returns to R&D investment by Lucking et al. (2019) and Bloom et al. (2019), to generate R&D investment values consistent with estimates by Stine (2008) of major historical U.S. R&D investment programs, and to produce an expected arrival time of a green technological innovation consistent with proposed policy timelines (between 30 and 40 years). The value of $\sigma_r$ the time-series volatility of the U.S. R&D capital stock estimates in the BLS database.

The climate dynamics and climate damage parameter values are set as in Barnett et al. (2023). Specifically, $\bar{\theta}$ is the average value across 144 climate model outcomes constructed from pulse experiments provided by Joos et al. (2013) and Geoffroy et al. (2013). The value of $\eta$ is set to fit estimates of annual carbon emissions from Figueres et al. (2018) based on World Bank, EIA, and IEA estimates of the clean and dirty capital split. The value of $\varsigma$ comes from Barnett et al. (2022). The values of $\lambda_1$, $\lambda_2$, $\lambda_3(\ell)$, and $\bar{y}$ are designed to incorporate the spread of potential climate damage outcomes based on Nordhaus (2019), Weitzman (2012), and the recent estimates by Waidelich et al. (2024).

Initial values for capital, temperature anomaly, the green capital ratio, and the knowledge capital stock come from counterparts in the World Bank National Accounts data, Masson-Delmotte et al. (2021b) estimates, Fried et al. (2021), Dhar and Davidson (2022), Carattini et al. (2023) and Berk and Van Binsbergen (2025) estimates, and the BLS database.

## 5.2 Model Solution Results

We now discuss the computational results of our model. The results shown are simulation pathway outcomes based on the solutions to the HJB equations and are initialized at today's values of the state variables and shown out to 60 years. The results provided are for the following specifications related to model uncertainty:

- Misspecification over climate, green and dirty capital, and knowledge capital dynamics
- Misspecification over the technological change and damage curvature jumps
- 5 equi-spaced post-jump damage model realizations: $\lambda_3(\ell) \in \{0, \ldots, \frac{1}{3}\}$
- Intermediate and breakthrough green technology: $A_g(\ell') = 0.1303, A_g(\ell'') = 0.1567$

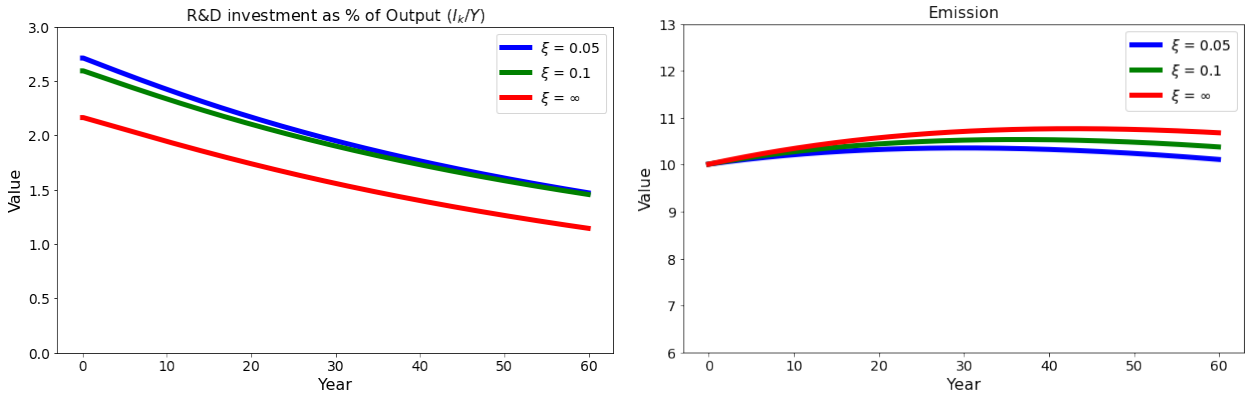
- Comparison across uncertainty parameters $\xi \in \{0.05, 0.1, \infty\}$

We focus on the simulated pathways from the pre-damage jump and pre-technology jump state for emissions, R&D investment, green and dirty capital investment, as well the distribution of climate, damage, and technology models, and the probabilities of a damage or technology jump occurring. We examine each of these results across different uncertainty aversion parameters to highlight the implications of increased model uncertainty on the optimal policy choices by our planner. For the simulated pathways, the red lines, represent the uncertainty neutral case when $\xi = \infty$, green lines represent the less uncertainty averse case when $\xi = 0.1$, and the blue lines represent the more uncertainty averse case when $\xi = 0.05$. For the model probabilities, the red histogram bars represent the uncertainty neutral case when $\xi = \infty$, and the blue histogram bars represent the less ($\xi = 0.1$) or more ($\xi = 0.05$) uncertainty averse as denoted in the figure.

### 5.2.1 Optimal R&D, Emissions, and Investment Choices

We start by examining the optimal economic control outcomes from our framework: R&D investment as a fraction of total output, carbon emissions in gigatons of carbon, and the investment amounts in level terms for green and dirty capital. Figure 3 shows the pathways for optimal R&D investment and emissions over time for different cases of uncertainty aversion. In the left panel, we show R&D investment as a percent of total output. Note first that the magnitude of R&D is fairly substantial in each case, ranging between about 2.0% and 2.75% of total output initially. For comparison's sake, Stine (2008) notes that expenditure on R&D directed towards the Manhattan, Apollo, and the Federal Energy Technology programs reach magnitudes around 0.5% of GDP. We can also see concerns about model uncertainty lead to noticeable impacts to the choice of R&D investment, though perhaps in a somewhat surprising way. Throughout the simulated pathways the more uncertainty averse the planner chooses increased levels of R&D investment. In each case, the magnitude of R&D diminishes over time, dropping down by about one percentage point after 60 years in each case. In the two uncertainty averse cases, the R&D investment starts higher, by about 50 bps for the less aversion case and nearly 75 bps for the high aversion case. The uncertainty aversion scenarios maintain this higher level throughout the simulated pathway. Thus, the endogenous uncertainty aversion probability adjustments in our framework strongly emphasizes additional R&D investment.

The pathway for emissions, shown in the right panel, also shows important implications generated by model uncertainty concerns. Mechanically, because emissions are determined by production using dirty capital, all the pathways start out at the same value. Also, in each




(a) R&D investment-to-output ratio

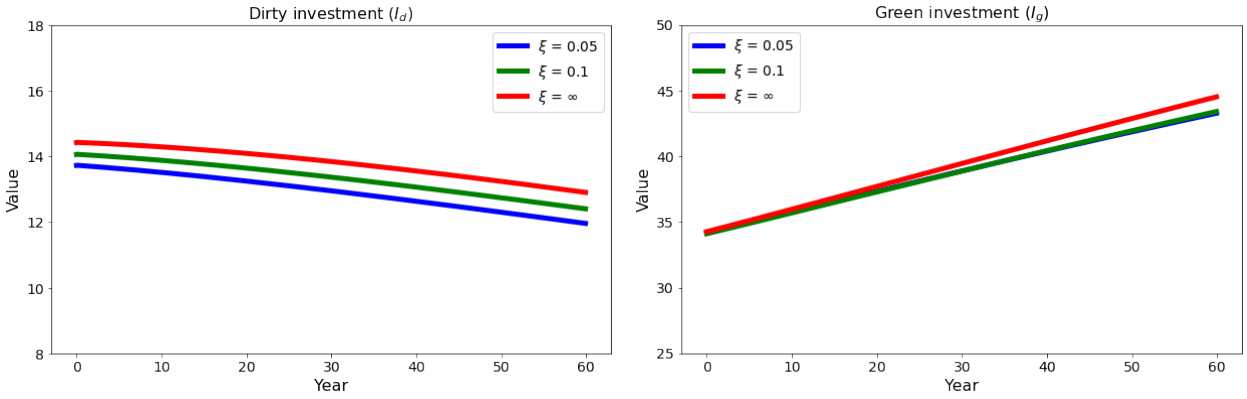
(b) Gigatons of carbon emissions

Figure 3: Simulated pathways of R&D investment and emissions based on the numerical solutions under different uncertainty penalty configurations. Panel (a) shows the pathway for R&D investment as a fraction of total output and Panel (b) shows the pathway for carbon emissions in gigatons of carbon (GtC). For each panel, results for three cases of model uncertainty are shown: $\xi = \infty$ (red line), $\xi = 0.1$ (green line), and $\xi = 0.05$ (blue line). The trajectories are simulated under the baseline probabilities abstracting from the intrinsic randomness.

case the pathways are relatively flat or even slightly declining over the simulated pathway. An important reason for these responses of emissions in our framework is related to the fact that emissions can only be reduced by reducing the dirty capital stock. This difficulty in reducing emissions is a feature of the model structure, as we have focused on a framework aimed at addressing concerns by energy economists and others that reducing emissions is likely to be far more difficult than put forward by many existing modeling frameworks[13]. Importantly, however, increased concerns about uncertainty do have a noticeable impact over emissions, with the more averse case of $\xi = 0.05$ seeing the strongest response in terms of reducing emissions over time. The difference in emissions between the uncertainty neutral and more aversion case is near 1 GtC of carbon by 60 years, with the response of the less averse case seeing a more modest reduction. We can therefore see that the optimal uncertainty-adjusted policy also leads to cautious emissions choices that become increasingly more pronounced the greater are the uncertainty concerns.

Next, we examine the pathways for the optimal level of investment in dirty and clean capital over time for different cases of uncertainty aversion, shown in Figure 4. The simulated pathways for dirty capital investment are shown in the left panel, while the simulated pathways for green capital investment are shown in the right panel. Comparing outcomes

[13]See Reguant (2021) for one example of existing work expressing such concerns.




(a) Dirty capital investment

(b) Green capital investment

Figure 4: Simulated pathways of investment in dirty and green capital stocks based on the numerical solutions under different uncertainty penalty configurations. Panel (a) shows the level of investment in the dirty capital stock and Panel (b) shows the level of investment in the green capital stock. For each panel, results for three cases of model uncertainty are shown: $\xi = \infty$ (red line), $\xi = 0.1$ (green line), and $\xi = 0.05$ (blue line). The trajectories are simulated under the baseline probabilities abstracting from the intrinsic randomness.

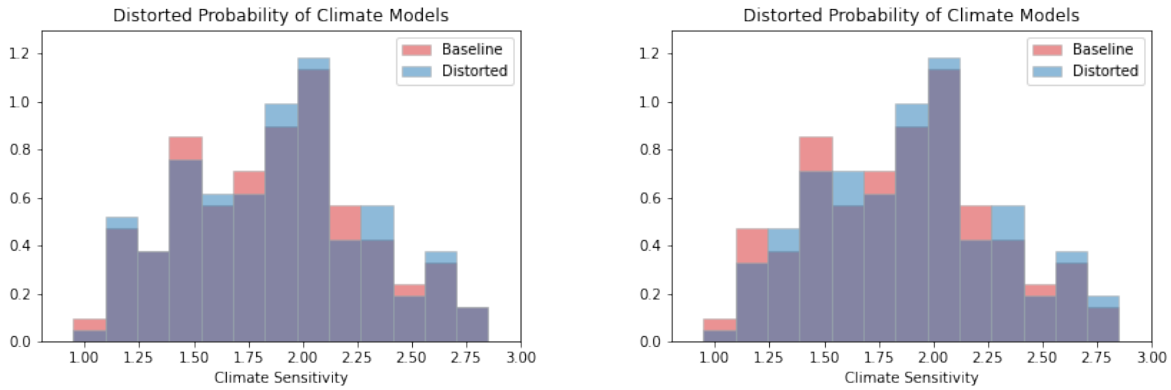
across the uncertainty aversion cases and across the two types of investment, we see three key takeaways from the planner's optimal policy choices as it relates to the two types of capital investment. First, comparing across the two types of investment, the level of investment in green capital is substantially higher than in dirty capital. At the beginning of the simulation pathway, the green investment level is about 2.5 times higher, \$34 trillion for all cases of green investment versus roughly \$14 trillion for each case of dirty investment. Yet, by the end of the pathway, investment in green capital is nearly four times higher than for dirty capital, around \$42-\$44 trillion for green investment as compared to \$12-\$13 trillion for dirty capital. This latter comparison also highlights the second key observation related to investment across different capital stocks in that green investment persistently increases during the simulation pathway, whereas dirty investment is always decreasing across the simulated pathway horizon.

The final investment result comes from comparing across uncertainty averse cases, where we see economically interesting implications for both dirty and green capital investment. First, uncertainty aversion leads to decreased investment in both green and dirty capital. The decrease in dirty capital investment is relatively more pronounced as $\xi$ decreases over the simulated pathway than for green capital investment. The immediate and persistent reduction in dirty capital investment is what is behind the widening gap in emissions for increased uncertainty concerns that we saw in the previous set of results. This result seems

intuitive, as concerns about increased climate damages would naturally incentivize the planner to want to reduce emissions, and even more so if uncertainty concerns lead to increased weight on more severe climate change consequences. The reduction in green investment is potentially more unexpected. While clearly the planner places substantial value on growing the green capital stock, this has to be weighed against choosing to instead engage in additional R&D investment or responding to uncertainty concerns that the productivity and growth of green capital might be less than the baseline model would anticipate. Therefore, we see that uncertainty also plays a notable role in the capital investment side, where incentives to proactively respond to potential worst-case climate consequences and concerns about economic productivity lead to reductions in both dirty and green capital investment.

### 5.2.2 Probability Distortions from Uncertainty Aversion

To more directly analyze the underlying uncertainty mechanisms driving the robustly optimal policy choices we have just seen, we now examine the endogenous optimal probability distortions derived by the uncertainty averse planner in our framework. Figures 5 - 7 show the probability distortions used by the planner in making robustly optimal policy choices for the climate model, the damage model, the technology model, and the technology and damage jump arrivals.



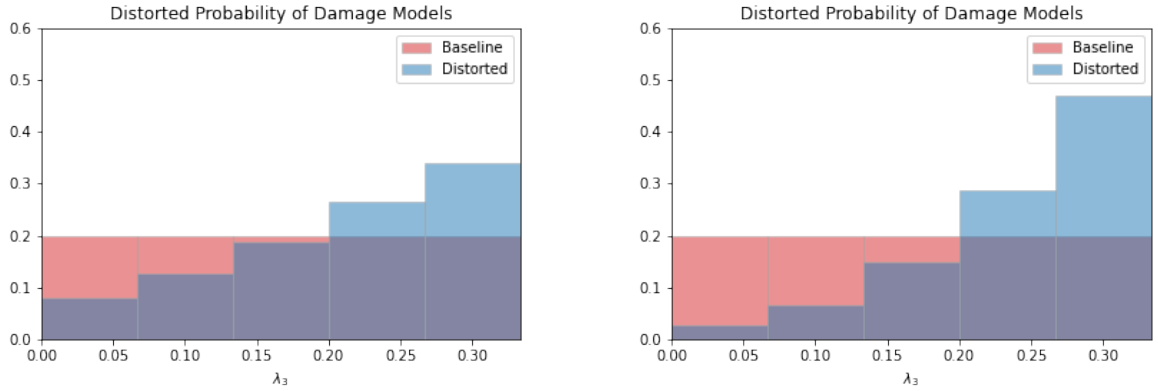

(a) Less aversion

(b) More aversion

Figure 5: Undistorted and distorted climate model distributions. The left plot shows the undistorted and distorted distributions with less misspecification aversion ($\xi = 0.1$). The right plot shows the undistorted and distorted distributions with more misspecification aversion ($\xi = 0.05$). The histograms are calculated at year 40, with the trajectories leading up to year 40 simulated under the baseline probabilities and abstracting from the intrinsic randomness.

We start by examining the probability distortions for the climate model in Figure 5. The

left panel compares the baseline probabilities in red to the distorted probabilities for the less averse case in blue, while the right panel compares the baseline probabilities in red to the distorted probabilities for the more averse case in blue. In each case, the distortion effect is fairly modest, with a slight adjustment leading to less weight in the left tail where the least severe climate sensitivity models are and more weight in the right tail where the most severe climate sensitivity models are. The effect is only modestly more pronounced for the increased level of uncertainty aversion. So while there is some incentive for the planner to act to protect against future climate change impacts, the effect through the climate model uncertainty channel is somewhat limited.




(a) Less aversion

(b) More aversion

Figure 6: Undistorted and distorted damage model distributions. The left plot shows the undistorted and distorted distributions with less misspecification aversion ($\xi = 0.1$). The right plot shows the undistorted and distorted distributions with more misspecification aversion ($\xi = 0.05$). The histograms are calculated at year 40, with the trajectories leading up to year 40 simulated under the baseline probabilities and abstracting from the intrinsic randomness.

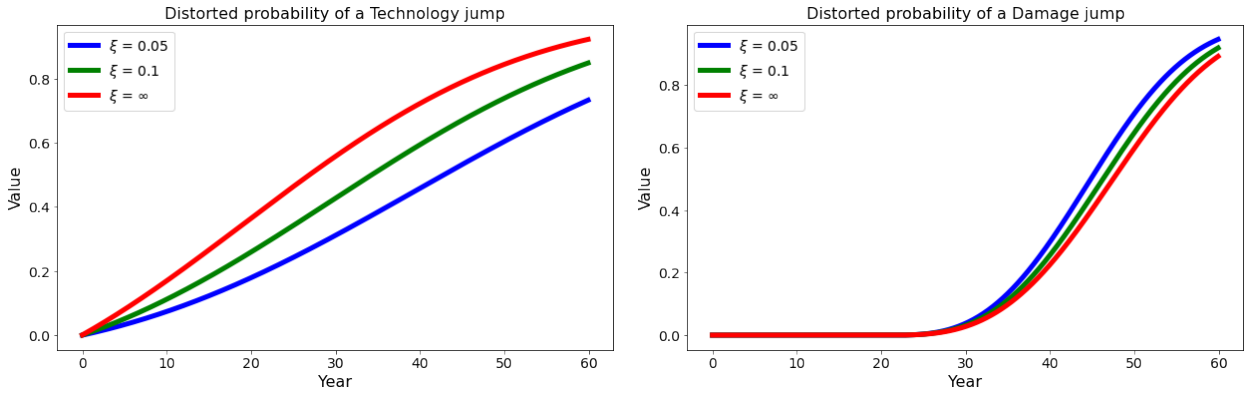
Figure 6 shows the probability distortions for the damage model. The impact of uncertainty aversion on the probabilities of the damage models used for by the planner for robust decisions-making is similar qualitatively to what we saw for the climate models, though effect is more pronounced along this channel. With uncertainty aversion, the planner tilts the probabilities towards the worst case damage models and away from the more favorable climate models, increasing the tilting as uncertainty aversion increases (going from the less aversion distortion in the left panel to the more aversion distortion in the right panel). These damage parameter probability distortions are more pronounced than we saw with the climate models, highlighting the more substantial concerns about potentially severe future climate damages than the climate model channel endogenously determined by the planner.

| Uncertainty Aversion | Prob("Catch Up") ($\tilde{\pi}$) | Prob("Breakthrough") ($1-\tilde{\pi}$) |
|---|---|---|
| Neutral ($\xi = \infty$) | 0.9631 | 0.0369 |
| Less Aversion ($\xi = 0.1$) | 0.9995 | 0.0005 |
| More Aversion ($\xi = 0.05$) | 1.0000 | 0.0000 |

Table 1: Undistorted and distorted technology model distributions. The table provides the probability of a "catch up" or "breakthrough" innovation realization conditional on a technology jump occurring. The top row shows the outcomes for uncertainty neutrality ($\xi = \infty$), the middle rows show the outcomes for less uncertainty aversion ($\xi = 0.1$), and the bottom row shows outcomes for more uncertainty aversion ($\xi = 0.05$). The probabilities are calculated at year 40, with the trajectories leading up to year 40 simulated under the baseline probabilities and abstracting from the intrinsic randomness.

The impact of uncertainty aversion on the implied probability distribution for technology models is given in Table 1. In this case, uncertainty aversion shifts the most of the already modest probability weight from the "breakthrough" realization to the "catch up" realization, with the effect amplified as we move from less aversion, in the left panel, to more aversion in the right panel. The intuition here is the same as with the previous two channels discussed. The planner makes optimal policy choices based on a distorted distribution that shifts probability from the best-case technological innovation outcome to the worst outcome technological innovation outcome. However, what is clear is that the probability distortion is quite substantial tilted towards concerns that the breakthrough shock will not come initially. Thus, the planner endogenously determines that technological innovation concerns are an economically important channel for robustly optimal policy choices.

The final probability distortions we examine are the pathways for the distorted jump probabilities, given in Figure 7. The right panel shows the distorted probability of a damage jump taking place the uncertainty neutral, less averse, and the more averse cases. It is immediately clear here that uncertainty concerns have only a modest impact to the probability of a damage jump taking place used by the planner. However, comparing across uncertainty aversion cases for the implied probability of a technology jump taking place, shown in the left panel, is a different story. The uncertainty averse planner down-weights the probability of a technology jump occurring as they incorporate more concern about model uncertainty into their robust decision making. This result highlights again, along with the distortion to the technology model probability distribution, that technological innovation uncertainty is a central focus of uncertainty aversion for the planner, as determined endogenously through the robust optimal decision making framework provided by dynamic decision theory. However, rather than our planner choosing to reduce R&D investment due to concerns about a




(a) Technology jump probability

(b) Damage jump probability

Figure 7: Simulated pathways of the cumulative probability of a technology or damage jump occurring based on the numerical solutions under different uncertainty penalty configurations. Panel (a) shows the distorted probability of a technology jump occurring and Panel (b) shows the distorted probability of a damage jump occurring. For each panel, results for three cases of model uncertainty are shown: $\xi = \infty$ (red line), $\xi = 0.1$ (green line), and $\xi = 0.05$ (blue line). The trajectories are simulated under the baseline probabilities abstracting from the intrinsic randomness.

decreased likelihood of the technological change taking place, the planner's cautious response is to augment R&D investment to try and increase the likelihood of the innovation shock occurring. This is somewhat offset by increasing concerns that the innovation realization could be more modest in magnitude than the baseline probabilities suggest. As a result, the planner does not go "all in" on R&D investment, but combines a proactive increased R&D investment response with a cautious dirty capital investment response, leading to constrained carbon emissions, to optimally respond to concerns about the different channels and forms of model uncertainty.

Our analysis of the impacts of uncertainty aversion, decomposed into the different probability distortion channels, highlights important effects influencing the optimal policy decisions of the planner in our model. First, the clear response of robust policy action in this setting is to avoid significant future climate change consequences. The question of how becomes clear from which probability distortions the planner finds most important in terms of welfare implications. The distortions to the probability weights used on the climate models is quite modest, with only limited effect in either the less or more averse case. The distortion effect is larger for the damage model probabilities, however the damage jump probability is quite limited. Importantly from the planner's perspective, notable probability distortions are made to the technology model weighting and the probability of a technology jump taking place. Thus, the policy response by the planner emphasizes adjustments to the R&D

policy in response to uncertainty concerns, with some restraint on dirty capital investment and modest additional adjustments to green investment as a result of uncertainty aversion. The social optimality of these responses depends importantly on the perceived social payoff of such actions. The sticky nature of emissions given their dependence on dirty capital, the higher productivity of dirty capital relative to green capital, and the potential welfare payoff of experiencing a significant green technological innovation leads our planner to make intersectoral and intertemporal investment trade-offs based on their uncertainty concerns. The perceived long-run social payoff of increased R&D investment and reduced carbon emissions outweighs alternative investment options choices, and emphasizes the importance of R&D and technological innovation even under model uncertainty concerns, and investment responses to support such long-term R&D targets.

# 6 Uncertainty Decomposition and Social Valuations

In continued work, we plan to further evaluate the impact of uncertainty channels in the model using decompositions that build on the work in Barnett et al. (2020) and Barnett et al. (2023). Specifically, our model allows us to evaluate and disentangle the consequences of the following types of uncertainty on optimal policy choices:

- Uncertainty resulting from the climate model, economic damages of climate change, technological innovation, and economic productivity model components
- Misspecification coming through drift distortions associated with the Brownian motion and misspecification coming through distortion of the jump processes
- Uncertainty related to potential post damage function jump realizations and uncertainty coming from potential post technological change realizations

To evaluate each of these channels and uncertainty forms, we will isolate each component individually through alternative model solutions and social valuations derived through Feynman-Kac characterizations of the value function marginal values. In doing so, we will highlight which model features and economic channels are endogenously determined by our planner to be of most significance when making valuation and optimal policy choices. This will further enhance the economic intuition and understanding of our model framework and optimal policy choices made by our uncertainty averse planner.

# 7 Algorithm Validation

An important issue to address for our numerical solutions is the validation of the accuracy of our results. We address this in various ways, with the results shown in Figure 8. First, because the post-jump model solutions require only 3 state variables, we can compare our augmented DGM-PIA algorithm results for this setting with the solution derived using more standard finite difference methods.[14] Figure 8b shows the neural net and finite difference solutions of the value function and optimal clean and dirty investment for different values of $R$, and values of $\log K$ and $Y$ fixed at the midpoints of our state space range ($\log K = 5.5, Y = 1.5$). We see that the two solutions are consistent for this part of the state space, and unreported results show the same consistency across the state space.

In addition, we examine the training loss for our neural net solutions to determine the magnitude of the HJB equation error for our solution from the DGM-PIA method. Figure 10a shows the training error for the post-technology, post-damage jump state solution across epochs. We reach a training error of approximately $10^{-3}$ or better for each of the four jump states, suggesting reasonable accuracy of our numerical solutions.

Finally, an important component for further evaluation of the validity of our solutions is to consider whether the training loss errors concentrate in economically meaningful parts of the state space. To explore this, Appendix F provides tables and figures with the training loss errors across the part of the state space most likely to be visited along our simulations. Specifically, we compute the 5th and 95th percentiles of the state variable values at years 20, 40, and 60 for 10,000 stochastic simulation pathways and show the range of training loss values across those economically relevant regions of the state space. We do this for all jump realization states and find that for all cases at all time periods considered our training loss errors are consistently on the order of $10^{-3}$. These outcomes further validate the accuracy of our numerical solutions and provide evidence that errors are not concentrated over economically relevant regions of the state space.

# 8 Conclusion

We analyze a particularly relevant economic example that integrates dynamic decision theory under uncertainty into a climate-economics-innovation framework with multiple capital stocks to guide uncertainty quantification of optimal R&D investment and investment in dirty and clean production capital. We examine different forms of uncertainty, including

[14] In particular, we use the false-transient, conjugate gradient-based numerical algorithm implemented in Barnett et al. (2020) and Barnett et al. (2022) for our finite difference solutions.

diffusion and jump process model misspecification, as well as the implications of uncertainty from climate sensitivity, climate damages, and technological innovation. Concerns about model uncertainty feed into the endogenous equilibrium policy responses in our model by "tilting" the stochastic discounting implicitly used in constructing the marginal valuations related to the externalities in our framework. This results in first-order implications for the socially optimal outcomes in our model as a result of incorporating aversion to model uncertainty in the planner's decision problem. In particular, uncertainty aversion amplifies the incentive to invest in technological innovation, with modest adjustments to reduce investment in dirty capital to constrain future carbon emissions, as the major policy response. Our algorithm allows us to derive the necessary global solutions required for computing social valuations from our continuous-time, infinite horizon social planner's problem that optimizes lifetime expected utility subject to model uncertainty concerns with numerous endogenous and potentially non-stationary state variables.

To solve our model, we develop and implement a deep learning-based algorithm to solve high-dimensional Hamilton-Jacobi-Bellman equations that account for the rich and complex nonlinearities that are of first-order importance in dynamic climate-economic models. Specifically, we focus on settings that incorporate endogenous nonlinearities that arise from model uncertainty aversion and jump processes related to climate damages and technological innovation. Model uncertainty along multiple dimensions, and jump processes characterizing economic and climate thresholds, are essential model components for a rigorous analysis of optimal climate policy for a carbon-neutral transition. Our algorithm provides one of the first approaches that we are aware of applying deep neural networks and pseudo-state variables in a way designed to handle such numerical complexities.

There are various dimensions of future research that our work helps open the door to. Along the climate-economics dimension, extensions include settings with heterogeneity across agents, goods, technologies, and climate damages. There are also important considerations related to allowing explicitly for learning about economic and climate dynamics, as well as decentralized equilibria with micro-founded market frictions and policy mechanisms to explore. In addition, our numerical algorithm is well suited to address broader applications in economics and asset pricing, such as settings with rare-disasters, long-run risk, financial frictions, tipping points or thresholds, and other economically and scientifically rich frameworks. Finally, because our algorithm is designed to handle general high-dimensional problems that incorporate endogenous nonlinearities, we see significant potential for application of our methods to more general stochastic optimal control settings. We leave exploration of such alternative settings for future work.

# References


Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous. 2012. The environment and directed technical change. *American economic review* 102 (1):131–66.

Acemoglu, Daron, Ufuk Akcigit, Douglas Hanley, and William Kerr. 2016. Transition to clean technology. *Journal of Political Economy* 124 (1):52–104.

Al-Aradi, Ali, Adolfo Correia, Gabriel Jardim, Danilo de Freitas Naiff, and Yuri Saporito. 2022. Extensions of the deep Galerkin method. *Applied Mathematics and Computation* 430:127287.

Allen, M, P Antwi-Agyei, F Aragon-Durand, M Babiker, P Bertoldi, M Bind, S Brown, M Buckeridge, I Camilloni, A Cartwright, et al. 2019. Technical Summary: Global warming of 1.5 °C. An IPCC Special Report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. Tech. rep., IPCC.

Anderson, Evan W, Lars Peter Hansen, and Thomas J Sargent. 2003. A quartet of semigroups for model specification, robustness, prices of risk, and model detection. *Journal of the European Economic Association* 1 (1):68–123.

Armstrong McKay, David I, Arie Staal, Jesse F Abrams, Ricarda Winkelmann, Boris Sakschewski, Sina Loriani, Ingo Fetzer, Sarah E Cornell, Johan Rockström, and Timothy M Lenton. 2022. Exceeding 1.5 C global warming could trigger multiple climate tipping points. *Science* 377 (6611):eabn7950.

Arrhenius, Svante. 1896. On the influence of Carbonic Acid in the Air upon Temperatue of the Ground. *Philosophical Magazine and Journal of Science Series 5* 41:237–276.

Azinovic, Marlon, Luca Gaegauf, and Simon Scheidegger. 2022. Deep equilibrium nets. *International Economic Review* 63 (4):1471–1525.

Bachouch, Achref, Côme Huré, Nicolas Langrené, and Huyen Pham. 2021. Deep neural networks algorithms for stochastic control problems on finite horizon: numerical applications. *Methodology and Computing in Applied Probability* 1–36.

Barnett, Michael. 2023. Climate change and uncertainty: An asset pricing perspective. *Management Science* .

Barnett, Michael, William A. Brock, and Lars Peter Hansen. 2020. Pricing Uncertainty Induced by Climate Change. *Review of Financial Studies* 33 (3):1024–1066.

Barnett, Michael, William Brock, and Lars Peter Hansen. 2022. Climate change uncertainty spillover in the macroeconomy. *NBER Macroeconomics Annual* 36 (1):253–320.

Barnett, Michael, William A Brock, Lars Peter Hansen, and Hong Zhong. 2023. How Should Climate Change Uncertainty Impact Social Valuation and Policy? *Working Paper* .

Baumol, William J. 1986. Productivity growth, convergence, and welfare: what the long-run data show. *The american economic review* 1072–1085.

Beck, C., W. E, and A. Jentzen. 2019. Machine learning approximation algorithms for high-dimensional fully nonlinear partial differential equations and second-order backward stochastic differential equations. *Journal of Nonlinear Science* 29 (4):1563–1619.

Beck, Christian, Sebastian Becker, Patrick Cheridito, Arnulf Jentzen, and Ariel Neufeld. 2021. Deep splitting method for parabolic PDEs. *SIAM Journal on Scientific Computing* 43 (5):A3135–A3154.

Berk, Jonathan B and Jules H Van Binsbergen. 2025. The impact of impact investing. *Journal of Financial Economics* 164:103972.

Bernanke, Ben S and Refet S Gürkaynak. 2001. Is growth exogenous? taking mankiw, romer, and weil seriously. *NBER macroeconomics annual* 16:11–57.

Bloom, Nicholas, John Van Reenen, and Heidi Williams. 2019. A toolkit of policies to promote innovation. *Journal of economic perspectives* 33 (3):163–84.

Bouckaert, Stéphanie, Araceli Fernandez Pales, Christophe McGlade, Uwe Remme, Brent Wanner, Laszlo Varro, Davide D'Ambrosio, and Thomas Spencer. 2021. Net zero by 2050: A roadmap for the global energy sector. Tech. rep., International Energy Agency.

Boussange, Victor, Sebastian Becker, Arnulf Jentzen, Benno Kuckuck, and Loïc Pellissier. 2022. Deep learning approximations for non-local nonlinear PDEs with Neumann boundary conditions. *arXiv preprint arXiv:2205.03672* .

Brock, William A. 1982. Asset prices in a production economy. In *The economics of information and uncertainty*, 1–46. University of Chicago Press.

Brock, William A and Leonard J Mirman. 1972. Optimal economic growth and uncertainty: the discounted case. *Journal of Economic Theory* 4 (3):479–513.

Brook, Barry W., Erle C. Ellis, Michael P. Perring, Anson W. Mackay, and Linus Blomqvist. 2013. Does the terrestrial biosphere have planetary tipping points? *Trends in Ecology and Evolution* 28:396–401.

Cai, Yongyang and Thomas S Lontzek. 2019. The social cost of carbon with economic and climate risks. *Journal of Political Economy* 127 (6):2684–2734.

Cai, Yongyang, Kenneth L Judd, Timothy M Lenton, Thomas S Lontzek, and Daiju Narita. 2015. Environmental Tipping Points Significantly Affect the Cost-Benefit Assessment of Climate Policies. *Proceedings of the National Academy of Sciences* 112 (15):4606–4611.

Cai, Yongyang, Kenneth L. Judd, and Thomas S. Lontzek. 2017. The Social Cost of Carbon with Climate Risk. Tech. rep., Hoover Institution, Stanford, CA.

Carattini, Stefano, Garth Heutel, and Givi Melkadze. 2023. Climate policy, financial frictions, and transition risk. *Review of Economic Dynamics* 51:778–794.

Carleo, G. and M. Troyer. 2017. Solving the quantum many-body problem with artificial neural networks. *Science* 355 (6325):602–606.

Carleton, Tamma and Michael Greenstone. 2021. Updating the United States Government's Social Cost of Carbon. SSRN Working Paper 2021-04, University of Chicago, Becker Friedman Institute for Economics.

Carmona, René and Mathieu Laurière. 2022. Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games: II—the finite horizon case. *The Annals of Applied Probability* 32 (6):4065–4105.

Castro, Javier. 2022. Deep learning schemes for parabolic nonlocal integro-differential equations. *Partial Differential Equations and Applications* 3 (6):77.

Cerreia-Vioglio, Simone, Lars Peter Hansen, Fabio Maccheroni, and Massimo Marinacci. 2021. Making Decisions under Model Misspecification.

Cochrane, John H. 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. *The Journal of Finance* 46 (1):209–237.

Colacito, Riccardo, Bridget Hoffmann, and Toan Phan. 2019. Temperature and growth: A panel analysis of the United States. *Journal of Money, Credit and Banking* 51 (2-3):313–368.

Dell, Melissa, Benjamin F Jones, and Benjamin A Olken. 2012. Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics* 4 (3):66–95.

Dhar, Shamik and Brian Davidson. 2022. An investor's guide to net zero by 2050. *Bnymellonim. Com* .

Dietz, Simon and Frank Venmans. 2019. Cumulative carbon emissions and economic policy: In search of general principles. *Journal of Environmental Economics and Management* 96:108–129.

Drijfhout, Sybren, Sebastian Bathiany, Claudie Beaulieu, Victor Brovkin, Martin Claussen, Chris Huntingford, Marten Scheffer, Giovanni Sgubin, and Didier Swingedouw. 2015. Catalogue of abrupt shifts in Intergovernmental Panel on Climate Change climate models. *Proceedings of the National Academy of Sciences* 112 (43):E5777–E5786.

Duarte, Victor. 2018. Gradient-Based Structural Estimation. *SSRN, 2023a. URL Available at SSRN: https://ssrn.com/abstract=* 3166273.

Duarte, Victor, Diogo Duarte, and Dejanir Silva. 2023. Machine learning for continuous-time finance. *SSRN, 2023a. URL https://ssrn. com/abstract=* 3012602.

E, Weinan, Jiequn Han, and Arnulf Jentzen. 2017. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Commun. Math. Stat.* 5 (4):349–380.

Eby, M., K. Zickfeld, A. Montenegro, D. Archer, K. J. Meissner, and A. J. Weaver. 2009. Lifetime of anthropogenic climate change: Millennial time scales of potential CO2 and surface temperature perturbations. *Journal of Climate* 22 (10):2501–2511.

Fernandez-Villaverde, Jesus, Galo Nuno, George Sorg-Langhans, and Maximilian Vogler. 2020. Solving high-dimensional dynamic programming problems using deep learning. *Unpublished working paper* .

Figueres, Christiana, Corinne Le Quéré, Anand Mahindra, Oliver Bäte, Gail Whiteman, Glen Peters, and Dabo Guan. 2018. Emissions Are Still Rising: Ramp Up the Cuts. *Nature* 564:27–30.

Franzke, Christian LE, Richard Blender, Terence J O'Kane, and Valerio Lembo. 2022. Stochastic methods and complexity science in climate research and modeling. *Frontiers in Physics* 10:931596.

Fried, Stephie, Kevin Michael Novan, and William Peterman. 2021. The macro effects of climate policy uncertainty .

Friedl, Aleksandra, Felix Kubler, Simon Scheidegger, and Takafumi Usui. 2023. Deep Uncertainty Quantification: With an Application to Integrated Assessment Models. *Working Paper* .

Friedlingstein, Pierre. 2015. Carbon cycle feedbacks and future climate change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 373 (20140421).

Gao, X. and L.-M. Duan. 2017. Efficient representation of quantum many-body states with deep neural networks. *Nature Communications* 8 (1):662.

Geoffroy, O, D Saint-Martin, D J L Olivié, A Voldoire, G Bellon, and S Tytéca. 2013. Transient Climate Response in a Two-Layer Energy-Balance Model. Part {I}: Analytical Solution and Parameter Calibration Using CMIP5 AOGCM Experiments. *Journal of Climate* 26 (6):1841–1857.

Gillett, Nathan P. 2023. Warming proportional to cumulative carbon emissions not explained by heat and carbon sharing mixing processes. *Nature Communications* 14 (1):6466.

Gnoatto, Alessandro, Marco Patacca, and Athena Picarelli. 2022. A deep solver for BSDEs with jumps. *arXiv preprint arXiv:2211.04349* .

Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski. 2014. Optimal Taxes on Fossil Fuel in General Equilibrium. *Econometrica* 82 (1):41–88.

Gupta, Maya, Andrew Cotter, Jan Pfeifer, Konstantin Voevodski, Kevin Canini, Alexander Mangylov, Wojciech Moczydlowski, and Alexander Van Esbroeck. 2016. Monotonic calibrated interpolated look-up tables. *The Journal of Machine Learning Research* 17 (1):3790–3836.

Han, J., A. Jentzen, and W. E. 2018. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences* 115 (34):8505–8510.

Han, J., C. Ma, Z. Ma, and W. E. 2019. Uniformly accurate machine learning-based hydrodynamic models for kinetic equations. *Proceedings of the National Academy of Sciences* 116 (44):21983–21991.

Han, Jiequn and Weinan E. 2016. Deep Learning Approximation for Stochastic Control Problems. *Deep Reinforcement Learning Workshop, arXiv preprint arXiv:1611.07422* .

Han, Jiequn and Ruimeng Hu. 2021. Recurrent neural networks for stochastic control problems with delay. *Mathematics of Control, Signals, and Systems* 33 (4):775–795.

Han, Jiequn and Jihao Long. 2020. Convergence of the deep BSDE method for coupled FBSDEs. *Probability, Uncertainty and Quantitative Risk* 5:1–33.

Hansen, James E, Pushker Kharecha, Makiko Sato, George Tselioudis, Joseph Kelly, Susanne E Bauer, Reto Ruedy, Eunbi Jeong, Qinjian Jin, Eric Rignot, et al. 2025. Global Warming Has Accelerated: Are the United Nations and the Public Well-Informed? *Environment: Science and Policy for Sustainable Development* 67 (1):6–44.

Hansen, Lars Peter and Jianjun Miao. 2018. Aversion to Ambiguity and Model Misspecification in Dynamic Stochastic Environments. *Proceedings of the National Academy of Sciences* 115 (37):9163–9168.

Hansen, Lars Peter and Thomas J. Sargent. 2007. Recursive Robust Estimation and Control without Commitment. *Journal of Economic Theory* 136 (1):1–27.

Hansen, Lars Peter and Thomas J Sargent. 2011. Robustness and ambiguity in continuous time. *Journal of Economic Theory* 146 (3):1195–1223.

Hasselmann, Klaus. 1976. Stochastic climate models part I. Theory. *tellus* 28 (6):473–485.

Hassler, John, Per Krusell, and Conny Olovsson. 2018. The Consequences of Uncertainty: Climate Sensitivity and Economic Sensitivity to the Climate. *Annual Review of Economics* 10:189–205.

Hsiang, Solomon, Robert Kopp, Amir Jina, James Rising, Michael Delgado, Shashank Mohan, DJ Rasmussen, Robert Muir-Wood, Paul Wilson, Michael Oppenheimer, et al. 2017. Estimating Economic Damage from Climate Change in the United States. *Science* 356 (6345):1362–1369.

Huré, Côme, Huyên Pham, and Xavier Warin. 2020. Deep backward schemes for high-dimensional nonlinear PDEs. *Mathematics of Computation* 89 (324):1547–1579.

Hutzenthaler, Martin, Arnulf Jentzen, Thomas Kruse, et al. 2019. On multilevel Picard numerical approximations for high-dimensional nonlinear parabolic partial differential equations and high-dimensional nonlinear backward stochastic differential equations. *Journal of Scientific Computing* 79 (3):1534–1571.

Jaakkola, Niko and Frederick Van der Ploeg. 2019. Non-cooperative and cooperative climate policies with anticipated breakthrough technology. *Journal of Environmental Economics and Management* 97:42–66.

Jenkins, JD, EN Mayfield, ED Larson, SW Pacala, and C Greig. 2021. Mission net-zero America: The nation-building path to a prosperous, net-zero emissions economy.

Jermann, Urban J. 1998. Asset pricing in production economies. *Journal of monetary Economics* 41 (2):257–275.

Ji, Shaolin, Shige Peng, Ying Peng, and Xichuan Zhang. 2020. Three algorithms for solving high-dimensional fully coupled FBSDEs through deep learning. *IEEE Intelligent Systems* 35 (3):71–84.

Joos, F., R. Roth, J. S. Fuglestvedt, G. P. Peters, I. G. Enting, W. Von Bloh, V. Brovkin, E. J. Burke, M. Eby, N. R. Edwards, T. Friedrich, T. L. Frölicher, P. R. Halloran, P. B. Holden, C. Jones, T. Kleinen, F. T. Mackenzie, K. Matsumoto, M. Meinshausen, G. K. Plattner, A. Reisinger, J. Segschneider, G. Shaffer, M. Steinacher, K. Strassmann, K. Tanaka, A. Timmermann, and A. J. Weaver. 2013. Carbon Dioxide and Climate Impulse Response Functions for the Computation of Greenhouse Gas Metrics: A Multi-Model Analysis. *Atmospheric Chemistry and Physics* 13 (5):2793–2825.

Kingma, Diederik P and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* .

Klibanoff, P, M Marinacci, and S Mukerji. 2009. Recursive {Smooth} {Ambiguity} {Preferences}. *Journal of Economic Theory* 144:930–976.

Kogan, Leonid and Dimitris Papanikolaou. 2014. Growth opportunities, technology shocks, and asset prices. *The journal of finance* 69 (2):675–718.

Kovachki, Nikola B, Samuel Lanthaler, and Andrew M Stuart. 2024. Operator learning: Algorithms and analysis. *arXiv preprint arXiv:2402.15715* .

Kung, Howard and Lukas Schmid. 2015. Innovation, growth, and asset prices. *The Journal of Finance* 70 (3):1001–1037.

Lemoine, Derek and Christian Traeger. 2014. Watch your step: optimal policy in a tipping climate. *American Economic Journal: Economic Policy* 6 (1):137–66.

Lenton, Timothy M, Hermann Held, Elmar Kriegler, Jim W Hall, Wolfgang Lucht, Stefan Rahmstorf, and Hans Joachim Schellnhuber. 2008. Tipping Elements in the Earth's Climate System. *Proceedings of the National Academy of Sciences* 105 (6):1786–1793.

Levitan, David. 2013. Quick-Change Planet: Do Global Climate Tipping Points Exist? *Scientific American* .

Li, Zongyi, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. 2020. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895* .

Lu, Liwei, Ruimeng Hu, Xu Yang, and Yi Zhu. 2024. Multi-Agent Relative Investment Games in a Jump Diffusion Market with Deep Reinforcement Learning Algorithm. *arXiv preprint arXiv:2404.11967* .

Lucas Jr, Robert E. 1988. On the mechanics of economic development. *Journal of monetary economics* 22 (1):3–42.

Lucking, Brian, Nicholas Bloom, and John Van Reenen. 2019. Have R&D spillovers declined in the 21st century? *Fiscal Studies* 40 (4):561–590.

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini. 2006. Dynamic Variational Preferences. *Journal of Economic Theory* 128 (1):4–44.

MacDougall, Andrew H., Neil C. Swart, and Reto Knutti. 2017. The Uncertainty in the Transient Climate Response to Cumulative $CO_2$ Emissions Arising from the Uncertainty in Physical Climate Parameters. *Journal of Climate* 30 (2):813–827.

Maliar, Lilia, Serguei Maliar, and Pablo Winant. 2021. Deep learning for solving dynamic economic models. *Journal of Monetary Economics* 122:76–101.

Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou. 2021a. Summary for Policymakers - Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Book section, IPCC, Cambridge, United Kingdom and New York, NY, USA.

Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B.u Zhou. 2021b. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Tech. rep., IPCC.

Matthews, H Damon, Nathan P Gillett, Peter A Stott, and Kirsten Zickfeld. 2009. The proportionality of global warming to cumulative carbon emissions. *Nature* 459 (7248):829–832.

McKinsey Global Initiative. 2022. The net-zero transition: What it would cost, what it could bring. Mckensey Global Institute.

Nordhaus, William. 2018. Projections and uncertainties about climate change in an era of minimal climate policies. *American Economic Journal: Economic Policy* 10 (3):333–360.

———. 2019. Economics of the disintegration of the Greenland ice sheet. *Proceedings of the National Academy of Sciences* 116 (25):12261–12269.

Norets, Andriy. 2012. Estimation of dynamic discrete choice models using artificial neural network approximations. *Econometric Reviews* 31 (1):84–106.

OECD. 2019. *Regions in industrial transition: Policies for people and places.* ORGANIZATION FOR ECONOMIC.

Olson, Roman, Ryan Sriver, Marlos Goes, Nathan M. Urban, H. Damon Matthews, Murali Haran, and Klaus Keller. 2012. A Climate Sensitivity Estimate using Bayesian Fusion of Instrumental Observations and an Earth System Model. *Journal of Geophysical Research Atmospheres* 117 (D04103):1–11.

Palmer, Tim and Bjorn Stevens. 2019. The scientific challenge of understanding and estimating climate change. *Proceedings of the National Academy of Sciences* 116 (49):24390–24395.

Papanikolaou, Dimitris. 2011. Investment shocks and asset prices. *Journal of Political Economy* 119 (4):639–685.

Payne, Jonathan, Adam Rebei, and Yucheng Yang. 2024. Deep Learning for Search and Matching Models. *Available at SSRN* .

Pham, Huyen, Xavier Warin, and Maximilien Germain. 2021. Neural networks-based backward scheme for fully nonlinear PDEs. *SN Partial Differential Equations and Applications* 2 (1):16.

Pierrehumbert, Ramond T. 2014. Short-Lived Climate Pollution. *Annual Review of Earth and Planetary Science* 42:341–379.

Pörtner, H. O., D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, and B. Rama. 2022. Summary for Policymakers - Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Book section, IPCC, Cambridge, UK. In Press.

Raissi, Maziar, Paris Perdikaris, and George E Karniadakis. 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* 378:686–707.

Reguant, Mar. 2021. Comment on "Climate Change Uncertainty Spillover in the Macroeconomy". Tech. rep., National Bureau of Economic Research.

Ricke, Katharine L. and Ken Caldeira. 2014. Maximum Warming Occurs about One Decade After a Carbon Dioxide Emission. *Environmental Research Letters* 9 (12):1–8.

Ritchie, Paul DL, Joseph J Clarke, Peter M Cox, and Chris Huntingford. 2021. Overshooting tipping point thresholds in a changing climate. *Nature* 592 (7855):517–523.

Rogelj, Joeri, Alexander Popp, Katherine V Calvin, Gunnar Luderer, Johannes Emmerling, David Gernaat, Shinichiro Fujimori, Jessica Strefler, Tomoko Hasegawa, Giacomo Marangoni, et al. 2018. Scenarios towards limiting global mean temperature increase below 1.5 C. *Nature Climate Change* 8 (4):325–332.

Rogelj, Joeri, Daniel Huppmann, Volker Krey, Keywan Riahi, Leon Clarke, Matthew Gidden, Zebedee Nicholls, and Malte Meinshausen. 2019. A new scenario logic for the Paris Agreement long-term temperature goal. *Nature* 573 (7774):357–363.

Romer, Paul M. 1990. Endogenous technological change. *Journal of political Economy* 98 (5, Part 2):S71–S102.

Rudik, Ivan. 2020. Optimal Climate Policy When Damages Are Unknown. *American Economic Journal: Economic Policy* 12 (2):340–73.

Runje, Davor and Sharath M Shankaranarayana. 2022. Constrained Monotonic Neural Networks. *arXiv preprint arXiv:2205.11775* .

Saporito, Yuri F and Zhaoyu Zhang. 2021. Path-dependent deep Galerkin method: A neural network approach to solve path-dependent partial differential equations. *SIAM Journal on Financial Mathematics* 12 (3):912–940.

Sauzet, Maxime. 2021. Projection methods via neural networks for continuous-time models. *Available at SSRN 3981838* .

Sill, Joseph. 1997. Monotonic networks. *Advances in neural information processing systems* 10.

Sirignano, J. and K. Spiliopoulos. 2018. DGM: A deep learning algorithm for solving partial differential equations. *Journal of computational physics* 375:1339–1364.

Stine, Deborah D. 2008. *The Manhattan Project, the Apollo program, and federal energy technology R & D programs: A comparative analysis*. Library of Congress.

Waidelich, Paul, Fulden Batibeniz, James Rising, Jarmo S Kikstra, and Sonia I Seneviratne. 2024. Climate damage projections beyond annual temperature. *Nature Climate Change* 14 (6):592–599.

Wehenkel, Antoine and Gilles Louppe. 2019. Unconstrained monotonic neural networks. *Advances in neural information processing systems* 32.

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

White House. 2021. The long-term strategy of the United States: pathways to net-zero greenhouse gas emissions by 2050.

Zhang, L., J. Han, H. Wang, R. Car, and W. E. 2018. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Physical Review Letters* 120 (14):143001.

# Appendix A Example Model HJB Equations

## A.1 Post Damage and Technology Jumps HJB Equation

The first jump-state we solve is for the post-technology jump, post-damage-function-jump state. We denote this value function by $V^{(\ell,\ell'')}(K_d, K_g, \hat{Y})$, indicating that we are at a given realization of $\lambda_3(\ell)$ and in the terminal technology state with green capital productivity $A_g(\ell'')$. The HJB equation characterizing the value function is given by:

$$
\begin{aligned}
0 = &\max_{i^g, i^d} \min_{h} \delta \log([A_d - i^d](1-Z) + [A_g(\ell'') - i^g]Z) + \delta \log K - \delta \log N(y) - \delta V^{(\ell,\ell'')} \\
&+ \left((1-Z)\phi_d(i^d) + Z\phi_g(i^g) - \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2}\right) V_{\log K}^{(\ell,\ell'')} + \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} V_{\log K, \log K}^{(\ell,\ell'')} \\
&+ \left(\phi_g(i^g) - Z\sigma_g^2 - \phi_d(i^d) + (1-Z)\sigma_d^2\right) Z(1-Z) V_Z^{(\ell,\ell'')} + \frac{1}{2} Z^2 (1-Z)^2 (\sigma_g^2 + \sigma_d^2) V_{ZZ}^{(\ell,\ell'')} \\
&+ \left[-Z(1-Z)^2\sigma_d^2 + Z^2(1-Z)\sigma_g^2\right] V_{\log K, Z}^{(\ell,\ell'')} + V_y^{(\ell,\ell'')} \bar{\theta} \mathcal{E}_t + \frac{|\varsigma|^2 (\mathcal{E}_t)^2}{2} V_{yy}^{(\ell,\ell'')}
\end{aligned}
$$

Taking first order conditions with respect to $i^d$ and $i^g$, we end up with the following two expressions for optimal investment choices

$$
\begin{aligned}
\delta([A_d - i^d](1-R) + [A_g'(\ell') - i^g]Z)^{-1} &= \Gamma_d \theta_d (1 + \theta_d i^d)^{-1} [V_{\log K}^{(\ell,\ell')} - Z V_Z^{(\ell,\ell')}], \\
\delta([A_d - i^d](1-Z) + [A_g'(\ell') - i^g]Z)^{-1} &= \Gamma_g \theta_g (1 + \theta_g i^g)^{-1} [V_{\log K}^{(\ell,\ell')} + (1-Z) V_Z^{(\ell,\ell')}].
\end{aligned}
$$

Each expression equates the marginal benefit of an additional unit of a given type of capital to the marginal utility of consumption, which is the utility loss of forgoing consumption for an additional unit of investment.

## A.2 Intermediate Jump State HJB Equations

The next HJB equations come from the intermediate jump states. The first case is breakthrough technology jump realization, pre-damage function jump realization. We denote this value function by $V^{(\ell'')}(K_d, K_g, Y)$, indicating that we have not had a realization of $\lambda_3(\ell)$, but we are in the terminal technology state with green capital productivity $A_g(\ell'')$. The HJB

equation in this case includes one adjustments relative to the post-technology, post-damage jump case, a term capturing the expectation of the damage function jump, given by

$$\sum_{\ell=1}^{L} \mathcal{J}_n^{\ell}(y)(V^{(\ell,\ell'')} - V^{(\ell'')}).$$

Expressions for the first order conditions with respect to $i^d$ and $i^g$ are unchanged from the post-damage, terminal technology jump case, equating the marginal benefit of an additional unit of a given type of capital to the utility loss of forgoing consumption for an additional unit of investment, which is the marginal utility of consumption.

The second and third intermediate cases are the pre-technology jump, post-damage function jump and intermediate technology jump, post-damage function jump settings. We denote these value functions by $V^{(\ell)}(K_d, K_g, Y, \log R)$ and $V^{(\ell,\ell')}(K_d, K_g, Y, \log R)$, respectively. Relative to the terminal technology, post-damage jump case, we introduce terms for the following features: the evolution of the knowledge capital accounting for the possibility of R&D investment to increase the knowledge capital stock, and the expectation of the green technology jump realization altering green capital productivity to $A_g(\ell'')$:

$$\left(\psi_r(i_r) - \frac{1}{2}|\sigma_r|^2\right) V^{(\ell)}_{\log R} + \frac{|\sigma_r|^2}{2} V^{(\ell)}_{\log R, \log R} + \mathcal{J}_g^{\ell'}(R)[V^{(\ell,\ell')} - V^{(\ell)}] + \mathcal{J}_g^{\ell''}(R)[V^{(\ell,\ell'')} - V^{(\ell)}].$$

if starting from the pre-technology jump state or

$$\left(\psi_r(i_r) - \frac{1}{2}|\sigma_r|^2\right) V^{(\ell,\ell')}_{\log R} + \frac{|\sigma_r|^2}{2} V^{(\ell,\ell')}_{\log R, \log R} + \mathcal{J}_g^{\ell''}(R)[V^{(\ell,\ell'')} - V^{(\ell,\ell')}].$$

if starting from the intermediate technology state.

In addition, allowing for R&D investment alters the output constraint in this setting, so that our flow utility term net of climate damages is now given by

$$\delta \log([A_d - i^d](1 - Z) + [A_g - i^g]Z - i^r).$$

if starting from the pre-technology jump state or

$$\delta \log([A_d - i^d](1 - Z) + [A_g(\ell') - i^g]Z - i^r).$$

if starting from the intermediate technology state. We now take first order conditions with

respect to $i^d$, $i^g$, and $i^r$, which give us expressions for optimal investment choices:

$$\begin{aligned}
\delta([A_d - i^d](1-Z) + [A_g - i^g]Z - i^r)^{-1} &= \phi_d'(i^d)[V^{(\ell)}_{\log K} - ZV^{(\ell)}_Z], \\
\delta([A_d - i^d](1-Z) + [A_g - i^g]Z - i^r)^{-1} &= \phi_g'(i^g)[V^{(\ell)}_{\log K} + (1-Z)V^{(\ell)}_Z], \\
\delta([A_d - i^d](1-Z) + [A_g - i^g]ZF - i^r)^{-1} &= \psi_r'(i^r)V^{(\ell)}_{\log R}.
\end{aligned}$$

if starting from the pre-technology jump state or

$$\begin{aligned}
\delta([A_d - i^d](1-Z) + [A_g(\ell') - i^g]Z - i^r)^{-1} &= \phi_d'(i^d)V^{(\ell,\ell')}_{\log K} - ZV^{(\ell,\ell')}_Z], \\
\delta([A_d - i^d](1-Z) + [A_g(\ell') - i^g]Z - i^r)^{-1} &= \phi_g'(i^g)[V^{(\ell,\ell')}_{\log K} + (1-Z)V^{(\ell,\ell')}_Z], \\
\delta([A_d - i^d](1-Z) + [A_g(\ell') - i^g]Z - i^r)^{-1} &= \psi_r'(i^r)V^{(\ell,\ell')}_{\log R}.
\end{aligned}$$

if starting from the intermediate technology state. As before, these expressions equate the marginal benefit of an additional unit of a given type of capital to the marginal utility of consumption, which is the utility loss of forgoing consumption for an additional unit of investment. However, given the altered output constraint expression, the marginal utility of consumption on the left-hand side of these equations now also incorporates the fact that some output is dedicated to R&D investment instead of consumption.

## A.3 Pre Damage and Technology Jumps HJB Equation

Finally, we can write down the value function associated with the pre-technology, pre-damage jump state. We denote this value function by $V(K_d, K_g, Y, \log R)$. The HJB equation incorporates each of the changes applied to the intermediate jump state cases: terms accounting for the evolution of the knowledge capital stock, the expectation of the technological innovation jump realizations, and the expectation of the damage function jump realization, as well as an altered output constraint to account for R&D investment. The first order conditions with respect to $i^d$, $i^g$, and $i^\kappa$ match those of the pre-technology, post-damage intermediate jump state, except that the relevant value function derivatives now pertain to $v$. As in each case, the FOCs equate the marginal benefit of an additional unit of a given type of capital to the marginal utility of consumption, which is the utility loss of forgoing consumption for an additional unit of investment.

# Online Appendix

This online appendix is provided in support of the paper "A Deep Learning Analysis of Climate Change, Innovation, and Uncertainty." Included here are additional details on theoretical derivations, the model parameters, and the numerical methods used in the paper.

# Appendix B Model Uncertainty Details

We now detail the introduction uncertainty aversion in the form of model misspecification across the following channels of interest in our analysis:

- Climate Sensitivity to Carbon Emissions
- Dirty, Green, and Knowledge Capital Stocks and Investments
- Climate Damage Function Severity and Arrival
- "Green" Technological Change Magnitude and Arrival

Aversion to these different channels of uncertainty are introduced into our baseline model using the tools of dynamic decision theory. Importantly, the planner in our model is wary of model uncertainty of these different forms and channels, and thus adopts a preference structure to identify potential worst-case models, constrained by a penalization function, and make optimal policy decisions that are robust to these possible outcomes. We outline the different forms of model uncertainty for our framework in what follows, focusing specifically on misspecification concerns about jump and diffusion processes.

## B.1 Jump Misspecification Concerns

We first consider the setting with misspecification concerns about the damage function and technological change jump channels. Before the jumps occur, the relevant contributions to the HJB equation related to the technological change and damage function jumps are

$$\mathcal{J}_g^{\ell'}(R)(V^{(\ell')} - V) + \mathcal{J}_g^{\ell''}(R)(V^{(\ell'')} - V) + \sum_{\ell=1}^{L} \mathcal{J}_n^{\ell}(y)(V^{(\ell)} - V).$$

Assuming a common robustness parameter $\xi$ across each channel, these terms are replaced

by the following uncertainty-adjusted contributions

$$\min_{g^{\ell},g^{\ell'},g^{\ell''}} \quad \mathcal{J}_g^{\ell'}(R)g^{\ell'}(V^{(\ell')}-V)+\mathcal{J}_g^{\ell''}(R)g^{\ell''}(V^{(\ell'')}-V)+\sum_{\ell=1}^{L}\mathcal{J}_n^{\ell}(y)g^{\ell}(V^{(\ell)}-V)$$

$$+\xi\left[\mathcal{J}_g^{\ell'}(R)\left(1-g^{\ell'}+g^{\ell'}\log g^{\ell'}\right)+\mathcal{J}_g^{\ell''}(R)\left(1-g^{\ell''}+g^{\ell''}\log g^{\ell''}\right)+\sum_{\ell=1}^{L}\mathcal{J}_n^{\ell}(y)\left(1-g^{\ell}+g^{\ell}\log g^{\ell}\right)\right].$$

The adjustment introduces probability distortions $g^{\ell}$, $g^{\ell'}$, and $g^{\ell''}$ to the jump processes for technological change and climate damages. These distortions impact expectations about the arrival rate and distribution of post-jump outcomes related to these jumps. To constrain these distortions, the relative entropy terms

$$\xi\left[\mathcal{J}_g^{\ell'}(R)\left(1-g^{\ell'}+g^{\ell'}\log g^{\ell'}\right)+\mathcal{J}_g^{\ell''}(R)\left(1-g^{\ell''}+g^{\ell''}\log g^{\ell''}\right)+\sum_{\ell=1}^{L}\mathcal{J}_n^{\ell}(y)\left(1-g^{\ell}+g^{\ell}\log g^{\ell}\right)\right]$$

are introduced into preferences, penalizing distortions that are "too large" based on the uncertainty aversion parameter $\xi$. Note that $g^{\ell}$, $g^{\ell'}$, and $g^{\ell''}$ are endogenous objects solved for by the decision maker by minimizing discounted lifetime expected utility, subject to the relative entropy penalties. The FOCs resulting from this minimization objective are

$$\begin{aligned} g^{\ell} &= \exp\left(-\frac{1}{\xi}(V^{(\ell)}-V)\right), \\ g^{\ell'} &= \exp\left(-\frac{1}{\xi}(V^{(\ell')}-V)\right), \\ g^{\ell''} &= \exp\left(-\frac{1}{\xi}(V^{(\ell'')}-V)\right). \end{aligned}$$

Each of these represents probability distortions that adjust the distributions associated with the jump processes related to the economic and climate channels for which there are concerns about model uncertainty. The inclusion of these probability distortions into the planner's problem leads to robustly-altered optimal policy choices that account for the distorted likelihood of outcomes based on potential worst-case outcomes.

We note that in each of the various jump realization states, there are modifications that need to be made to these components. After the breakthrough technology and damage jumps have both taken place, jump uncertainty is no longer relevant for the social planner. After the breakthrough technology jump has occurred, but not the damage jump, we replace $A_g$ or $A_g(\ell')$ with $A_g(\ell'')$ and the relevant uncertainty adjustments to the HJB equation and

minimization FOC pertaining to the damage function jump are given by

$$\min_{g^\ell} \sum_{\ell=1}^{L} \mathcal{J}_n^\ell(y) g^\ell (V^{(\ell,\ell'')} - V^{(\ell'')}) + \xi \sum_{\ell=1}^{L} \mathcal{J}_n^\ell(y) \left(1 - g^\ell + g^\ell \log g^\ell\right),$$
$$g^\ell = \exp\left(-\frac{1}{\xi}(V^{(\ell,\ell'')} - V^{(\ell'')})\right).$$

After an intermediate technology jump has occurred, but not the damage jump, we replace $A_g$ with $A_g(\ell')$ and the relevant uncertainty adjustments to the HJB equation and minimization FOC pertaining to the technological change and damage function jumps are

$$\min_{g^\ell, g^{\ell''}} \quad \mathcal{J}_g^{\ell''}(R) g^{\ell''}(V^{(\ell'')} - V^{(\ell')}) + \sum_{\ell=1}^{L} \mathcal{J}_n^\ell(y) g^\ell (V^{(\ell,\ell')} - V^{(\ell')})$$
$$+\xi \left[\mathcal{J}_g^{\ell''}(R) \left(1 - g^{\ell''} + g^{\ell''} \log g^{\ell''}\right) + \sum_{\ell=1}^{L} \mathcal{J}_n^\ell(y) \left(1 - g^\ell + g^\ell \log g^\ell\right)\right],$$

$$\begin{aligned} g^\ell &= \exp\left(-\tfrac{1}{\xi}(V^{(\ell,\ell')} - V^{(\ell')})\right), \\ g^{\ell''} &= \exp\left(-\tfrac{1}{\xi}(V^{(\ell'')} - V^{(\ell')})\right). \end{aligned}$$

Before a technology jump has occurred, but after the damage jump, the relevant uncertainty adjustments to the HJB equation and minimization FOC pertaining to the technology jump are as follows

$$\min_{g^{\ell'}, g^{\ell''}} \quad \mathcal{J}_g^{\ell'}(R) g^{\ell'}(V^{(\ell,\ell')} - V^{(\ell)}) + \mathcal{J}_g^{\ell''}(R) g^{\ell''}(V^{(\ell,\ell'')} - V^{(\ell)})$$
$$+\xi \left[\mathcal{J}_g^{\ell'}(R) \left(1 - g^{\ell'} + g^{\ell'} \log g^{\ell'}\right) + \mathcal{J}_g^{\ell''}(R) \left(1 - g^{\ell''} + g^{\ell''} \log g^{\ell''}\right)\right],$$

$$\begin{aligned} g^{\ell'} &= \exp\left(-\tfrac{1}{\xi}(V^{(\ell,\ell')} - V^{(\ell)})\right), \\ g^{\ell''} &= \exp\left(-\tfrac{1}{\xi}(V^{(\ell,\ell'')} - V^{(\ell)})\right). \end{aligned}$$

## B.2 Diffusion Misspecification Concerns

We now consider the setting with misspecification concerns about the climate and capital stocks diffusion channels. In th pre-technology, pre-damage function justmp state the

relevant contributions to the HJB equation for the diffusion components are of the form

$$\begin{aligned}
&\left((1-Z)\phi_d(i^d) + Z\phi_g(i^g) - \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2}\right) V_{\log K} \\
&+ \left(\phi_d(i^d) - Z\sigma_g^2 - \phi_g(i^g) + (1-Z)\sigma_d^2\right) Z(1-Z)V_Z \\
&+ V_y\bar{\theta}\mathcal{E}_t + \left(\psi_r(i^r) - \frac{1}{2}|\sigma_r|^2\right) V_{\log R}.
\end{aligned}$$

Assuming the same common robustness parameter $\xi$ for each channel as before, these terms are replaced by the following uncertainty-adjusted contributions

$$\begin{aligned}
\min_h\ &\left((1-Z)\phi_d(i^d) + Z\phi_g(i^g) - \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2}\right) V_{\log K} \\
&+ \left(\phi_g(i^g) - Z\sigma_g^2 - \phi_d(i^d) + (1-Z)\sigma_d^2\right) Z(1-Z)V_Z \\
&+ \left(\left(V_{\log K} - ZV_Z\right)(1-Z)\sigma_d + \left(V_{\log K} + (1-Z)V_Z\right) Z\sigma_g\right) \cdot h \\
&+ V_y(\bar{\theta} + \varsigma \cdot h)\mathcal{E}_t + \left(\psi_r(i^r) - \frac{1}{2}|\sigma_r|^2 + \sigma_r \cdot h\right) V_{\log R} + \xi\frac{|h|^2}{2}.
\end{aligned}$$

This adjustment introduces a drift distortion $h$ to the diffusion processes for the different types of capitals and for temperature, disguised by the Brownian shocks. This type of drift distortion is a known result from the Girsanov theorem, and captures a distributional change represented by a likelihood ratio. To constrain this distortion, again, a relative entropy term is introduced into preferences, in this setting, the quadratic expression $\xi|h|^2/2$, penalizing distortions that are "too large" based on the uncertainty aversion parameter $\xi$. As with $g^\ell$ and $g^{\ell'}$, $h$ is an endogenous object solved by the decision maker by minimizing discounted lifetime expected utility, subject to the relative entropy penalties. The FOC resulting from this minimization objective is given by

$$h = -\frac{1}{\xi}[\{V_{\log K} - V_Z Z\}(1-Z), \{V_{\log K} + V_Z(1-Z)\}Z, V_{\log R}, V_y] \cdot [\sigma_d, \sigma_g, \sigma_r, \varsigma\mathcal{E}_t]'$$

Given the volatilities are assumed to be independent, we can separate $h$ into four different

scalar components for each dimension of the Brownian as follows:

$$\begin{aligned} h_d &= -\frac{1}{\xi}\{V_{\log K} - V_Z Z\}(1-Z)\cdot\sigma_d \\ h_g &= -\frac{1}{\xi}\{V_{\log K} + V_Z(1-Z)\}Z\cdot\sigma_g \\ h_r &= -\frac{1}{\xi}V_r\cdot\sigma_r \\ h_y &= -\frac{1}{\xi}V_y\mathcal{E}_t\cdot\varsigma \end{aligned}$$

The inclusion into the planner's problem of this probability distortion in the form of a drift distortion leads to further robustly-altered optimal policy choices that account for the distorted likelihood of outcomes based on potential worst-case outcomes.

As with the jump misspecification concerns, in each of the various jump realization states there are modifications that need to be made to these components. In the post-technology, post-damage function jump state the remaining uncertainty implications come from the climate sensitivity channel and the green and dirty capital stocks. Therefore, in addition to replacing $A_g$ with $A_g'(\ell')$, and denoting the realized $\lambda_3(\ell)$ value, the uncertainty-adjusted contribution to the HJB and the relevant minimization FOC are given by

$$\begin{aligned} \min_h\ &\left((1-Z)\phi_d(i^d) + Z\phi_g(i^g) - \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2}\right)V_{\log K}^{(\ell,\ell')} \\ &+ \left(\phi_g(i^g) - Z\sigma_g^2 - \phi_d(i^d) + (1-Z)\sigma_d^2\right)Z(1-Z)V_Z^{(\ell,\ell')} \\ &+ \left(\left(V_{\log K}^{(\ell,\ell')} - ZV_Z^{(\ell,\ell')}\right)(1-Z)\sigma_d + \left(V_{\log K}^{(\ell,\ell')} + (1-Z)V_Z^{(\ell,\ell')}\right)Z\sigma_g\right)\cdot h \\ &+ V_y^{(\ell,\ell')}(\bar\theta + \varsigma\cdot h)\mathcal{E}_t + \xi\frac{|h|^2}{2}, \\ h = -\frac{1}{\xi}&[\{V_{\log K}^{(\ell,\ell')} - V_Z^{(\ell,\ell')}Z\}(1-Z), \{V_{\log K}^{(\ell,\ell')} + V_Z^{(\ell,\ell')}(1-Z)\}Z, V_y^{(\ell,\ell')}]\cdot[\sigma_d, \sigma_g, \varsigma\mathcal{E}_t]' \end{aligned}$$

In the post-technology, pre-damage function jump state we replace $A_g$ with $A_g'(\ell)$ and the relevant uncertainty adjustments to the HJB equation and minimization FOC pertaining

to the diffusion terms are given by

$$\begin{aligned}
\min_h \ & \left((1-Z)\phi_d(i^d) + Z\phi_g(i^g) - \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2}\right) V^{(\ell')}_{\log K} \\
& + \left(\phi_g(i^g) - Z\sigma_g^2 - \phi_d(i^d) + (1-Z)\sigma_d^2\right) Z(1-Z) V_Z^{(\ell')} \\
& + \left(\left(V^{(\ell')}_{\log K} - ZV_Z^{(\ell')}\right)(1-Z)\sigma_d + \left(V^{(\ell')}_{\log K} + (1-Z)V_Z^{(\ell')}\right) Z\sigma_g\right) \cdot h \\
& + V_y^{(\ell')}(\bar{\theta} + \varsigma \cdot h)\mathcal{E}_t + \xi \frac{|h|^2}{2}, \\
h = & -\frac{1}{\xi}[\{V^{(\ell')}_{\log K} - V_Z^{(\ell')} Z\}(1-Z), \{V^{(\ell')}_{\log K} + V_Z^{(\ell')}(1-Z)\}Z, V_y^{(\ell')}] \cdot [\sigma_d, \sigma_g, \varsigma \mathcal{E}_t]'
\end{aligned}$$

In the pre-technology, post-damage function jump state the relevant uncertainty adjustments to the HJB equation and minimization FOC pertaining to the diffusion terms are

$$\begin{aligned}
\min_h \ & \left((1-Z)\phi_d(i^d) + Z\phi_g(i^g) - \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2}\right) V^{(\ell)}_{\log K} \\
& + \left(\phi_g(i^g) - Z\sigma_g^2 - \phi_d(i^d) + (1-Z)\sigma_d^2\right) Z(1-Z) V_Z^{(\ell)} \\
& + \left(\left(V^{(\ell)}_{\log K} - ZV_Z^{(\ell)}\right)(1-Z)\sigma_d + \left(V^{(\ell)}_{\log K} + (1-Z)V_Z^{(\ell)}\right) Z\sigma_g\right) \cdot h \\
& + V_y^{(\ell)}(\bar{\theta} + \varsigma \cdot h)\mathcal{E}_t \\
& + \left(\psi_r(i^r) - \frac{1}{2}|\sigma_r|^2 + \sigma_r \cdot h\right) V^{(\ell)}_{\log R} + \xi \frac{|h|^2}{2}, \\
h = & -\frac{1}{\xi}[\{V^{(\ell)}_{\log K} - V_Z^{(\ell)} Z\}(1-Z), \{V^{(\ell)}_{\log K} + V_Z^{(\ell)}(1-Z)\}Z, V^{(\ell)}_{\log R}, V_y^{(\ell)}] \cdot [\sigma_d, \sigma_g, \sigma_r, \varsigma \mathcal{E}_t]'
\end{aligned}$$

## B.3 Full Misspecification Concerns

We have so far shown the introduction of the jump and diffusion uncertainty separately in our analysis. However, in our analysis, we incorporate each of the uncertainty channels simultaneously, allowing for broadly conceived uncertainty considerations in our results. There are explicit contributions related to diffusion uncertainty and jump uncertainty in the model. However, allowing for all of the uncertainty channels simultaneously allows for important interaction effects as the jump misspecification distortions $g^\ell$, $g^{\ell'}$, and $g^{\ell''}$, and the diffusion misspecification distortion $h$, will alter the value functions and their derivatives across jump states and states of nature. These more implicit effects will influence the optimized distortions arising from uncertainty aversion, and as a result, impact the social valuations and optimal policy choices of the social planner.

# Appendix C Full HJB Equations

## C.1 Post Damage and Technology Jumps Setting

The HJB equation for the post damage and technology jumps setting is given by

$$
\begin{aligned}
0 = & \max_{i^g, i^d} \min_{h} \delta \log([A_d - i^d](1-Z) + [A_g(\ell'') - i^g]Z) + \delta \log K - \delta \log N(y;\ell) - \delta V^{(\ell,\ell'')} \\
& + \left( (1-Z)\phi_d(i^d) + Z\phi_g(i^g) - \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} \right) V_{\log K}^{(\ell,\ell'')} + \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} V_{\log K, \log K}^{(\ell,\ell'')} \\
& + \left( \phi_g(i^g) - Z\sigma_g^2 - \phi_d(i^d) + (1-Z)\sigma_d^2 \right) Z(1-Z) V_Z^{(\ell,\ell'')} + \frac{1}{2} Z^2 (1-Z)^2 (\sigma_g^2 + \sigma_d^2) V_{ZZ}^{(\ell,\ell'')} \\
& + \left( \left( V_{\log K}^{(\ell,\ell'')} - Z V_Z^{(\ell,\ell'')} \right) (1-Z)\sigma_d + \left( V_{\log K}^{(\ell,\ell'')} + (1-Z) V_Z^{(\ell,\ell'')} \right) Z\sigma_g \right) \cdot h \\
& + \left[ -Z(1-Z)^2 \sigma_d^2 + Z^2 (1-Z)\sigma_g^2 \right] V_{\log K, Z}^{(\ell,\ell'')} \\
& + V_y^{(\ell,\ell'')} (\bar{\theta} + \varsigma \cdot h) \mathcal{E}_t + \frac{|\varsigma|^2 (\mathcal{E}_t)^2}{2} V_{yy}^{(\ell,\ell'')} + \xi \frac{|h|^2}{2}
\end{aligned}
$$

## C.2 Post Damage and Intermediate Technology Jumps Setting

The HJB equation for the post damage and Intermediate technology jumps setting is

$$
\begin{aligned}
0 = & \max_{i^g, i^d, i^r \ g^{\ell'}, h} \min \delta \log([A_d - i^d](1-Z) + [A_g(\ell') - i^g]Z - i_r) + \delta \log K - \delta \log N(y;\ell) - \delta V^{(\ell,\ell')} \\
& + \left( (1-Z)\phi_d(i^d) + Z\phi_g(i^g) - \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} \right) V_{\log K}^{(\ell)} + \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} V_{\log K, \log K}^{(\ell,\ell')} \\
& + \left( \phi_g(i^g - Z\sigma_g^2 - \phi_d(i^d) + (1-Z)\sigma_d^2 \right) Z(1-Z) V_Z^{(\ell,\ell')} + \frac{1}{2} Z^2 (1-Z)^2 (\sigma_g^2 + \sigma_d^2) V_{ZZ}^{(\ell,\ell')} \\
& + \left( \left( V_{\log K}^{(\ell,\ell')} - Z V_Z^{(\ell,\ell')} \right) (1-Z)\sigma_d + \left( V_{\log K}^{(\ell,\ell')} + (1-Z) V_Z^{(\ell,\ell')} \right) Z\sigma_g \right) \cdot h \\
& + \left[ -Z(1-Z)^2 \sigma_d^2 + Z^2 (1-Z)\sigma_g^2 \right] V_{\log K, Z}^{(\ell,\ell')} \\
& + V_y^{(\ell,\ell')} (\bar{\theta} + \varsigma \cdot h) \mathcal{E}_t + \frac{|\varsigma|^2 (\mathcal{E}_t)^2}{2} V_{yy}^{(\ell,\ell')} \\
& + \left( \psi_r(i^r) - \frac{1}{2}|\sigma_r|^2 + \sigma_r \cdot h \right) V_{\log R}^{(\ell,\ell')} + \frac{|\sigma_r|^2}{2} V_{\log R, \log R}^{(\ell,\ell')} + \xi \frac{|h|^2}{2} \\
& + \mathcal{J}_g^{\ell''}(R) g^{\ell''} (V^{(\ell,\ell'')} - V^{(\ell,\ell')}) + \xi \mathcal{J}_g^{\ell''}(R) \left( 1 - g^{\ell''} + g^{\ell''} \log g^{\ell''} \right).
\end{aligned}
$$

## C.3 Post Damage and Pre Technology Jumps Setting

The HJB equation for the post damage and pre technology jumps setting is given by

$$
\begin{aligned}
0 = \max_{i^g,i^d,i^r} \min_{g^{\ell'},h} & \delta \log([A_d - i^d](1-Z) + [A_g - i^g]Z - i_r) + \delta \log K - \delta \log N(y;\ell) - \delta V^{(\ell)} \\
& + \left( (1-Z)\phi_d(i^d) + Z\phi_g(i^g) - \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} \right) V^{(\ell)}_{\log K} + \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} V^{(\ell)}_{\log K, \log K} \\
& + \left( \phi_g(i^g - Z\sigma_g^2 - \phi_d(i^d) + (1-Z)\sigma_d^2 \right) Z(1-Z) V_Z^{(\ell)} + \frac{1}{2} Z^2 (1-Z)^2 (\sigma_g^2 + \sigma_d^2) V_{ZZ}^{(\ell)} \\
& + \left( \left( V^{(\ell)}_{\log K} - Z V_Z^{(\ell)} \right) (1-Z)\sigma_d + \left( V^{(\ell)}_{\log K} + (1-Z) V_Z^{(\ell)} \right) Z\sigma_g \right) \cdot h \\
& + \left[ -Z(1-Z)^2 \sigma_d^2 + Z^2 (1-Z)\sigma_g^2 \right] V^{(\ell)}_{\log K, Z} \\
& + V_y^{(\ell)} (\bar{\theta} + \varsigma \cdot h) \mathcal{E}_t + \frac{|\varsigma|^2 (\mathcal{E}_t)^2}{2} V_{yy}^{(\ell)} \\
& + \left( \psi_r(i^r) - \frac{1}{2}|\sigma_r|^2 + \sigma_r \cdot h \right) V^{(\ell)}_{\log R} + \frac{|\sigma_r|^2}{2} V^{(\ell)}_{\log R, \log R} + \xi \frac{|h|^2}{2} \\
& + \mathcal{J}_g^{\ell'}(R) g^{\ell'} (V^{(\ell,\ell')} - V^{(\ell)}) + \xi \mathcal{J}_g^{\ell'}(R) \left( 1 - g^{\ell'} + g^{\ell'} \log g^{\ell'} \right) \\
& + \mathcal{J}_g^{\ell''}(R) g^{\ell''} (V^{(\ell,\ell'')} - V^{(\ell)}) + \xi \mathcal{J}_g^{\ell''}(R) \left( 1 - g^{\ell''} + g^{\ell''} \log g^{\ell''} \right).
\end{aligned}
$$

## C.4 Pre Damage and Post Technology Jumps Setting

The HJB equation for the pre damage and post technology jumps setting is given by

$$
\begin{aligned}
0 = \max_{i^g,i^d} \min_{g^{\ell},h} & \delta \log([A_d - i^d](1-Z) + [A_g(\ell'') - i^g]Z - i^r) + \delta \log K - \delta \log N(y) - \delta V^{(\ell'')} \\
& + \left( (1-Z)\phi_d(i^d) + Z\phi_g(i^g) - \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} \right) V^{(\ell'')}_{\log K} + \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} V^{(\ell'')}_{\log K, \log K} \\
& + \left( \phi_g(i^g) - Z\sigma_g^2 - \phi_d(i^d) + (1-Z)\sigma_d^2 \right) Z(1-Z) V_Z^{(\ell'')} + \frac{1}{2} Z^2 (1-Z)^2 (\sigma_g^2 + \sigma_d^2) V_{ZZ}^{(\ell'')} \\
& + \left( \left( V^{(\ell'')}_{\log K} - Z V_Z^{(\ell'')} \right) (1-Z)\sigma_d + \left( V^{(\ell'')}_{\log K} + (1-Z) V_Z^{(\ell'')} \right) Z\sigma_g \right) \cdot h \\
& + \left[ -Z(1-Z)^2 \sigma_d^2 + Z^2 (1-Z)\sigma_g^2 \right] V^{(\ell'')}_{\log K, Z} \\
& + V_y^{(\ell'')} (\bar{\theta} + \varsigma \cdot h) \mathcal{E}_t + \frac{|\varsigma|^2 (\mathcal{E}_t)^2}{2} V_{yy}^{(\ell'')} + \xi \frac{|h|^2}{2} \\
& + \sum_{\ell=1}^{L-S} \mathcal{J}_n^{\ell}(y) g^{\ell} (V^{(\ell,\ell'')} - V^{(\ell'')}) + \xi \sum_{\ell=1}^{L-S} \mathcal{J}_n^{\ell}(y) \left( 1 - g^{\ell} + g^{\ell} \log g^{\ell} \right).
\end{aligned}
$$

## C.5 Pre Damage and Intermediate Technology Jumps Setting

The HJB equation for the pre damage and intermediate technology jumps setting is

$$
\begin{aligned}
0 = \max_{i^g, i^d, i^r} \min_{g^\ell, g^{\ell'}, h} \; & \delta \log([A_d - i^d](1-Z) + [A_g(\ell') - i^g]Z - i^r) + \delta \log K - \delta \log N(y) - \delta V^{(\ell')} \\
& + \left( (1-Z)\phi_d(i^d) + Z\phi_g(i^g) - \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} \right) V^{(\ell')}_{\log K} + \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} V^{(\ell')}_{\log K, \log K} \\
& + \left( \phi_g(i^g) - Z\sigma_g^2 - \phi_d(i^d) + (1-Z)\sigma_d^2 \right) Z(1-Z) V^{(\ell')}_Z + \frac{1}{2} Z^2 (1-Z)^2 (\sigma_g^2 + \sigma_d^2) V^{(\ell')}_{ZZ} \\
& + \left( \left( V^{(\ell')}_{\log K} - Z V^{(\ell')}_Z \right) (1-Z)\sigma_d + \left( V^{(\ell')}_{\log K} + (1-Z) V^{(\ell')}_Z \right) Z\sigma_g \right) \cdot h \\
& + \left[ -Z(1-Z)^2 \sigma_d^2 + Z^2 (1-Z)\sigma_g^2 \right] V^{(\ell')}_{\log K, Z} \\
& + V^{(\ell')}_y (\bar{\theta} + \varsigma \cdot h) \mathcal{E}_t + \frac{|\varsigma|^2 (\mathcal{E}_t)^2}{2} V^{(\ell')}_{yy} \\
& + \left( \psi_r(i^r) - \frac{1}{2}|\sigma_r|^2 + \sigma_r \cdot h \right) V^{(\ell')}_{\log R} + \frac{|\sigma_r|^2}{2} V^{(\ell')}_{\log R, \log R} + \xi \frac{|h|^2}{2} \\
& + \mathcal{J}_g^{\ell''}(R) g^{\ell''} (V^{(\ell'')} - V^{(\ell')}) + \sum_{\ell=1}^{L-S} \mathcal{J}_n^{\ell}(y) g^{\ell} (V^{(\ell, \ell')} - V^{(\ell')}) \\
& + \xi \left[ \mathcal{J}_g^{\ell''}(R) \left( 1 - g^{\ell''} + g^{\ell''} \log g^{\ell''} \right) + \sum_{\ell=1}^{L-S} \mathcal{J}_n^{\ell}(y) \left( 1 - g^{\ell} + g^{\ell} \log g^{\ell} \right) \right].
\end{aligned}
$$

## C.6 Pre Damage and Technology Jumps Setting

The HJB equation for the pre damage and technology jumps setting is given by

$$
\begin{aligned}
0 = \max_{i^g, i^d, i^r} \min_{g^\ell, g^{\ell'}, h} \; & \delta \log([A_d - i^d](1-Z) + [A_g - i^g]Z - i^r) + \delta \log K - \delta \log N(y) - \delta V \\
& + \left( (1-Z)\phi_d(i^d) + Z\phi_g(i^g) - \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} \right) V_{\log K} + \frac{\sigma_d^2(1-Z)^2 + \sigma_g^2 Z^2}{2} V_{\log K, \log K} \\
& + \left( \phi_g(i^g) - Z\sigma_g^2 - \phi_d(i^d) + (1-Z)\sigma_d^2 \right) Z(1-Z) V_Z + \frac{1}{2} Z^2 (1-Z)^2 (\sigma_g^2 + \sigma_d^2) V_{ZZ} \\
& + \left( (V_{\log K} - Z V_Z)(1-Z)\sigma_d + (V_{\log K} + (1-Z)V_Z) Z\sigma_g \right) \cdot h \\
& + \left[ -Z(1-Z)^2 \sigma_d^2 + Z^2 (1-Z)\sigma_g^2 \right] V_{\log K, Z} \\
& + V_y(\bar{\theta} + \varsigma \cdot h)\mathcal{E}_t + \frac{|\varsigma|^2 (\mathcal{E}_t)^2}{2} V_{yy} \\
& + \left( \psi_r(i^r) - \frac{1}{2}|\sigma_r|^2 + \sigma_r \cdot h \right) V_{\log R} + \frac{|\sigma_r|^2}{2} V_{\log R, \log R} + \xi \frac{|h|^2}{2} \\
& + \mathcal{J}_g^{\ell'}(R) g^{\ell'} (V^{(\ell')} - V) + \mathcal{J}_g^{\ell''}(R) g^{\ell''} (V^{(\ell'')} - V) + \sum_{\ell=1}^{L-S} \mathcal{J}_n^{\ell}(y) g^{\ell} (V^{(\ell)} - V) \\
& + \xi \left[ \mathcal{J}_g^{\ell'}(R) \left( 1 - g^{\ell'} + g^{\ell'} \log g^{\ell'} \right) + \mathcal{J}_g^{\ell''}(R) \left( 1 - g^{\ell''} + g^{\ell''} \log g^{\ell''} \right) + \sum_{\ell=1}^{L-S} \mathcal{J}_n^{\ell}(y) \left( 1 - g^{\ell} + g^{\ell} \log g^{\ell} \right) \right].
\end{aligned}
$$

# Appendix D Parameter Values

For our computations, we must specify ranges and initial values for our state variables. These values are given in Table 2, and are set similarly to Barnett et al. (2023). The initial value of global mean temperature anomaly $Y_0 = 1.1$ degrees Celsius matches the estimated current value from Masson-Delmotte et al. (2021b). The initial value of the green capital-to-total capital ratio $Z_0$ is chosen as an intermediate value from estimates in the literature using data on fixed assets and durable goods, the fraction of green-firm to total market capitalization, the electricity and transportation sectors from the EIA, BEA, FAA, and FTSE.[15] The initial value of total capital is set as $K_0 = 880$ so that our initial GDP matches the 2023 World GDP value of \$101.2 trillion estimated by the World Bank National Accounts data given our assumed initial average aggregate productivity of $\bar{\alpha} = (1 - Z_0) \times A_d + Z_0 \times A_g = 0.115$. The initial value of knowledge stock is set to $R_0 = 11.2$, which matches estimates from the

[15] Specifically, Fried et al. (2021), Dhar and Davidson (2022), Carattini et al. (2023), Berk and Van Binsbergen (2025) derive values ranging between 0.5 and 0.82 from data on fixed assets and durable goods, the fraction of green-firm to total market capitalization, the electricity and transportation sectors, and adjustments for uncertainty about the distinctions between clean and dirty capital in the data and ubiquity of fossil fuels in modern economic production.

BLS database of the current Total US R&D capital stock value, converted and scaled up to a global value based on the US-to-world productive capital ratio.

Table 2: State Variable Initial Values and Ranges

| State variables | values |
|---|---|
| $K_0$ | 880 |
| $Z_0$ | 0.7 |
| $Y_0$ | 1.1 |
| $R_0$ | 11.2 |
| State variables | range |
| $\log(K)$ | $[4, 7]$ |
| $Z$ | $[0.01, 0.99]$ |
| $Y$ | $[0, 4]$ |
| $\log(R)$ | $[1, 6]$ |

The economic parameter values are given in Table 3. The choice of $\delta$ is consistent with values of the subjective discount rate used in the macroeconomics and asset pricing literature. The values of $(\Gamma_d, \theta_d)$ and $(\Gamma_g, \theta_g)$ are chosen so that the no-climate, one capital version of the model satisfies three conditions, consistent with estimated values in the macroeconomics literature and empirical values from the BEA and World Bank databases, as in Barnett et al. (2020): a Tobin's q value of 2.5 ($q = \frac{A-i}{\delta} v_{\log K} = \frac{1+\theta i}{\Gamma\theta} = 2.5$); a growth rate of capital of 2% ($\mathbb{E}[dK/K] = \alpha + \Gamma \log(1 + \theta i) = 0.02$); and a ratio of the adjustment cost parameters normalized to one ($\Gamma\theta = 1$). The values of $\sigma_d$ and $\sigma_g$ are chosen to match the annual percent changes in the time series of GDP from the World Bank database. Given the initial values of global GDP and green-to-total capital ratio from the data, our assumed initial average aggregate productivity of $\bar{\alpha} = 0.115$ means the values of $A_d$ and $A_g$ must be jointly determined in order to generate output consistent with World Bank World GDP values.

To determine the values of $A_d$ and $A_g$, as well as the values $A_g(\ell')$ and $A_g(\ell'')\}$ for the intermediate and breakthrough technological innovations, we build off of the empirical estimation of Acemoglu et al. (2016). In their framework, an incremental technology step is defined by a factor $\lambda$ such that the resulting incremental productivity gain is given by $A_j \times \lambda$. In addition, they define $n$ as the number of steps between the leading and lagging technology in the economy $A_i/A_j = \lambda^n$. Given our assumption that $A_d > A_g$, consistent with the findings in Acemoglu et al. (2016), and because of our assumed AK production

technology and additivity across output types, the following two conditions must hold

$$\begin{aligned} A_d &= \lambda^n \times A_g \\ \bar{\alpha} &= Z_0 \times A_g + (1 - Z_0) \times A_d \end{aligned}$$

Using Acemoglu et al. (2016)'s estimates of the incremental step size of $\lambda = 1.063$ and of the leading and lagging technology gap of $n = 3$, along with the initial values of global GDP and green-to-total capital ratio from the data, our initial productivity parameters are

$$A_d = 0.1303 \text{ and } A_g = 0.1085$$

We define our intermediate technological innovation as a "catching up" shock such that green productivity matches dirty productivity

$$A_g(\ell') = A_d = 0.1303$$

and our "breakthrough" shock such that it creates a symmetrically sized technology gap between green productivity and dirty productivity

$$A_g(\ell') = \lambda^n \times A_d = 0.1567.$$

We again use the empirical estimate of Acemoglu et al. (2016) for a breakthrough shock to set the probability of a breakthrough shock, conditional on the occurrence of the first technology jump, as $\pi = 0.04$. This leaves the probability of the intermediate shock, conditional on the occurrence of the first technology jump occurring, as $1 - \pi = 0.96$.

Table 3: Economic Parameters

| Parameters | values |
|---|---|
| $\delta$ | 0.01 |
| $(\alpha_d, \Gamma_d, \theta_d, \sigma_d)$ | ( -0.035, 0.060, 16.7, 0.01) |
| $(\alpha_g, \Gamma_g, \theta_g, \sigma_g)$ | ( -0.035, 0.060, 16.7, 0.01) |
| $A_d$ | 0.1303 |
| $\{A_g, A_g(\ell'), A_g(\ell'')\}$ | $\{0.1085, 0.1303, 0.1567\}$ |
| $(\lambda, n)$ | (1.063, 3) |
| $(\zeta, \psi_0, \psi_1, \sigma_\kappa)$ | (0, 0.10583, 0.5, 0.0078) |
| $\varrho$ | 746.67 |

We assume the depreciation of the knowledge stock to be $\zeta = 0$ for simplicity and set the R&D investment curvature parameter as $\psi_1 = 0.5$ for computational tractability. The initial value of our knowledge stock is based on estimates from the BLS of the total US R&D stock values (scaled up to World values). We set $\psi_0 = 0.1$, guided by estimates for the returns to R&D investment by Lucking et al. (2019) and Bloom et al. (2019). Barnett et al. (2023) show in a simplified version of their model, with no damage jumps or uncertainty version, the combination of our initial knowledge stock value with these R&D investment parameters and choice of scaling parameter $\varrho = 746.67$ generates R&D investment that peaks just below .5% of GDP, consistent with estimates by Stine (2008) for major U.S. R&D investment programs, and an anticipated arrival time of the green technological innovation of around 30-40 years. The value of $\sigma_r$ is chosen to match the annual percent changes in the time series of U.S. R&D capital stock from the BLS database.

Table 4: Climate Dynamics and Damages Parameters

| Parameters | values |
|---|---|
| $\bar{\theta}$ | 1.86 / 1000 |
| $\eta$ | 0.291 |
| $\varsigma$ | $1.2 \times 1.86/1000$ |
| $(\lambda_1, \lambda_2)$ | $(0.00017675, 2 \times 0.0022)$ |
| $\{\lambda_3(\ell)\}_{\ell=1,\ldots,L-S}$ | $\{\frac{1}{3}\frac{\ell-1}{L-1}\}_{\ell=1,\ldots,L}$ |
| $(r_1, r_2)$ | $(1.5, 0.36)$ |
| $(\underline{y}, \bar{y})$ | $(1.5, 2.5)$ |

The climate dynamics and climate damage parameter values are given in Table 4. The values of $\theta(m)$ come from pulse experiments estimates produced by Barnett et al. (2022), based on the results from Joos et al. (2013) and Geoffroy et al. (2013), and are consistent with values reported in Masson-Delmotte et al. (2021b). The value $\bar{\theta}$ is the average value across all climate 144 models. The value of $\eta$ is chosen to match recent estimates of annual carbon emissions of 10 GtC from Figueres et al. (2018) and the Global Carbon Project, as well as the estimates of World GDP and the clean and dirty capital split. Specifically, for our initial values of green-to-total capital and total capital stock of $Z_0 = 0.7$ and $K_0 = 880$, and assumed dirty sector productivity of $A_d = 0.13$, the empirical target for initial emissions of $\mathcal{E}_t = 10$ GtC produces an emissions intensity value of $\eta = 0.291$. The value of $\varsigma$ matches volatility used by Barnett et al. (2022). The values of $\lambda_1$, $\lambda_2$, $\lambda_3(\ell)$, and $\bar{y}$ match damage function parameters used by Barnett et al. (2023), which are designed to incorporate the spread of potential climate damage outcomes based on Nordhaus (2019), Weitzman (2012), and the very recent work by Waidelich et al. (2024).

Calibration of the climate damage jump intensity function is motivated by the climate

thresholds and climate tipping point literature. Drijfhout et al. (2015), Armstrong McKay et al. (2022), Rogelj et al. (2018) and Rogelj et al. (2019) note thresholds beginning as earlier as $1^{\circ}C$, multiple thresholds above $\sim 1.5^{\circ}C$, and increasing probability and risks of more severe damages at and beyond $2.0^{\circ}C$. Ritchie et al. (2021) and Hansen et al. (2025) suggest points of no return, beyond which tipping points are not reversible, are highly uncertain based on dynamical systems theory modeling and are dependent upon long-term feedback responses. Thus, we first choose baseline values of $\underline{y} = 1.5$ and $\overline{y} = 2.5$ for the range over which the damage jump can occur. We then set $\mathsf{r}_0 = 1.5$ and target a jump realization occurring by $y_m = 2.0$, which we achieve by imposing $\mathsf{r}_1 = 0.36$ based on a simplified temperature pathway.

# Appendix E Neural Nets Implementations

Below we give the pseudo-code for the deep Galerkin method - policy improvement algorithms (DGM-PIA) described in Section 4.5, and implementation details, including network architectures and training hyperparameters. For the DGM-PIA described in Section 4.5, we use feedforward neural networks with 4 hidden layers of width 32. For the hidden layers, we use `tanh` and `softplus` activation functions to approximate the optimal controls and `swish` activation functions to approximate the value function. For the the output layer, we use a customized hyperbolic tangent function and `softplus` activation functions. To train the neural nets, we run 2000000 epochs with a batch size of 128. The learning rates $\text{lr}_V = 10e^{-5}$ and $\text{lr}_{\boldsymbol{\alpha}} = 10e^{-5}$ and we use the ADAM optimizer proposed in Kingma and Ba (2014). The training can be done on Google Colab with a total runtime of 2.29 hours.

**Algorithm 1** The DGM-PIA algorithm for solving the generic HJB (4.5)

**Require:** a maximum number of epochs $N$, initial neural networks for the value function $V(\boldsymbol{x};\theta_0)$ and the control $\boldsymbol{\alpha}(\boldsymbol{x};\varphi_0)$, learning rates for the value function and the control $\mathrm{lr}_V$, $\mathrm{lr}_{\boldsymbol{\alpha}}$

1: **for** $n = 0, 1, 2, \ldots, N$ **do**

2: Generate random samples $\{\boldsymbol{x}_m\}_{m=1}^M$ from the domain $\Omega$ according to $\nu$

3: Compute the value function loss in (4.5) using samples $\{\boldsymbol{x}_m\}_{m=1}^M$:

$$L_V\left(\theta_n; \{\boldsymbol{x}_m\}_{m=1}^M\right) = \frac{1}{M}\sum_{m=1}^{M}\left[-\delta V(\boldsymbol{x}_m;\theta_n) + \mathcal{L}^{\boldsymbol{\alpha}(\boldsymbol{x}_m;\varphi_n)}V(\boldsymbol{x}_m;\theta_n) + f(\boldsymbol{x}_m, \boldsymbol{\alpha}(\boldsymbol{x}_m;\varphi_n))\right]^2$$

4: Take a gradient descent step to update $\theta_{n+1}$:

$$\theta_{n+1} = \theta_n - \mathrm{lr}_V \nabla_\theta L_V(\theta_n; \{\boldsymbol{x}_m\}_{m=1}^M)$$

5: Compute the control loss in (4.5) using samples $\{\boldsymbol{x}_m\}_{m=1}^M$:

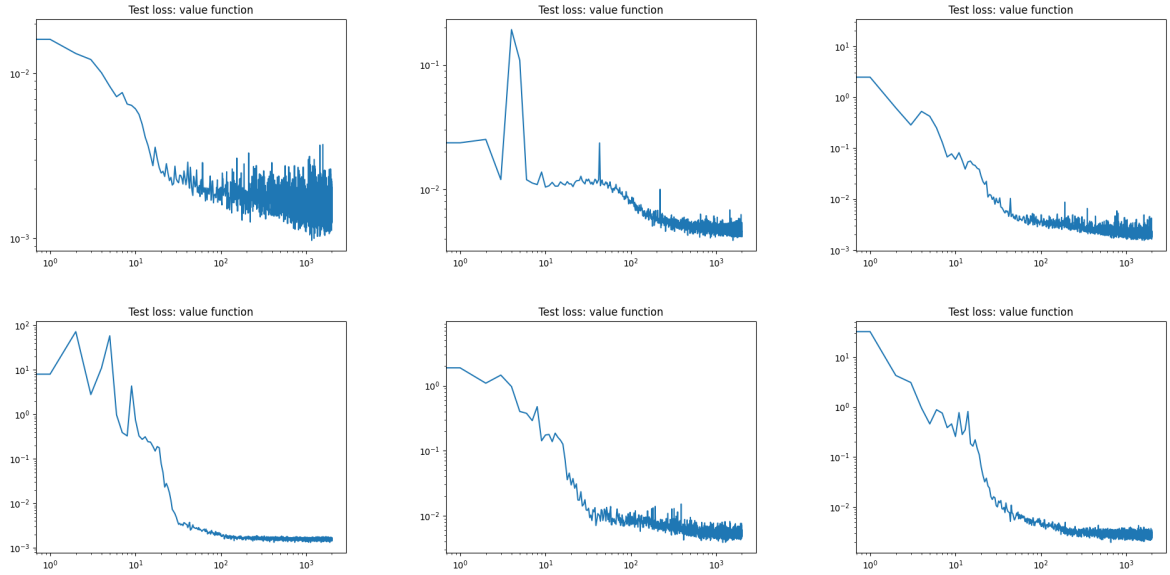
$$L_{\boldsymbol{\alpha}}\left(\varphi_n; \{\boldsymbol{x}_m\}_{m=1}^M\right) = -\frac{1}{M}\sum_{m=1}^{M}\left[\mathcal{L}^{\boldsymbol{\alpha}(\boldsymbol{x}_m;\varphi_n)}V(\boldsymbol{x}_m;\theta_n) + f(\boldsymbol{x}_m, \boldsymbol{\alpha}(\boldsymbol{x}_m;\varphi_n))\right]^2$$

6: Take a gradient descent step to update $\varphi_{n+1}$:

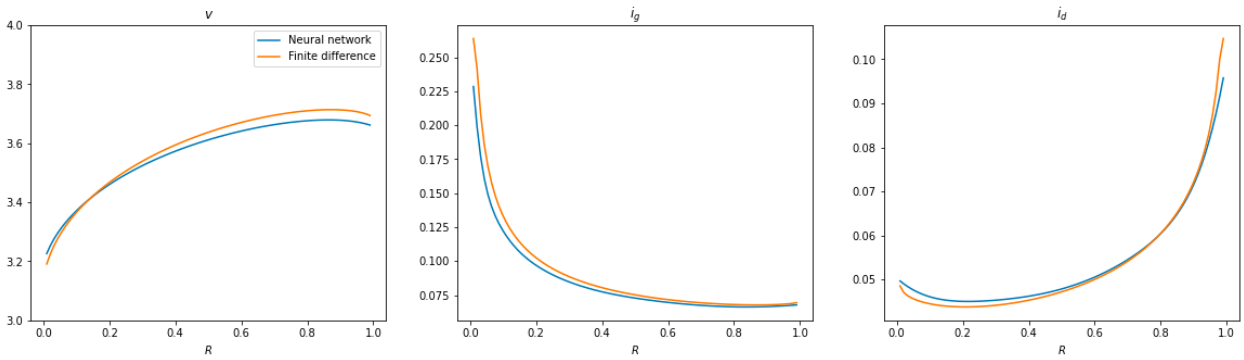
$$\varphi_{n+1} = \varphi_n - \mathrm{lr}_{\boldsymbol{\alpha}} \nabla_\varphi L_{\boldsymbol{\alpha}}(\varphi_n; \{\boldsymbol{x}_m\}_{m=1}^M)$$

7: **end for**

# Appendix F Algorithm Validation Results




(a) HJB training loss error for each jump state solution.




(b) Finite difference and DGM-PIA solution comparison

Figure 8: Model solution validation. Panel (a) plot the training errors over epochs for each of the HJB equations corresponding to the different jump states in our coupled system: post-tech, post-damage (top left); interim-tech, post-damage (top middle); pre-tech, post-damage (top right); post-tech, pre-damage (bottom left); interim-tech, pre-damage (bottom middle); pre-tech, pre-damage (bottom right). Panel (b) plots the value function and optimal controls for the FD and NN solution in the post-technology and post-damage jump state.

## F.1 Economically Relevant Training Loss Ranges: $\xi = \infty$

| Scenario | $T = 20$ | $T = 40$ | $T = 60$ |
|---|---|---|---|
| Pre-Damage-Pre-Tech | $2.69 \times 10^{-3}$ | $2.40 \times 10^{-3}$ | $1.86 \times 10^{-3}$ |
| Post-Damage-Pre-Tech | NA | $0.76 \times 10^{-3}$ | $1.16 \times 10^{-3}$ |
| Pre-Damage-Interm-Tech | $5.96 \times 10^{-3}$ | $5.18 \times 10^{-3}$ | $4.19 \times 10^{-3}$ |
| Post-Damage-Interm-Tech | $4.40 \times 10^{-3}$ | $4.11 \times 10^{-3}$ | $4.58 \times 10^{-3}$ |
| Pre-Damage-Post-Tech | $3.13 \times 10^{-3}$ | $3.32 \times 10^{-3}$ | $3.84 \times 10^{-3}$ |
| Post-Damage-Post-Tech | NA | $1.97 \times 10^{-3}$ | $2.24 \times 10^{-3}$ |

Table 5: Validation loss for the PDEs across the four jump realization states, transposed. The values are computed by averaging the PDE residuals evaluated at locations likely to be visited during simulation. NA represents states where no more than one path entered the particular jump state space at the specified year.

| Regime (Tech, Damage) | State Variable | $t$=20 | $t$=40 | $t$=60 |
|---|---|---|---|---|
| | $\log K_t$ | [6.80, 6.91] | [6.84, 6.99] | [6.87, 7.06] |
| Pre Damage | $Z_t$ | [0.69, 0.73] | [0.70, 0.75] | [0.73, 0.78] |
| Pre Tech | $Y_t$ | [1.34, 1.62] | [1.61, 2.05] | [1.89, 2.37] |
| | $\log R_t$ | [3.13, 3.23] | [3.65, 3.76] | [4.01, 4.12] |
| | $\log K_t$ | NA | [6.86, 7.01] | [6.88, 7.07] |
| Post Damage | $Z_t$ | NA | [0.70, 0.75] | [0.71, 0.78] |
| Pre Tech | $Y_t$ | NA | [2.50, 2.75] | [2.53, 3.16] |
| | $\log R_t$ | NA | [3.65, 3.76] | [4.01, 4.15] |
| | $\log K_t$ | [6.83, 6.97] | [6.90, 7.13] | [6.97, 7.29] |
| Pre Damage | $Z_t$ | [0.70, 0.75] | [0.72, 0.79] | [0.74, 0.82] |
| Interm Tech | $Y_t$ | [1.34, 1.63] | [1.66, 2.02] | [1.91, 2.35] |
| | $\log R_t$ | [3.14, 3.24] | [3.64, 3.77] | [4.02, 4.18] |
| | $\log K_t$ | [6.83, 6.85] | [6.91, 7.12] | [6.99, 7.28] |
| Post Damage | $Z_t$ | [0.71, 0.71] | [0.72, 0.78] | [0.74, 0.82] |
| Interm Tech | $Y_t$ | [2.49, 2.50] | [2.50, 2.85] | [2.53, 3.13] |
| | $\log R_t$ | [3.12, 3.13] | [3.65, 3.79] | [4.03, 4.18] |
| | $\log K_t$ | [6.88, 7.07] | [6.94, 7.38] | [7.15, 7.74] |
| Pre Damage | $Z_t$ | [0.72, 0.79] | [0.74, 0.86] | [0.78, 0.91] |
| Post Tech | $Y_t$ | [1.34, 1.56] | [1.62, 2.05] | [1.95, 2.30] |
| | $\log R_t$ | NA | NA | NA |
| | $\log K_t$ | NA | [6.95, 7.33] | [7.09, 7.66] |
| Post Damage | $Z_t$ | NA | [0.75, 0.85] | [0.78, 0.89] |
| Post Tech | $Y_t$ | NA | [2.49, 2.76] | [2.53, 3.03] |
| | $\log R_t$ | NA | NA | NA |

Table 6: The 5th and 95th percentiles of state variable values computed at the specified year based on 10,000 stochastic simulation pathways. NA represents states where no more than one path entered the particular jump state space at the specified year, or the state variable is no longer relevant in the computations.

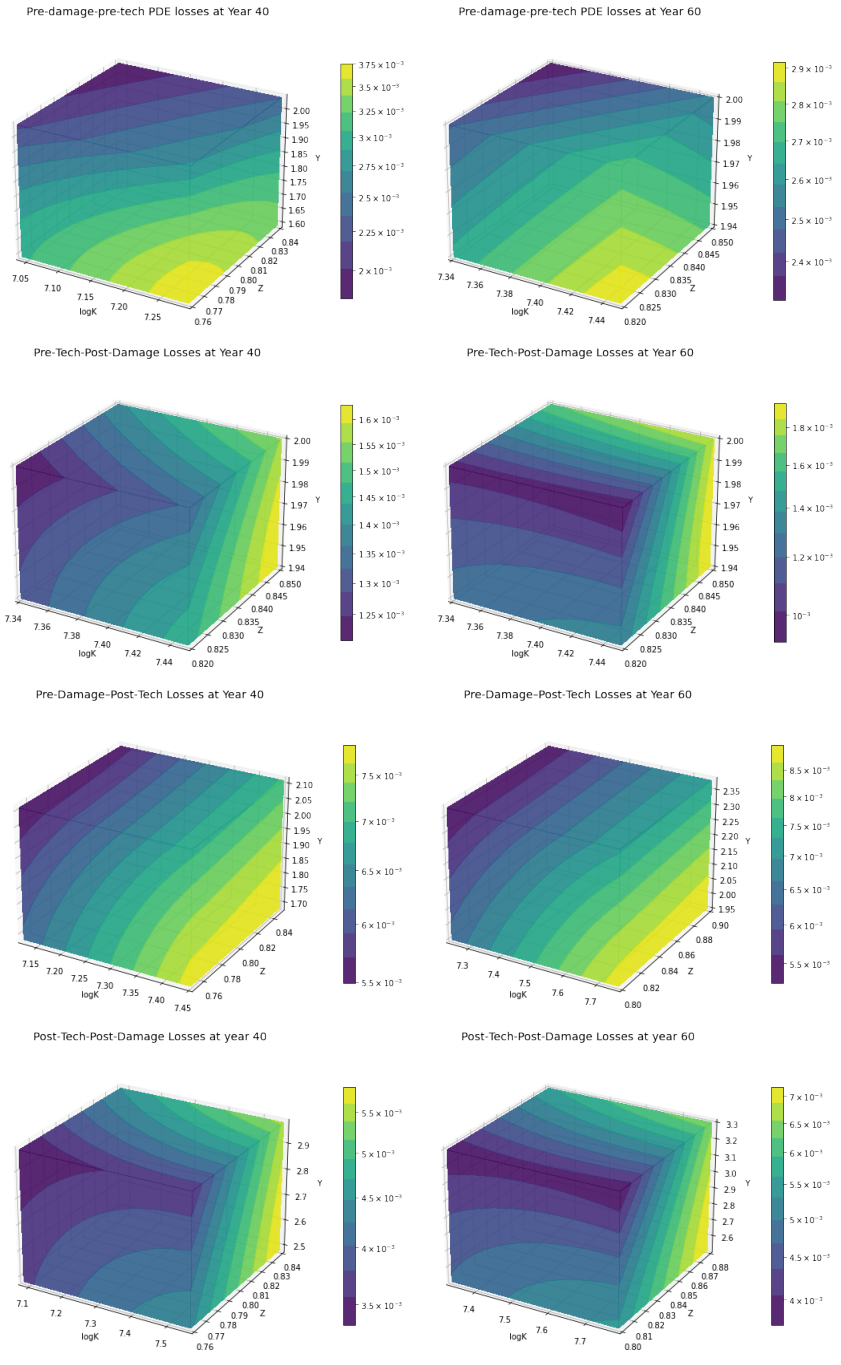



Figure 9: Three dimensional plots of the training loss errors shown over the 5th and 95th percentiles of state variable values computed at the specified year based on 10,000 stochastic simulation pathways. The first row shows the no jump state. The second row shows the damage only jump state. The third row shows the technology only jump state. The last row shows the both jump state. The left column shows the errors at year 40. The right right column shows the errors at year 60.

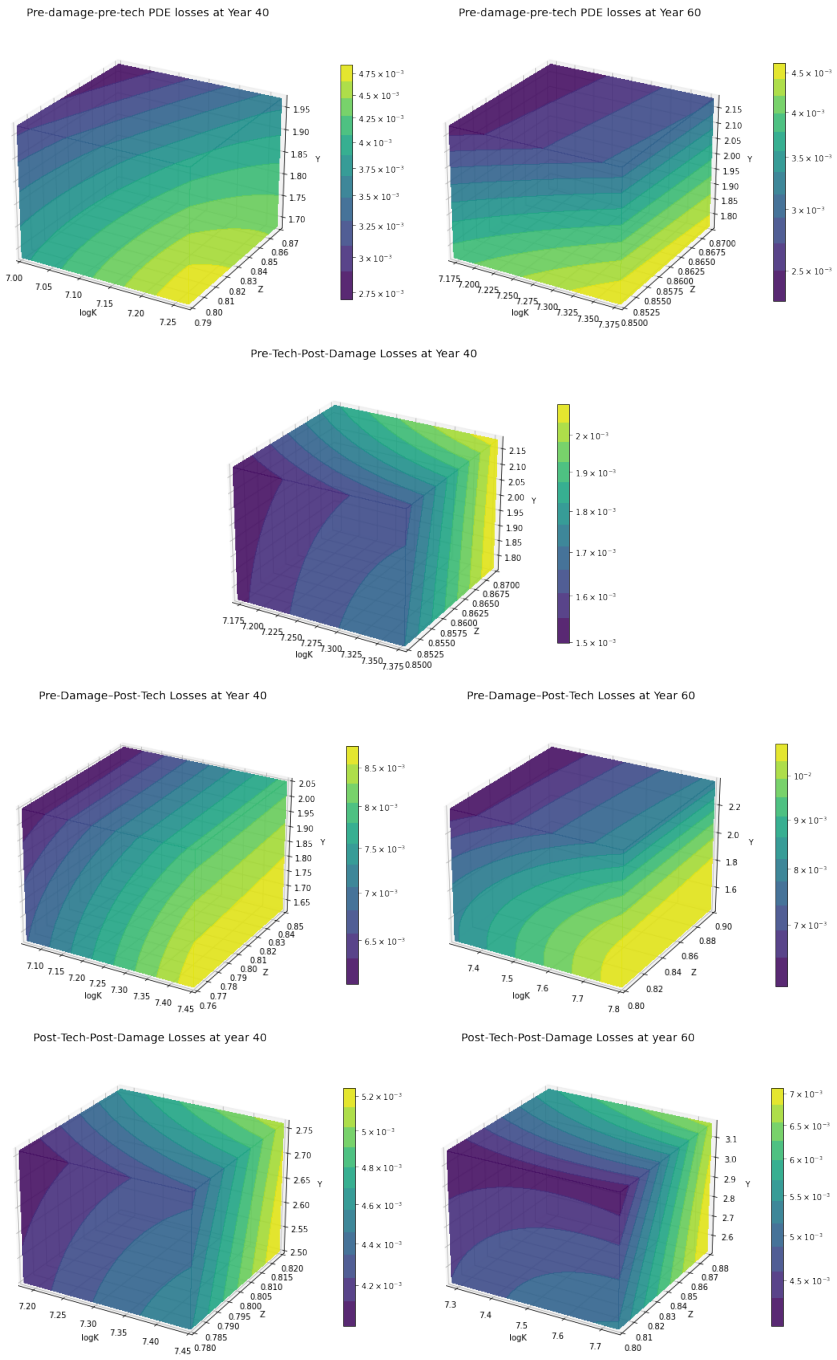
## F.2 Economically Relevant Training Loss Ranges: $\xi = 0.1$

| Scenario | $T = 20$ | $T = 40$ | $T = 60$ |
|---|---|---|---|
| Pre-Damage-Pre-Tech | $3.10 \times 10^{-3}$ | $2.22 \times 10^{-3}$ | $1.24 \times 10^{-3}$ |
| Post-Damage-Pre-Tech | NA | $0.46 \times 10^{-3}$ | $0.40 \times 10^{-4}$ |
| Pre-Damage-Interm-Tech | $7.16 \times 10^{-3}$ | $5.86 \times 10^{-3}$ | $4.28 \times 10^{-3}$ |
| Post-Damage-Interm-Tech | $4.40 \times 10^{-3}$ | $4.11 \times 10^{-3}$ | $4.58 \times 10^{-3}$ |
| Pre-Damage-Post-Tech | $3.35 \times 10^{-3}$ | $3.57 \times 10^{-3}$ | $4.22 \times 10^{-3}$ |
| Post-Damage-Post-Tech | NA | $1.91 \times 10^{-3}$ | $2.16 \times 10^{-3}$ |

Table 7: Validation loss for the PDEs across the four jump realization states, transposed. The values are computed by averaging the PDE residuals evaluated at locations likely to be visited during simulation. NA represents states where no more than one path entered the particular jump state space at the specified year.

| Regime (Tech, Damage) | State Variable | $t$=20 | $t$=40 | $t$=60 |
|---|---|---|---|---|
| | $\log K_t$ | [6.80, 6.90] | [6.84, 6.99] | [6.84, 7.07] |
| Pre Damage | $Z_t$ | [0.69, 0.73] | [0.70, 0.75] | [0.71, 0.77] |
| Pre Tech | $Y_t$ | [1.34, 1.62] | [1.66, 2.04] | [1.97, 2.30] |
| | $\log R_t$ | [3.17, 3.26] | [3.69, 3.82] | [4.09, 4.21] |
| | $\log K_t$ | [6.86, 6.89] | [6.85, 6.99] | [6.87, 7.06] |
| Post Damage | $Z_t$ | NA | [0.70, 0.75] | [0.71, 0.78] |
| Pre Tech | $Y_t$ | NA | [2.52, 2.83] | [2.55, 3.06] |
| | $\log R_t$ | NA | [3.69, 3.81] | [4.05, 4.18] |
| | $\log K_t$ | NA | [6.89, 7.12] | [6.95, 7.28] |
| Pre Damage | $Z_t$ | [0.70, 0.74] | [0.71, 0.77] | [0.73, 0.81] |
| Interm Tech | $Y_t$ | [1.33, 1.61] | [1.63, 2.02] | [1.89, 2.37] |
| | $\log R_t$ | [3.18, 3.28] | [3.72, 3.85] | [4.11, 4.27] |
| | $\log K_t$ | [6.83, 6.86] | [6.90, 7.12] | [6.98, 7.27] |
| Post Damage | $Z_t$ | [0.70, 0.73] | [0.71, 0.77] | [0.73, 0.80] |
| Interm Tech | $Y_t$ | [2.51, 2.53] | [2.50, 2.84] | [2.55, 3.15] |
| | $\log R_t$ | [3.19, 3.23] | [3.70, 3.84] | [4.08, 4.25] |
| | $\log K_t$ | [6.88, 7.07] | [6.95, 7.40] | [7.11, 7.59] |
| Pre Damage | $Z_t$ | [0.72, 0.79] | [0.73, 0.86] | [0.77, 0.89] |
| Post Tech | $Y_t$ | [1.31, 1.53] | [1.65, 2.06] | [1.93, 2.30] |
| | $\log R_t$ | NA | NA | NA |
| | $\log K_t$ | NA | [7.04, 7.32] | [7.05, 7.66] |
| Post Damage | $Z_t$ | NA | [0.76, 0.85] | [0.76, 0.90] |
| Post Tech | $Y_t$ | NA | [2.50, 2.72] | [2.51, 3.01] |
| | $\log R_t$ | NA | NA | NA |

Table 8: The 5th and 95th percentiles of state variable values computed at the specified year based on 10,000 stochastic simulation pathways. NA represents states where no more than one path entered the particular jump state space at the specified year, or the state variable is no longer relevant in the computations.




(a) HJB training loss error for each jump state solution.

Figure 10: Three dimensional plots of the training loss errors shown over the 5th and 95th percentiles of state variable values computed at the specified year based on 10,000 stochastic simulation pathways. The first row shows the no jump state. The second row shows the damage only jump state. The third row shows the technology only jump state. The last row shows the both jump state. The left column shows the errors at year 40. The right right column shows the errors at year 60. Missing plot times are when no more than one path entered the particular jump state space at the specified year.

## F.3 Economically Relevant Training Loss Ranges: $\xi = 0.05$

| Scenario | $T = 20$ | $T = 40$ | $T = 60$ |
|---|---|---|---|
| Pre-Damage-Pre-Tech | $3.39 \times 10^{-3}$ | $2.29 \times 10^{-3}$ | $1.25 \times 10^{-3}$ |
| Post-Damage-Pre-Tech | NA | $0.44 \times 10^{-3}$ | $0.36 \times 10^{-3}$ |
| Pre-Damage-Interm-Tech | $7.61 \times 10^{-3}$ | $5.99 \times 10^{-3}$ | $4.00 \times 10^{-3}$ |
| Post-Damage-Interm-Tech | $5.41 \times 10^{-3}$ | $4.55 \times 10^{-3}$ | $4.53 \times 10^{-3}$ |
| Pre-Damage-Post-Tech | $3.29 \times 10^{-3}$ | $3.46 \times 10^{-3}$ | $4.05 \times 10^{-3}$ |
| Post-Damage-Post-Tech | NA | $1.62 \times 10^{-3}$ | $1.87 \times 10^{-3}$ |

Table 9: Validation loss for the PDEs across the four jump realization states, transposed. The values are computed by averaging the PDE residuals evaluated at locations likely to be visited during simulation. NA represents states where no more than one path entered the particular jump state space at the specified year.

| Regime (Tech, Damage) | State Variable | $t$=20 | $t$=40 | $t$=60 |
|---|---|---|---|---|
| | $\log K_t$ | [6.79, 6.90] | [6.84, 6.99] | [6.92, 7.01] |
| Pre Damage | $Z_t$ | [0.69, 0.73] | [0.71, 0.75] | [0.73, 0.77] |
| Pre Tech | $Y_t$ | [1.34, 1.62] | [1.66, 2.01] | [1.92, 2.27] |
| | $\log R_t$ | [3.19, 3.28] | [3.71, 3.84] | [4.13, 4.23] |
| | $\log K_t$ | [6.89, 6.89] | [6.85, 6.99] | [6.89, 7.08] |
| Post Damage | $Z_t$ | [0.73, 0.73] | [0.70, 0.75] | [0.72, 0.78] |
| Pre Tech | $Y_t$ | [2.50, 2.50] | [2.50, 2.80] | [2.56, 3.03] |
| | $\log R_t$ | [3.21, 3.21] | [3.71, 3.83] | [4.08, 4.21] |
| | $\log K_t$ | [6.83, 6.96] | [6.90, 7.12] | [6.96, 7.28] |
| Pre Damage | $Z_t$ | [0.70, 0.75] | [0.72, 0.78] | [0.74, 0.81] |
| Interm Tech | $Y_t$ | [1.34, 1.63] | [1.65, 2.04] | [1.92, 2.42] |
| | $\log R_t$ | [3.19, 3.28] | [3.72, 3.85] | [4.12, 4.28] |
| | $\log K_t$ | [6.91, 6.91] | [6.89, 7.11] | [6.97, 7.28] |
| Post Damage | $Z_t$ | [0.71, 0.71] | [0.71, 0.78] | [0.73, 0.81] |
| Interm Tech | $Y_t$ | [2.49, 2.49] | [2.50, 2.83] | [2.52, 3.14] |
| | $\log R_t$ | [3.24, 3.24] | [3.71, 3.84] | [4.09, 4.27] |
| | $\log K_t$ | [6.85, 7.10] | [6.96, 7.39] | [7.12, 7.64] |
| Pre Damage | $Z_t$ | [0.72, 0.79] | [0.75, 0.86] | [0.79, 0.89] |
| Post Tech | $Y_t$ | [1.37, 1.57] | [1.62, 1.92] | [1.91, 2.28] |
| | $\log R_t$ | NA | NA | NA |
| | $\log K_t$ | NA | [6.99, 7.34] | [7.04, 7.59] |
| Post Damage | $Z_t$ | NA | [0.76, 0.85] | [0.76, 0.90] |
| Post Tech | $Y_t$ | NA | [2.52, 2.72] | [2.51, 3.09] |
| | $\log R_t$ | NA | NA | NA |

Table 10: The 5th and 95th percentiles of state variable values computed at the specified year based on 10,000 stochastic simulation pathways. NA represents states where no more than one path entered the particular jump state space at the specified year, or the state variable is no longer relevant in the computations.

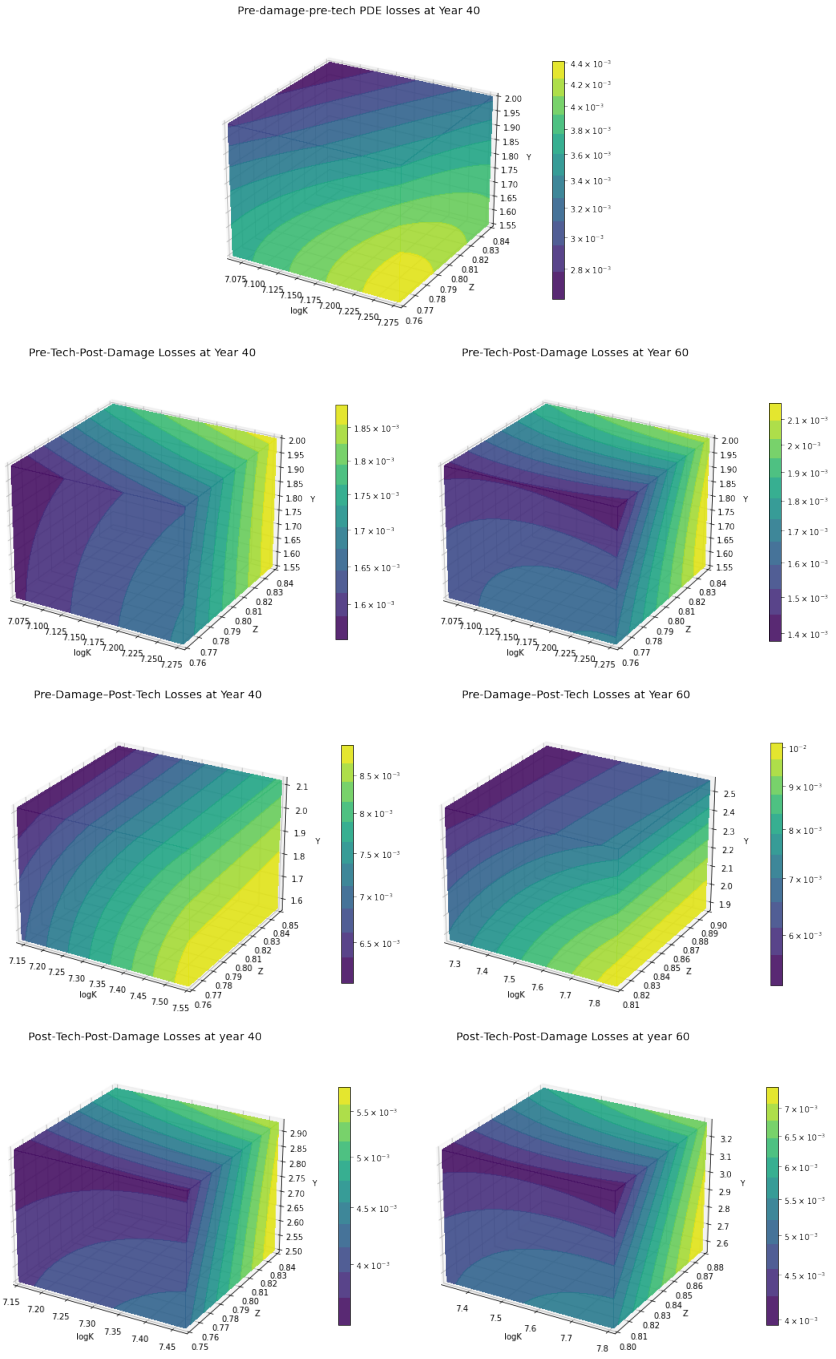



Figure 11: Three dimensional plots of the training loss errors shown over the 5th and 95th percentiles of state variable values computed at the specified year based on 10,000 stochastic simulation pathways. The first row shows the no jump state. The second row shows the damage only jump state. The third row shows the technology only jump state. The last row shows the both jump state. The left column shows the errors at year 40. The right right column shows the errors at year 60. Missing plot times are when no more than one path entered the particular jump state space at the specified year.