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

PRELIMINARY DRAFT

Abstract

We study the implications of model uncertainty in a climate-economics framework with three types of capital: "dirty" capital that produces carbon emissions when used for production, "clean" capital that generates no emissions but is initially less productive than dirty capital, and knowledge capital that increases with R&D investment and leads to technological innovation in green sector productivity. To solve our high-dimensional, non-linear model framework we implement a neural-network-based global solution method. We show there are first-order impacts of model uncertainty on optimal decisions and social valuations in our integrated climate-economic-innovation framework. Accounting for interconnected uncertainty over climate dynamics, economic damages from climate change, and the arrival of a green technological change leads to substantial adjustments to investment in the different capital types in anticipation of technological change and the revelation of climate damage severity.

*We are grateful for valuable suggestions and comments from Victor Duarte (discussant), Harrison Hong, Felix Kubler, Simon Scheidegger, Jose Scheinkman, and Rebecca Willett. We also thank participants at 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). This draft is preliminary and comments are welcome.

1 Introduction

The potential consequences of climate change are becoming increasingly apparent. The Sixth Assessment Report (AR6) of the Intergovernmental Panel on Climate Change (IPCC) states 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.

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 means 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.

An important consideration in our analysis is the substantial uncertainty about the central mechanisms in our model. Specifically, our framework allows for uncertainty as it pertains to carbon-climate dynamics, or the mapping from carbon emissions into atmospheric carbon into temperature changes; economic damage functions, or the negative impact on output due to changes in atmospheric temperature; and technological innovation, or the likelihood of a technology shock that augments the productivity of green capital via R&D investment. While each

[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).

of these model components allows for risk in the form of stochastic realizations, we are interested in exploring alternative forms of uncertainty. This includes ambiguity about possible model parameterizations, as well as 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. Our analysis highlights the endogenous interdependencies and feedback effects that substantially impact optimal policy choices and social valuations related to green innovation coming from uncertainty aversion.

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 for our multi-dimensional, continuous-time, climate-economic framework. Our algorithm augments the deep Galerkin method-policy improvement algorithm setting (DGM-PIA) by incorporating key parameters as pseudo-states. While using parameters as pseudo-states 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. The use of pseudo states is critical for numerically approximating the unique monotonicity and nonlinearity features in our framework arising from model uncertainty aversion. As a result, the algorithm is able to handle multiple non-stationary, endogenous state variables with considerable non-linearity in an infinite horizon setting. Our numerical methodology is an important contribution of our work to the fields of macroeconomics, finance, and climate-economics. The algorithm provides a toolbox that considerably expands the ability of researchers to address research questions in these areas that can quickly be overwhelmed by the "curse of dimensionality" when accounting for regional, household, and firm heterogeneities across production technologies, economic frictions, and policy objectives.

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

[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.

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. 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. 2021, 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, mathematically, 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).

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 Climate-Economics Model

We now outline the main model for our analysis. The model incorporates components related to climate dynamics, economic elements and dynamics related to damages from climate change, production technologies, preferences, and technological innovation through R&D, as well as model uncertainty aversion. We outline each of these model pieces in what follows, and then derive a number of theoretical results before presenting the numerical solutions to the model.

2.1 Climate Dynamics

Recent work in geosciences has focused on constructing simplified approximations of climate dynamics generated by large-scale atmosphere-ocean general circulation models (AOGCM) for use in policy analysis. These include a tractable framework demonstrated by Matthews et al.

(2009) and Friedlingstein (2015) as a proportional relationship between temperature change and cumulative carbon emissions. This relationship is of the general form:

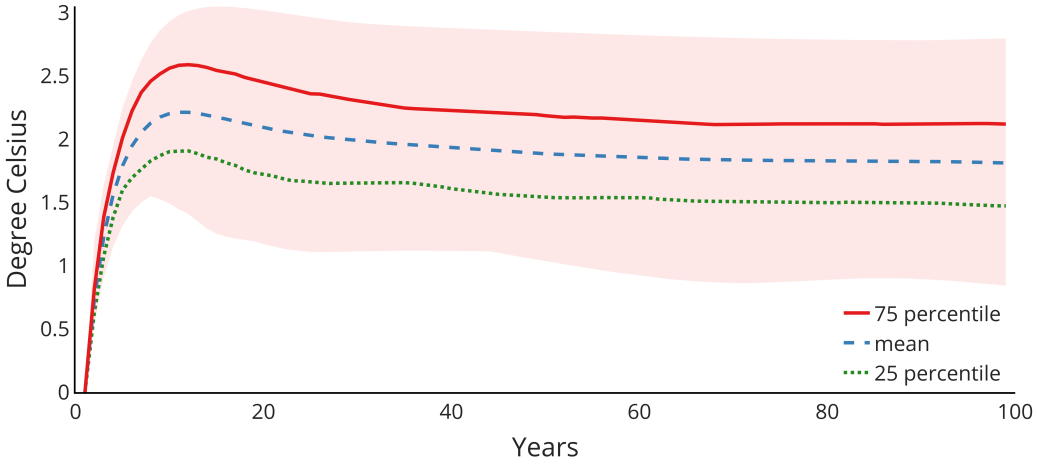
$$\text{temperature anomaly} \quad \approx \underline{\text{climate sensitivity}} \ (\beta_f) \ \times \ \text{cumulative emissions.}$$

Not only is this approximate relationship straightforward to incorporate into most continuous-time macroeconomic and asset pricing models, it also directly provides a model comparison set to use in our uncertainty analysis based on the constant of proportionality implied by different large-scale climate models. For our analysis, we use a stochastic variant of this relationship

$$dY_t = E_t(\beta_f dt + \varsigma dW_t),$$

where Y_t is the global mean temperature anomaly, with respect to the preindustrial level, E_t is global carbon emissions, W_t is a Brownian motion process with filtration $\mathcal{F}_t$, ς is the volatility loading for temperature, and β_f is the climate sensitivity proportionality constant. This model abstracts from transitory "weather" fluctuations in temperature and assumes that emissions today have a long-lasting, i.e., permanent, impact on future temperature. The inclusion of stochastic variation, which allows for a meaningful uncertainty analysis by "disguising" the climate model in the statistical sense that a small set of observations do not immediately reveal the true model, is motivated by the argument about climate model predictability put forward in Palmer and Stevens (2019). The parameter β_f varies across different carbon-climate dynamics models, and we will denote these different values as $\beta_{f,\ell}$ for a given model $\ell \in \{1, ..., \mathcal{L}\}$. In the geoscience literature, this parameter is known as the transient climate response to cumulative emissions (TCRE) parameter and serves as a climate sensitivity measure.

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. 2021, Barnett 2023) and others have shown that this relationship is well suited for frequencies as short as one year. We follow Barnett et al. (2021) 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. Joos et al. (2013) examines carbon dynamics variation and uncertainty based on responses of atmospheric carbon concentration to emission pulses of 100 GtC 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 Arrhenius equation. We combine these two sets of pulse experiments by taking 9 different atmospheric carbon responses as inputs into 16 temperature dynamics approximations to derive 144 different carbon-climate TCRE models. Figure 1 shows the dynamic pathways and the implied TCRE parameters for these 144 models. We can see from the pathways that the temperature response peaks around 10 years, and flattens out thereafter,



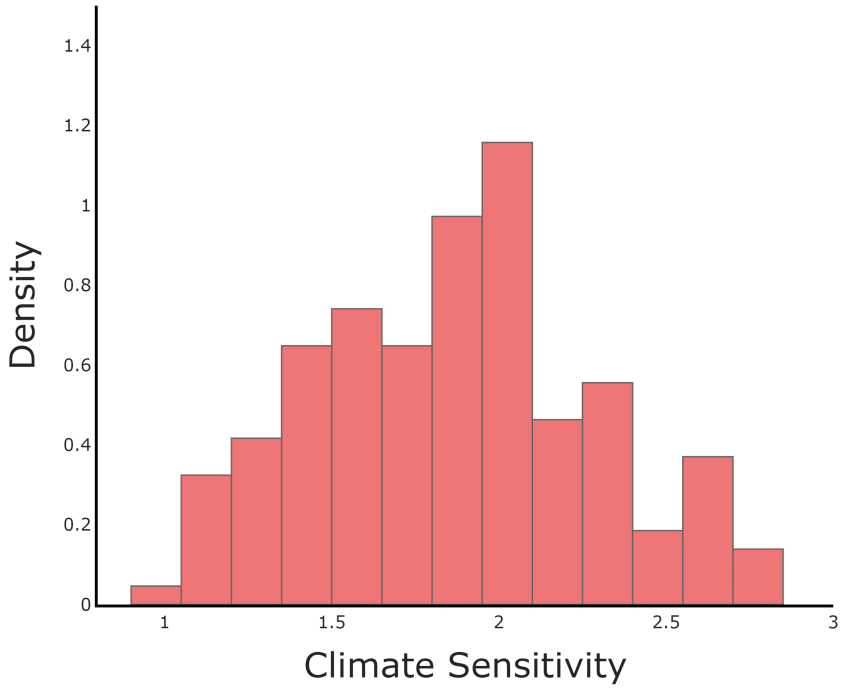


Figure 1: Implied TCRE coefficients across time and models, capturing carbon and temperature uncertainty. The top figure shows the percentiles for temperature responses to emission pulses for all carbon and temperature model combinations. The bottom figure is a histogram for the exponentially weighted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

and the implied TCRE parameters are quite dispersed, with values ranging from around 1 to just below 3, with the average value being 1.86.

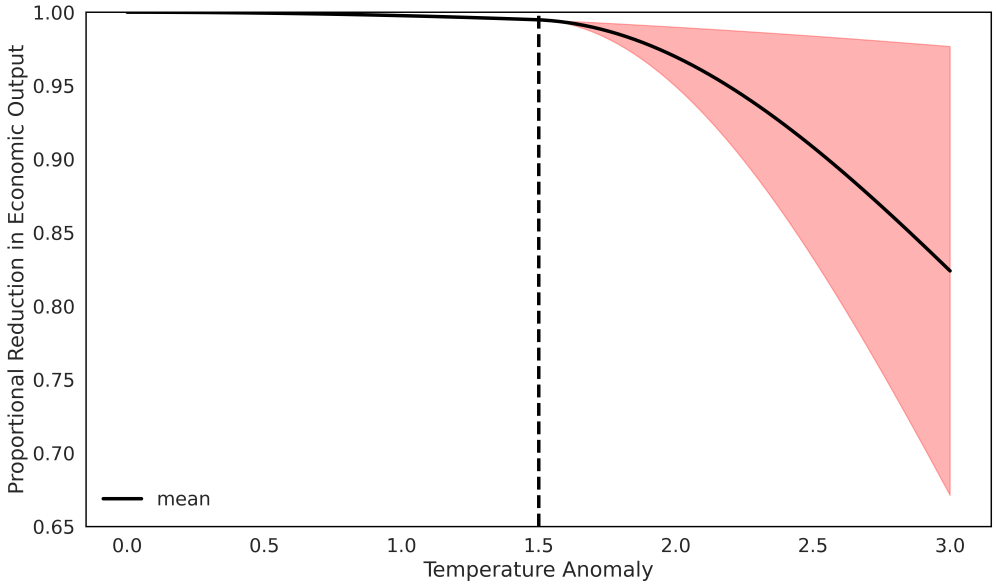
2.2 Climate Damages

Following Barnett et al. (2023), our model of economic damages from climate change, or climate damages, is piece-wise log quadratic. Damages N_t are assumed to impact capital, output, investments, and consumption in a proportional manner. This specification allows for a change in the evolution of damages that is triggered by a jump process realization. Uncertainty about climate damages is connected to the values of $\hat{Y}$, which is defined by $\hat{Y} = Y + \bar{y} - \tilde{Y}$ for $\underline{y} < \tilde{y} \leqslant \bar{y}$, and γ_3^m, for which there are alternatives values denoted by $m \in \mathcal{M}$. There is a stochastic jump process that determines dynamically which value of γ_3^m will be realized for the damage function and at the value of $\tilde{Y}$ the jump will be realized. Specifically, the jump process has m absorbing states, where each state m corresponds to a value of γ_3^m. Before the realization of the jump, each γ_3^m value has a prior probability π_m^p. When the jump is realized, the value of γ_3^m 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 $\mathcal{I}_d(y)$ that is increasing in the endogenous temperature anomaly state variable y and is localized to the range of temperature values $[\underline{y}, \overline{y}]$. At the time of the jump τ, the state variable $\hat{Y}$ shifts from the temperature anomaly $Y = \tilde{Y}_\tau$ to the damage jump threshold $\bar{y}$. Thus $\hat{Y}$ is a state variable that shifts from capturing just the temperature anomaly to a jump-adjusted measure capturing the derivative shift of the damage function as well. The uncertainty is associated both with the probability distribution on possible γ_3^m model realizations, as well as the probability of when a damage jump will occur in the $[\underline{y}, \overline{y}]$. We leave the full functional form for the Appendix, but the evolution of log damages is

$$d \log N_t = \begin{cases} \left(\gamma_1 + \gamma_2 \widehat{Y}_t\right) E_t \left(\beta_{f,\ell} dt + \varsigma dW_t\right) + \frac{\gamma_2 |\varsigma|^2 (E_t)^2}{2} dt, & t \leqslant \tau, \\ \left(\gamma_1 + \gamma_2 \widehat{Y}_t + \gamma_3^m \left(\widehat{Y}_t - \bar{y}\right)\right) E_t \left(\beta_{f,\ell} dt + \varsigma dW_t\right) + \frac{\left(\gamma_2 + \gamma_3^m\right) |\varsigma|^2 (E_t)^2}{2} dt, & t \leqslant \tau. \end{cases}$$

For our main setting, we assume $[\underline{y}, \overline{y}] = [1.5, 2.0]$, 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 γ_3^m values for the jump realization occurring at the temperature anomaly values of $\tilde{y} = \underline{y} = 1.5$ and $\tilde{y} = \bar{y} = 2$ which bound our range for the damage jumps occurring. The red shaded region shows the range of possible damage outcomes across the possible values of γ_3^m, and the black line gives the mean value.[4] While there is significant discussion about the likelihood of tipping points or thresholds at the global scale captured by discrete jumps

[4]While these plots are nearly identical to those found in Barnett et al. (2021), there are key differences. Specifically, in Barnett et al. (2021), there is a possibility for discontinuities in the damage function and jumps in the climate damages, which are ignored in their analysis. Our specification removes such discontinuities, and allows only for smooth shifts in the derivative of the damage function.



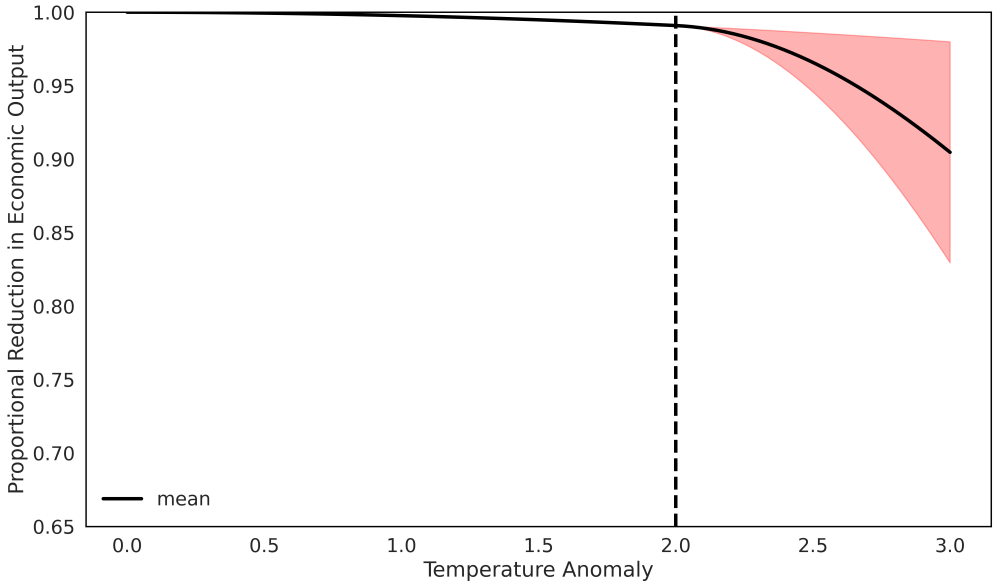


Figure 2: Range of possible damage functions for different jump thresholds for two cases. The shaded region in each plot gives the range of possible values for exp(-n), which measures the proportional reduction of the productive capacity of the economy. The top figure shows the damage function curvature when the jump occurs at $Y_t = \bar{y} = 1.5$. The bottom figure shows the damage function curvature when the jump occurs at $Y_t = \underline{y} = 2.0$.

or shifts in climate damages[5], our setting is instead a smooth shift where the planner realizes whether the damage function is more concave than previously known. In this sense, we 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. Moreover, the structure of our uncertainty is novel with regards to much of the existing climate-economics literature, as the information dynamics in most settings is static, or uncertainty about the damage function is never resolved.

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:

$$\begin{aligned} F_i(K_{i,t}) &= A_i K_{i,t}, i \in \{d, g\}, \\ F(K_{d,t}, K_{g,t}) &= F_d(K_{d,t}) + F_g(K_{g,t}) = A_d K_{d,t} + A_g K_{g,t}. \end{aligned}$$

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

$$\begin{aligned} dK_d/K_d &= [\alpha_d + \Gamma_d \log(1 + \theta_d i_d)]dt + \sigma_d dW, \\ dK_g/K_g &= [\alpha_g + \Gamma_g \log(1 + \theta_g i_g)]dt + \sigma_g dW. \end{aligned}$$

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 $R = \frac{K_g}{K_d + K_g}$. By Ito's lemma, these two state variables have the following evolution processes:

$$\begin{aligned} d\log K = {} & (1 - R)[\alpha_d + \Gamma_d \log(1 + \theta_d i_d)]dt + R[\alpha_g + \Gamma_g \log(1 + \theta_g i_g)]dt \\ & - \frac{1}{2}|\sigma_d(1 - R)|^2 - \frac{1}{2}|\sigma_g R|^2 dt \\ & + (1 - R)\sigma_d dW_d + R\sigma_g dW_g, \\ dR/R = {} & (1 - R)\{[\alpha_g + \Gamma_g \log(1 + \theta_g i_g)] - [\alpha_d + \Gamma_d \log(1 + \theta_d i_d)]\}dt \\ & + (1 - R)\{(1 - R)|\sigma_d|^2 - R|\sigma_g|^2\}dt \\ & + (1 - R)\sigma_g dW_g - (1 - R)\sigma_d dW_d. \end{aligned}$$

There are two key differences across these different consumption good production sectors in the economy. First, Sector d is the only sector that generates emissions, so that E_t in our temperature

[5]See Brook et al. (2013) and Levitan (2013) for examples on this discussion.

evolution equation is given by

$$E_t = \eta A_d K_d,$$

where η 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, we also assume that there is the potential for a "green" technology shock that would augment Sector g productivity. The arrival rate of this one-time jump in Sector g productivity is an increasing function of the aggregate knowledge capital stock in the economy κ. The evolution of the aggregate knowledge stock is given by

$$d\log\kappa = -\zeta dt + \psi_0(\frac{I_\kappa}{\kappa})^{\psi_1}dt - \frac{|\sigma_\kappa|^2}{2}dt + \sigma_\kappa \cdot dW_t,$$

where I_κ is the level of R&D investment made to increase the knowledge capital stock, ζ is the decay rate of the knowledge stock, ψ_0 and ψ_1 capture the effectiveness of R&D investment, and σ_κ is the volatility associated with the knowledge capital stock.

We include uncertainty about the realization of the technology shock in the form of a discrete set of possible realizations for the post-technology jump Sector g productivity A_g^j. The dynamic realization of the technological change shock that determines the value of A_g^j is through a stochastic jump process. There are j absorbing states, where alternative values are denoted by $j \in \mathcal{J}$, with each state j corresponding to a particular value of A_g^j and each value of A_g^j having a prior probability π_g^j. For each potential realization of the technology shock, the value of A_g^j is such that $A_g^j \geqslant A_g$. The likelihood of a technological change jump is dependent upon a jump intensity $\mathcal{I}_g(\kappa)$. We choose the simple functional form $\mathcal{I}_g(\kappa) = \kappa\varrho$, where $\kappa > 0$ is a constant, so that the arrival rate is increasing in the endogenous knowledge stock state variable κ. The information dynamics and framework and structure of the technological change jump are similar to the damage function jump. Once the jump is realized, the true technological change outcome for the value of A_g^j is known to the social planner. In addition, the uncertainty pertains to the probability distribution of potential A_g^j realizations and the arrival rate of the technological change shock. This structure focuses on the economic implications related to the uncertainty of breakthrough green technologies, again a relatively novel channel of innovation compared to the climate-economics literature. We find this specification appealing as it allows for a broader interpretation of the uncertainty of "green" technology change that further enriches our analysis and uncertainty quantification as it relates to optimal policy considerations related to technological innovation and climate change.

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

$$U(\tilde{C}) = U(C/N) = \delta \log(A_d K_d - i_d K_d + A_g K_g - i_g K_g - i_\kappa (K_g + K_d)) - \delta \log N,$$

where δ 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

$$C = A_d K_d - i_d K_d + A_g K_g - i_g K_g - i_\kappa (K_g + K_d),$$

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 $i_\kappa = I_\kappa/(K_d + K_g)$ is the R&D investment-to-total capital ratio.

3 Model Solution

With the model framework laid out, we can not 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:

$$\{i_g^*, i_d^*, i_\kappa^*\}$$

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

$$\{\log K, R, Y_t, \log \kappa, \log N_t\}.$$

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.

3.1 Post Damage and Technology Jumps HJB Equation

The first jump-state we must solve is after the technology jump and damage jump have already occurred, or the post-technology jump, post-damage-function-jump state. We denote this value function by $V^{(m,j)}(K_d, K_g, \hat{Y}, \log N)$, indicating that we are at a given realization of γ_3^m and in the second technology state with green capital productivity A_g^j. We can guess-and-verify that the value function solution in this stage can be analytically simplified as follows

$$V^{(m,j)}(K_d, K_g, \hat{Y}, \log N) = v^{(m,j)}(\log K, R, \hat{Y}) - \log N_t.$$

The remaining HJB equation characterizing the simplified value function is given by:

$$\begin{aligned}
\delta v^{(m,j)} = & \max_{i_g, i_d} \delta \log([A_d - i_d](1-R) + [A_g' - i_g]R - i_\kappa) + \delta \log K \\
& + \left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \right) v^{(m,j)}_{\log K} \\
& + \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} v^{(m,j)}_{\log K, \log K} \\
& + \left(\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - R\sigma_g^2 - \left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + (1-R)\sigma_d^2 \right) R(1-R) v_R^{(m,j)} \\
& + \frac{1}{2} R^2 (1-R)^2 (\sigma_g^2 + \sigma_d^2) v_{RR}^{(m,j)} \\
& + \left[-R(1-R)^2 \sigma_d^2 + R^2(1-R)\sigma_g^2 \right] v^{(m,j)}_{\log K, R} \\
& + \beta_f \eta A_d (1-R) K v_y^{(m,j)} + \frac{|\varsigma|^2 (\eta A_d (1-R) K)^2}{2} v_{yy}^{(m,j)} \\
& - \left(\{\gamma_1 + \gamma_2 \hat{Y} + \gamma_3^m (\hat{Y} - \underline{y})\} \beta_f \eta A_d (1-R) K + \frac{1}{2}(\gamma_2 + \gamma_3^m) |\varsigma|^2 \eta A_d (1-R) K)^2 \right).
\end{aligned}$$

Taking first order conditions with respect to i_d and i_g, we end up with the following two expressions for optimal investment choices

$$\begin{aligned}
\delta([A_d - i_d](1-R) + [A_g' - i_g]R)^{-1} &= \Gamma_d \theta_d (1+\theta_d i_d)^{-1} [v^{(m,j)}_{\log K} - R v_R^{(m,j)}], \\
\delta([A_d - i_d](1-R) + [A_g' - i_g]R)^{-1} &= \Gamma_g \theta_g (1+\theta_g i_g)^{-1} [v^{(m,j)}_{\log K} + (1-R) v_R^{(m,j)}].
\end{aligned}$$

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.

3.2 Intermediate Jump State HJB Equations

Our next two HJB equations come from the intermediate jump states. The first case is when the technology jump has already occurred, but the damage function jump has not. In this case, we can apply a similar analytical simplification to the value function for the social planner. We denote this value function by $V^{(j)}(K_d, K_{g,}Y, \log N)$, indicating that we have not had a realization of γ_3^m, but we are in the second technology state with green capital productivity A_g^j. The value function in this stage is given by

$$V^{(j)}(K_d, K_{g,}Y, \log N) = v^{(j)}(\log K, R, Y_t) - \log N_t.$$

There are two adjustments made to the simplified HJB equation in this case relative to the post-technology and -damage jump case. First, because the damage jump has not yet occurred, the equation capturing the evolution of damages is altered to be

$$-\left(\{\gamma_1 + \gamma_2 Y)\}\beta_f \eta A_d (1-R)K + \frac{1}{2}(\gamma_2)|\varsigma|^2(\eta A_d(1-R)K)^2\right).$$

Second, the expectation of damage function jump introduces the additional term

$$\mathcal{I}_d(y)\sum_{m=1}^{M}\pi_d^m(v^{(m,j)} - v^{(j)}).$$

We again take first order conditions with respect to i_d and i_g. These expressions are unchanged from the post-damage and 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 intermediate case is when the technology jump has not yet occurred, but the damage function jump has. We can again apply the same analytical simplification to the value function for the social planner. We denote this value function by $V^{(m)}(K_d, K_{g,}Y, \log\kappa, \log N)$, indicating that we have had a realization of γ_3^m, but we are in the technology state with green capital productivity A_g. The value function in this stage is given by

$$V^{(m)}(K_d, K_g, \hat{Y}, \log\kappa, \log N) = v^{(m)}(\log K, R, \hat{Y}, \log\kappa) - \log N_t.$$

Relative to the post-technology and -damage jump case, we introduce the following new terms

to the simplified HJB equation

$$(-\zeta + \psi_0 i_\kappa^{\psi_1} \exp(\psi_1(\log K - \log \kappa)) - \frac{1}{2}|\sigma_\kappa|^2)v^{(m)}_{\log \kappa} + \frac{|\sigma_\kappa|^2}{2} v^{(m)}_{\log \kappa, \log \kappa}$$
$$+ \mathcal{I}_g(\kappa) \sum_{j=1}^{J} \pi_g^j [v^{(m,j)} - v^{(m)}].$$

The first two terms account for the evolution of the knowledge capital and the possibility for R&D investment to increase the knowledge capital stock. The last term accounts for the possibility of the green technology jump that alters green capital productivity to A_g^j and is increasingly more likely as the knowledge capital stock increases.

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

$$\delta \log([A_d - i_d](1 - R) + [A_g - i_g]R - i_\kappa).$$

We now take first order conditions with respect to i_d, i_g, and i_κ, which give us expressions for optimal investment choices

$$\delta([A_d - i_d](1 - R) + [A_g - i_g]R - i_\kappa)^{-1} = \Gamma_d \theta_d (1 + \theta_d i_d)^{-1} [v^{(m)}_{\log K} - R v^{(m)}_R],$$
$$\delta([A_d - i_d](1 - R) + [A_g - i_g]R - i_\kappa)^{-1} = \Gamma_g \theta_g (1 + \theta_g i_g)^{-1} [v^{(m)}_{\log K} + (1 - R) v^{(m)}_R],$$
$$\delta([A_d - i_d](1 - R) + [A_g - i_g]R - i_\kappa)^{-1} = \psi_0 \psi_1 i_\kappa^{\psi_1 - 1} \exp(\psi_1(\log K - \log \kappa)) v^{(m)}_{\log \kappa}.$$

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.

3.3 Pre Damage and Technology Jumps HJB Equation

Finally, we have the components needed for the value function associated with pre-technology and -damage jumps. We denote this value function by $V(K_d, K_g, Y, \log \kappa, \log N)$, indicating that neither a realization of γ_3^m nor the jump to the green capital productivity has occurred. The same analytical simplification can be applied here, and so the value function in this stage is given by

$$V(K_d, K_g, Y, \log \kappa, \log N) = v(\log K, R, Y_t, \log \kappa) - \log N_t.$$

The simplified HJB equation incorporates each of the changes applied to the two intermediate

jump state cases. In particular, relative to the post-damage and technology jump state HJB equation, we introduce the following additional terms

$$
\begin{aligned}
&(-\zeta + \psi_0 i_\kappa^{\psi_1} \exp(\psi_1(\log K - \log \kappa)) - \frac{1}{2}|\sigma_\kappa|^2) v_{\log \kappa} + \frac{|\sigma_\kappa|^2}{2} v_{\log \kappa, \log \kappa} \\
&- \left(\{\gamma_1 + \gamma_2 Y)\} \beta_f \eta A_d (1-R) K + \frac{1}{2}(\gamma_2)|\varsigma|^2 (\eta A_d (1-R) K)^2 \right) \\
&+ \mathcal{I}_g(\kappa) \sum_{j=1}^{J} \pi_g^j [v^{(j)} - v] + \mathcal{I}_d(y) \sum_{m=1}^{M} \pi_d^m (v^{(m)} - v).
\end{aligned}
$$

The first line accounts for the evolution of the knowledge capital and the possibility for R&D investment to increase the knowledge capital stock. The second line captures the altered evolution of damages because the damage jump has not yet occurred. The last line accounts for the possibility of the green technology jump that alters green capital productivity to A_g^j and is increasingly more likely as the knowledge capital stock increases, as well as the expectation of damage function jump that becomes more likely as temperature increases.

We again alter the output constraint in this setting relative to the post-damage and technology jump case due to the possibility of R&D investment. As a result, the first order conditions with respect to i_d, i_g, and i_κ match those of the intermediate jump state before the technology jump and after the damage jump, except that the relevant value function derivatives do not 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.

4 Model Uncertainty

With the baseline model results in place, we now introduce the various channels of model uncertainty of focus in our analysis. We explore uncertainty aversion in the form of model misspecification and/or smooth ambiguity across the following channels:

- Climate Sensitivity to Carbon Emissions
- Climate Damage Function Severity and Arrival
- "Green" Technological Change

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. Further elaboration on alternative settings, including smooth ambiguity aversion, and complete details on the HJB and FOC equations characterizing the model solutions, is provided in the appendix.

4.1 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

$$\mathcal{I}_g(\kappa)\sum_{j=1}^{J}\pi_g^j(v^{(j)}-v)+\mathcal{I}_d(y)\sum_{m=1}^{M}\pi_d^m(v^{(m)}-v).$$

Assuming a common robustness parameter ξ across each channel, these terms are replaced by the following uncertainty-adjusted contributions

$$\begin{aligned}\min_{g_j,f_m}\ &\mathcal{I}_g(\kappa)\sum_{j=1}^{J}\pi_g^j g_j(v^{(j)}-v)+\mathcal{I}_d(y)\sum_{m=1}^{M}\pi_d^m f_m(v^{(m)}-v)\\ &+\xi\left[\mathcal{I}_g(\kappa)\sum_{j=1}^{J}\pi_g^j\left(1-g_j+g_j\log g_j\right)+\mathcal{I}_d(y)\sum\pi_d^m\left(1-f_m+f_m\log f_m\right)\right].\end{aligned}$$

The adjustment introduces probability distortions g_j and f_m 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 $\xi\left[\mathcal{I}_g(\kappa)\sum_{j=1}^{J}\pi_g^j\left(1-g_j+g_j\log g_j\right)+\mathcal{I}_d(y)\sum\pi_d^m\left(1-f_m+f_m\log f_m\right)\right]$ are introduced into preferences, penalizing distortions that are "too large" based on the uncertainty aversion parameter ξ. Note that g_j, and f_m 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 given by

$$\begin{aligned}g_j &= \exp\left(-\frac{1}{\xi}(v^{(j)}-v)\right),\\ f_m &= \exp\left(-\frac{1}{\xi}(v^{(m)}-v)\right).\end{aligned}$$

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 prob-

lem 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 technology and damage jumps have both taken place, jump uncertainty is no longer relevant for the social planner. After the technology jump has occurred, but not the damage jump, we replace A_g with A_g^j and the relevant uncertainty adjustments to the HJB equation and minimization FOC pertaining to the climate model and damage function jump are given by

$$\min_{f_m} \mathcal{I}_d(y) \sum_{m=1}^{M} \pi_d^m f_m (v^{(m,j)} - v^{(j)}) + \xi \mathcal{I}_d(y) \sum \pi_d^m \left(1 - f_m + f_m \log f_m\right),$$
$$f_m = \exp\left(-\frac{1}{\xi}(v^{(m,j)} - v^{(j)})\right).$$

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

$$\min_{g_j} \mathcal{I}_g(\kappa) \sum_{j=1}^{J} \pi_g^j g_j (v^{(m,j)} - v^{(m)}) + \xi \mathcal{I}_g(\kappa) \sum_{j=1}^{J} \pi_g^j \left(1 - g_j + g_j \log g_j\right),$$
$$g_j = \exp\left(-\frac{1}{\xi}(v^{(m,j)} - v^{(m)})\right).$$

4.2 Diffusion Misspecification Concerns

We now consider the setting with misspecification concerns about the climate and capital stocks diffusion channels. Before the technology and damage jumps occur, the relevant contributions to the HJB equation related to the diffusion components are of the form

$$\left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2}\right) v_{\log K}$$
$$+ \left(\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - R\sigma_g^2 - \left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + (1-R)\sigma_d^2\right) R(1-R) v_R$$
$$+ \left(\left[v_y - \{\gamma_1 + \gamma_2 Y\}\right] \beta_f \eta A_d (1-R) K + \frac{1}{2}\gamma_2 |\varsigma|^2 (\eta A_d (1-R) K)^2\right)$$
$$+ \left(-\zeta + \psi_0 i_\kappa^{\psi_1} \exp(\psi_1(\log K - \log \kappa)) - \frac{1}{2}|\sigma_\kappa|^2\right) v_{\log \kappa}.$$

Assuming the same common robustness parameter ξ for each channel as before, these terms

are replaced by the following uncertainty-adjusted contributions

$$\begin{aligned}
\min_h \; & \left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \right) v_{\log K} \\
& + \left(\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - R\sigma_g^2 - \left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + (1-R)\sigma_d^2 \right) R(1-R) v_R \\
& + \left(\left(v_{\log K} - R v_R\right)(1-R)\sigma_d + \left(v_{\log K} + (1-R)v_R\right) R\sigma_g \right) \cdot h \\
& + \left(\left[v_y - \{\gamma_1 + \gamma_2 Y\}\right](\beta_f + \varsigma \cdot h)\eta A_d (1-R)K + \frac{1}{2}\gamma_2 |\varsigma|^2 (\eta A_d (1-R)K)^2 \right) \\
& + \left(-\zeta + \psi_0 i_\kappa^{\psi_1} \exp(\psi_1(\log K - \log \kappa)) - \frac{1}{2}|\sigma_\kappa|^2 + \sigma_\kappa \cdot h \right) v_{\log \kappa} + \xi \frac{|h|^2}{2}.
\end{aligned}$$

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 ξ. As with g_j and f_m, 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

$$\begin{aligned}
h \;=\; & -\frac{1}{\xi}\left[\left(v_{\log K} - R v_R\right)(1-R)\sigma_d + \left(v_{\log K} + (1-R)v_R\right) R\sigma_g\right] \\
& -\frac{1}{\xi}\varsigma \eta A_d (1-R)K\left(v_y - \{\gamma_1 + \gamma_2 Y\}\right) \\
& -\frac{1}{\xi}\sigma_\kappa v_{\log \kappa}.
\end{aligned}$$

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. After the technology and damage jumps have both taken place, 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^j, and denoting the realized γ_3^m value, the uncertainty-adjusted contribution to the HJB

and the relevant minimization FOC are given by

$$\begin{aligned}
\min_h \; & \left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \right) v_{\log K}^{(m,j)} \\
& + \left(\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - R\sigma_g^2 - \left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + (1-R)\sigma_d^2 \right) R(1-R) v_R^{(m,j)} \\
& + \left(\left(v_{\log K}^{(m,j)} - R v_R^{(m,j)} \right)(1-R)\sigma_d + \left(v_{\log K}^{(m,j)} + (1-R) v_R^{(m,j)} \right) R\sigma_g \right) \cdot h \\
& + \left(\left[v_y^{(m,j)} - \{\gamma_1 + \gamma_2 \hat{Y} + \gamma_3^m (\hat{Y} - \tilde{y})\} \right] (\beta_f + \varsigma \cdot h) \eta A_d (1-R) K \right) + \xi \frac{|h|^2}{2}, \\
h = & -\frac{1}{\xi} \left[\left(v_{\log K}^{(m,j)} - R v_R^{(m,j)} \right)(1-R)\sigma_d + \left(v_{\log K}^{(m,j)} + (1-R) v_R^{(m,j)} \right) R\sigma_g \right] \\
& - \frac{1}{\xi} \varsigma \eta A_d (1-R) K \left(v_y^{(m,j)} - \{\gamma_1 + \gamma_2 \hat{Y} + \gamma_3^m (\hat{Y} - \tilde{y})\} \right).
\end{aligned}$$

After the technology jump has occurred, but not the damage jump, we replace A_g with A_g^j and the relevant uncertainty adjustments to the HJB equation and minimization FOC pertaining to the diffusion terms are given by

$$\begin{aligned}
\min_h \; & \left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \right) v_{\log K}^{(j)} \\
& + \left(\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - R\sigma_g^2 - \left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + (1-R)\sigma_d^2 \right) R(1-R) v_R^{(j)} \\
& + \left(\left(v_{\log K}^{(j)} - R v_R^{(j)} \right)(1-R)\sigma_d + \left(v_{\log K}^{(j)} + (1-R) v_R^{(j)} \right) R\sigma_g \right) \cdot h \\
& + \left(\left[v_y^{(j)} - \{\gamma_1 + \gamma_2 Y\} \right] (\beta_f + \varsigma \cdot h) \eta A_d (1-R) K \right) + \xi \frac{|h|^2}{2}, \\
h = & -\frac{1}{\xi} \left[\left(v_{\log K}^{(j)} - R v_R^{(j)} \right)(1-R)\sigma_d + \left(v_{\log K}^{(j)} + (1-R) v_R^{(j)} \right) R\sigma_g \right] \\
& - \frac{1}{\xi} \varsigma \eta A_d (1-R) K \left(v_y^{(j)} - \{\gamma_1 + \gamma_2 Y\} \right).
\end{aligned}$$

Before the technology jump has occurred, but after the damage jump, the relevant uncertainty

adjustments to the HJB equation and minimization FOC pertaining to the diffusion terms are

$$\begin{aligned}
\min_h \Bigg(& (1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \Bigg) v^{(m)}_{\log K} \\
& + \left(\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - R\sigma_g^2 - \left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + (1-R)\sigma_d^2\right) R(1-R) v^{(m)}_R \\
& + \left(\left(v^{(m)}_{\log K} - R v^{(m)}_R\right)(1-R)\sigma_d + \left(v^{(m)}_{\log K} + (1-R)v^{(m)}_R\right) R\sigma_g\right) \cdot h \\
& + \left(\left[v^{(m)}_y - \{\gamma_1 + \gamma_2 \hat{Y} + \gamma_3^m(\hat{Y} - \tilde{y})\}\right](\beta_f + \varsigma \cdot h)\eta A_d(1-R)K + \frac{1}{2}\gamma_2 |\varsigma|^2 (\eta A_d (1-R)K)^2\right) \\
& + \left(-\zeta + \psi_0 i_\kappa^{\psi_1} \exp(\psi_1(\log K - \log \kappa)) - \frac{1}{2}|\sigma_\kappa|^2 + \sigma_\kappa \cdot h\right) v^{(m)}_{\log \kappa} + \xi \frac{|h|^2}{2}, \\
h = & -\frac{1}{\xi}\left[\left(v^{(m)}_{\log K} - R v^{(m)}_R\right)(1-R)\sigma_d + \left(v^{(m)}_{\log K} + (1-R)v^{(m)}_R\right) R\sigma_g\right] \\
& - \frac{1}{\xi}\varsigma \eta A_d (1-R)K\left(v^{(m)}_y - \{\gamma_1 + \gamma_2 \hat{Y} + \gamma_3^m(\hat{Y} - \tilde{y})\}\right) \\
& - \frac{1}{\xi}\sigma_\kappa v^{(m)}_{\log \kappa}.
\end{aligned}$$

4.3 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 f_m and g_j, 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.

5 Computational Method

Before delving into numerical results, we outline the numerical method used to derive those results. The computational algorithm developed here is a critical contribution of our paper to the literature. Much of the theoretical work in climate economics and climate finance requires, by necessity, the use of computational methods to derive solutions. The integrated structure of these models requires a multi-dimensional state space, often times with non-linear dynamic relationships and functional forms, in order to rigorously account for relevant model features. The result of specifying such models is the need to confront the "curse of dimensionality", or the fact that the complexity of deriving numerical solutions exponentially increases in the number of dimensions.

As noted by (Han et al. 2018), numerical solutions are almost always unavailable for problems where the number of dimensions is greater than or equal to 4 when solving with standard finite difference or finite element methods. Incorporating model uncertainty, which introduces non-linear endogenous responses to model uncertainty, such as exponential tilting to the distribution of expected future model outcomes, can further exacerbate the computational burden. As we will outline below, we provide a novel implementation of deep neural networks for deriving the numerical solutions to our HJB equations.

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 with high dimensions where the "curse of dimensionality" can make computational solutions infeasible for other methods, or where significant non-linearities can cause other computational methodologies to fail, deep learning solution methods have been seen as a potentially "game-changing" toolset. By approximating the representative agent's value function and optimal controls with deep neural networks, the problems of exponentially increasing complexity from high-dimensional state spaces 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) and finance (Duarte et al. 2023, Sauzet 2021) implementing these types of solutions methods are only beginning to scratch the surface of the potential value and importance for study key problems in the literature more broadly.

While this existing work has explored settings with (potentially many) more state variables than our current setting, some rely on assumptions such as state variables that are symmetric, stationary, or exogenously specificed, as well as linearity in the model framework, to maintain tractability with such scale. Importantly, our methodology relaxes such assumptions, which is essential to analyzing our climate-economics model with model uncertainty aversion. 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 such as ours. 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 sufficiently handle the nonlinearities arising from our application of dynamic decision theory to address concerns about model uncertainty.

We address these issues in our framework by implementing an extended deep Galerkin method algorithm that considers critical parameters as additional (pseudo) 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 and multiple endogenous state variables 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, avoiding 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 areas of research that require solving dynamic stochastic optimal control problems.

5.1 Implementing Neural Nets for Numerical Solutions

We now outline our algorithm, and provide pseudo-code and further details about implementation in the appendix. We use 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 δ 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^{θ} 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 θ_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^{θ_n}, by update φ_{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 Ω of $\boldsymbol{x}$ characterizing the different regions' relative importance.

In our problem, depending on the setting we are solving, the state processes could contain $(\log K, R, Y_t, \log \kappa, \log N_t)$, the control variables could contain $(i_d, i_g, i_\kappa, g_j, f_m, h)$. In addition, in order to keep the value function's monotonicity with respect to γ_3^m, we take γ_3^m as an input or "pseudo-state" of the parameterized neural network.

6 Numerical Results

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

6.1 Functional Forms and Assumptions

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

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

Second, we use the following functional forms for our jump arrival rates. For the technology change jump, we assume an arrival rate that is proportional to the knowledge stock: $\mathcal{I}_g(\kappa) = \kappa/\varrho$. The parameter ϱ scales the knowledge capital stock variable to change the units into arrival rate

units, and is chosen based on expected green policy implementation timelines proposed by various countries. Section 6.2 provides further details about the choice of this parameter value.

The damage jump intensity $\mathcal{I}_d(y)$ follows the arrival rate proposed in Barnett et al. (2021):

$$\mathcal{I}_d(y) = r_1 \left(\exp\left[\frac{r_2}{2} (y - \underline{y})^2 \right] - 1 \right) \mathbb{1}_{y \geqslant \underline{y}}.$$

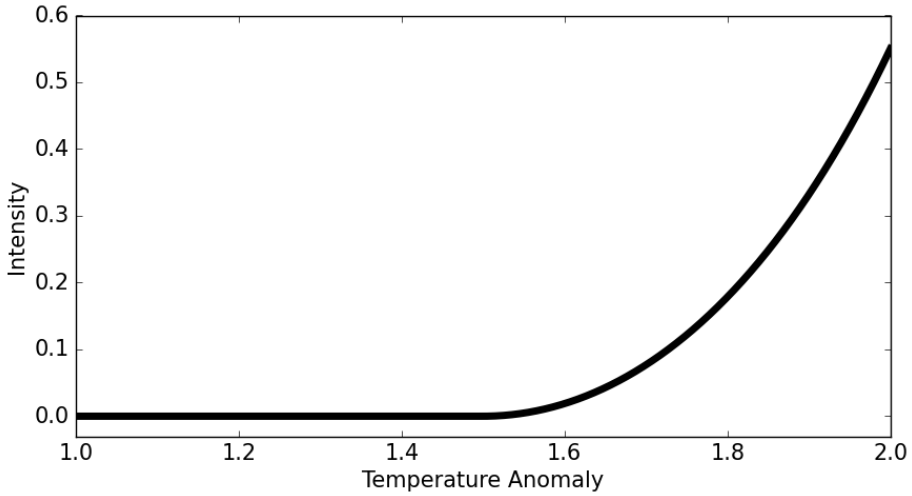


Figure 3: Intensity function, $r_1 = 1.5$ and $r_2 = 2.5$. With this intensity function, the probability of a jump at an anomaly of 1.6 is approximately .02 per annum, increasing to about .08 per annum at an anomaly of 1.7, increasing further to approximately .18 per annum at an anomaly of 1.8 and then to about one third per annum when the anomaly is 1.9.

Figure 3 shows the increasing nature of the arrival rate as temperature anomaly y increases. The calibration is such that the probability of a jump taking place by $y = 2$ is essentially 1.

6.2 Parameter Values

We now outline how we chose the parameter values used in our numerical analysis. The economic parameter values are given in Table 1. We note that the special case of our model without climate change and technological innovation is the model given in Eberly and Wang (2009). We therefore use their assumed parameters values of the relevant economic parameters as our baseline solution values. These parameters are not without justification from empirical estimates and previous modeling set-ups. The choice of δ is consistent with values of the subjective discount rate used in the macroeconomics and asset pricing literature. (Γ_d, θ_d) produces a value of Tobin's q of 2.5, which is consistent with estimated values in the macroeconomics literature. The values of σ_d and σ_g are within the range of values used in the macro-asset pricing literature to produce values of the price of risk observed in the data. The values of A_d and A_g generate output consistent with

World Bank World GDP values. Way et al. (2022) provide estimates of future costs of green technology across various scenarios (fast transition, slow transition, no transition) by estimating coefficients for stochastic Wright's law and Moore's law. The values for A_g^j coincide with implied productivity gains in different green technologies over the next 50 years for these estimates.

Table 1: Economic Parameters

Parameters	values
δ	0.025
$(\alpha_d, \Gamma_d, \theta_d, \sigma_d)$	(-0.035, 0.025, 100, 0.15)
$(\alpha_g, \Gamma_g, \theta_g, \sigma_g)$	(-0.035, 0.025, 100, 0.15)
(A_d, A_g)	(0.12, 0.10)
$\{A_g^j\}$	$A_g \times \{1 + \frac{(j-1)}{J-1}\}_{j=1,\ldots,3}$
$(\zeta, \psi_0, \psi_1, \sigma_\kappa)$	(0, 0.10583, 0.5, 0.016)
ϱ	448

We assume $\zeta = 0$. Lucking et al. (2019) and Bloom et al. (2019) have provided estimates for the returns to R&D investment which guide our choice of (ψ_0, ψ_1). The choice of ϱ translates our initial value of knowledge stock, which is based on estimates from the BLS of the total US R&D stock values (scaled up to World values), to an expected arrival time of a green technological innovation occurring between 30 and 80 years. The value of σ_κ is chosen to match the other capital volatilities, which are based on estimates from the World Bank database.

Table 2: Climate Dynamics and Damages Parameters

Parameters	values
β_f	1.86 / 1000
η	0.17
ς	$1.2 \times 1.86/1000$
(γ_1, γ_2)	$(0.00017675, 2 \times 0.0022)$
$\{\gamma_3^m\}$	$\{\frac{1}{3}\frac{m-1}{M-1}\}_{m=1,\ldots,5}$
$(r_1, r_2, \underline{y})$	$(1.5, 2.5, 1.5)$
$\bar{y}$	2

The climate dynamics and climate damage parameter values are given in Table 2. The values of $\beta_{f,\ell}$ come from pulse experiments estimates produced by Barnett et al. (2021), 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 β_f is the average value across all climate 144 models. The value of η is chosen to match the current estimate of annual carbon emissions of 10

GtC from Figueres et al. (2018), based on World Bank estimates of World GDP and EIA/IEA estimates of the clean and dirty capital split. The value of ς matches volatility used by Barnett et al. (2021). The values of γ_1, γ_2, γ_3^m, and $\bar{y}$ match damage function parameters used by Barnett et al. (2021), which are designed to incorporate the spread of potential climate damage outcomes based on Nordhaus (2019), Weitzman (2012), and others in the literature.

Table 3: State Variable Initial Values and Ranges

State variables	values
K_0	739
R_0	0.5
Y_0	1.1
κ_0	11.2
State variables	range
$\log(K)$	$[4, 8.5]$
R	$[0.01, 0.99]$
Y	$[0, 4]$
$\log(\kappa)$	$[1, 6]$

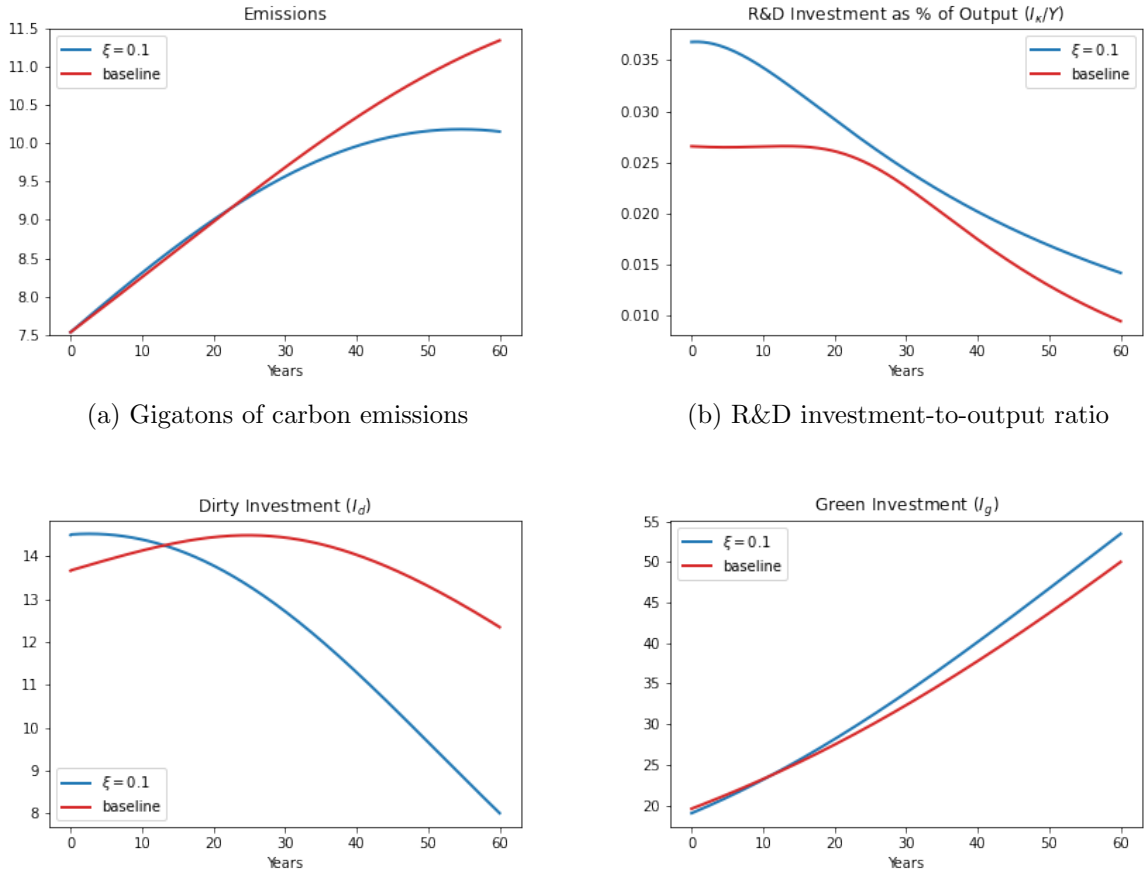
For our computations, we must also specify ranges and initial values for our state variables. These values are given in Table 3. The initial value of total capital K_0 matches estimates from the World Bank of World GDP. The initial value of global mean temperature anomaly Y_0 matches the estimated current value from Masson-Delmotte et al. (2021b). The initial value of the green capital-to-total capital ratio R_0 is based on estimates of clean and dirty capital splits from the EIA and IEA. The initial value of knowledge stock κ_0 matches estimates from the BLS of Total US R&D stock values (scaled up to World values).

6.3 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 30 years, which is near the time when the temperature anomaly hits $1.5^{\circ}C$ and the probability of a damage jump occurring becomes non-zero. For the probability of the technology and damage jumps occurring, we examine the outcomes out to 40 years. The results provided are for the following specifications related to model uncertainty:

- β_f as the average of the 144 climate models
- 20 damage models $\gamma_3 \in \{0, ..., 1/3\}$

- Misspecification over technology and damage jumps and the climate model dynamics
- Comparison across uncertainty parameters $\xi \in \{0.1, \infty\}$



(a) Gigatons of carbon emissions

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

(c) Dirty capital investment

(d) Green capital investment

Figure 4: Simulated outcomes under different uncertainty penalty configurations based on the numerical solutions. Panel (a) shows the pathway for carbon emissions. Panel (b) shows the pathway for R&D investment as a fraction of total output. Panel (c) shows the level of dirty capital investment. Panel (d) shows the level of green capital investment. The trajectories are simulated under the baseline probabilities abstracting from the intrinsic randomness. The pathways stop after 30 years of simulated outcomes, near the time that the temperature anomaly reaches $1.5^{\circ}C$.

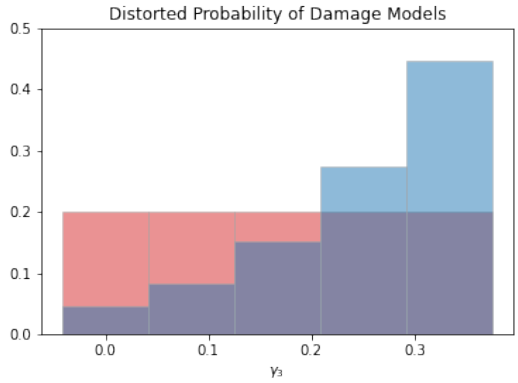
We focus on the simulated pathways from the pre-damage jump and pre-technology jump state for emissions, R&D investment, green and dirty capital investment, as well the distribution of climate and damage models, and the probabilities of a damage or technology jump occurring. We examine each of these results across different uncertainty aversion parameters to highlight the implications of increased model uncertainty on the optimal policy choices by our planner. The

red lines, or red histogram bars, represent the uncertainty neutral case when $\xi = \infty$, and the blue lines, or blue histogram bars, represent the uncertainty averse case when $\xi = 0.1$.

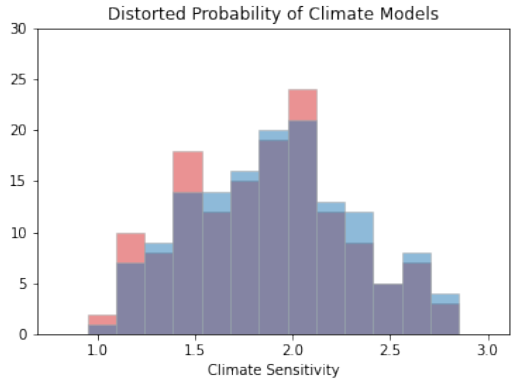
Figure 4 provides the main economic policy outcomes of interest. Figures 4a and 4b show the pathways for emissions and R&D investment across the uncertainty averse and uncertainty neutral cases. First, in the top left panel we see that emissions in each case start out similarly, and though they are increasing over time for both cases, we can see that the emissions diverge and are lower for the uncertainty averse case of $\xi = 0.1$. In the top right panel, we see 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 2.5% and 3.5% 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 near 0.5%, with total R&D levels reaching over 2%. We can also see that when the planner is concerned about model uncertainty, i.e., for $\xi = 0.1$, the fraction of output committed to R&D investment starts substantially higher, and though the R&D investment beings to decrease before the uncertaintry neutral case, it always remains higher when there are concerns about model uncertainty.

Figures 4c and 4d show the levels of dirty and clean investment across the uncertainty aversion cases. Comparing the dirty investment in the bottom left panel and the clean investment in the bottom right panel, we see three key points. First, 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 50% higher, but by the end of the pathway, it is nearly four to five times higher. This highlights the second key observation related to investment, that green investment persistently increases during the simulation pathway, whereas dirty investment begins to taper off and diminish over time. Finally, we see that uncertainty concerns lead to initially higher dirty investment, that decreases more rapidly than in the uncertainty neutral case, whereas there is only a slight amplification of green investment due to difference in uncertainty aversion.

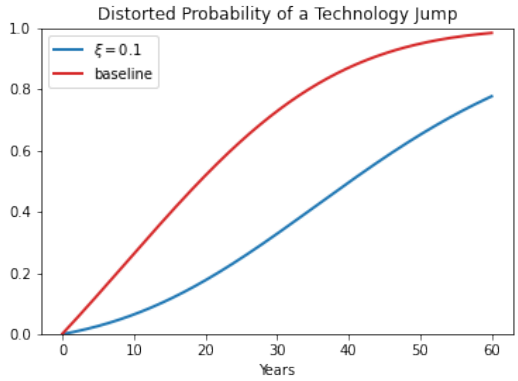
From Figure 5 we can see why the planner responds the way they do. The histograms in Figures 5a and 5b show the baseline and distorted probabilities given to damage models (left panel) and climate models (right panel) by the planner. For each type of model, the planner adjusts the probability to give more weight to the right end of the distribution where the implied level of climate damage and climate change are both more severe. The effect is fairly modest for the climate models, and somewhat more pronounced for the damage models. Thus, the planner has an increased incentive to move away from dirty capital and towards green capital when concerns about model uncertainty are present. The pathways in Figures 5c and 5d show the baseline and distorted probability of a technology jump (left panel) and damage jump (right panel) taking place by the planner. We see only a small adjustment to the jump probability for damages. However, the uncertainty averse planner significantly down-weights the probability of a technology jump occurring when they incorporate concerns about model uncertainty. This adjustment is the larger



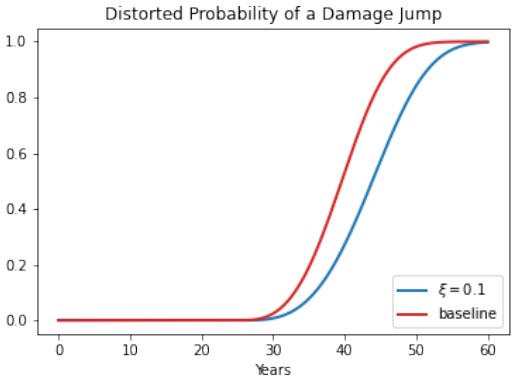
(a) Gigatons of carbon emissions



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



(c) Dirty capital investment



(d) Green capital investment

Figure 5: Distorted model distributions and distorted probability of a jump under different uncertainty penalty configurations based on the numerical solutions. Panel (a) shows the distorted distribution of damage models. Panel (b) shows the distorted distribution of climate models. Panel (c) shows the distorted probability of a technology jump occurring. Panel (d) shows the distorted probability of a damage jump occurring. The trajectories are simulated under the baseline probabilities abstracting from the intrinsic randomness. The histograms are calculated at year 40, and the pathways stop after 40 years of simulated outcomes.

distributional impact in relative terms, and highlights how the optimal endogenous response by the planner for determining robust policy is to focus their model uncertainty concerns on the technology jump. However, rather than reduce their R&D investment, the planner emphasizes even further the technological change channel and increases their R&D investment to try and increase the likelihood of the innovation shock occurring.

In summary, it is clear that climate concerns lead to substantial policy action, where the social planner allocates significant resources to R&D investment and green capital investment, while diverting some of those resources away from dirty investment. However, concerns about uncertainty aversion have relatively modest impacts on investment choices in clean production capital, instead focusing somewhat more on reducing dirty production capital and leaning even more heavily into the policy response of amplifying R&D investment. The social optimality of this response is driven by the potentially significant payoff from technological innovation in this setting, the fact that emissions are sticky in the sense that they are proportional to output from production using dirty capital and there is no direct mechanism for removing or transforming capital, and the costly nature of accumulating new green capital that is less productive than dirty capital until a technology jump occurs. As a result, the planner places a substantial social valuation on R&D investment as a policy response tool in order to potentially initiate the technology jump as relying mainly on emissions reductions through additional dirty capital investment reduction or further investment in green capital.

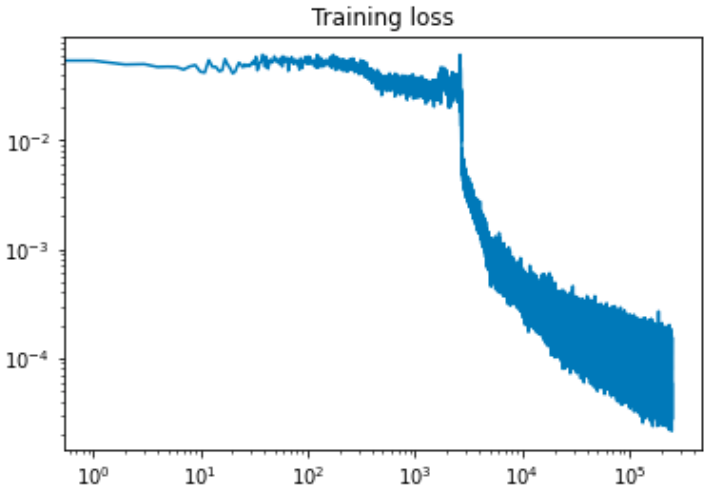
6.4 Validation of Neural Network Solutions

An important issue to address for our numerical solutions is the validation of the accuracy of our results. We address this in three ways, with the results shown in Figure 6. First, 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 6a shows the training error for the post-technology, post-damage jump state solution across epochs. We are able to reach a training error of approximately 10^{-4}, suggesting our solution is reasonably accurate.

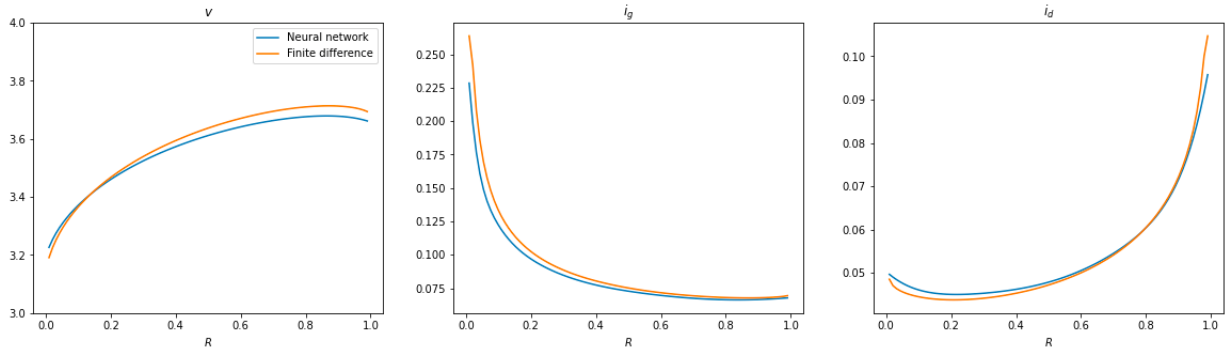
In addition, because the post-jump model solutions require only 3 state variables, we can compare and validate these results with solutions derived using the solution derived from more standard finite difference methods. In particular, we use the false-transient, conjugate gradient-based numerical algorithm implemented in Barnett et al. (2020) and Barnett et al. (2021) for our finite difference solutions. Figure 6b 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 methods provide consistent outcomes.

Finally, for the pre-technology, pre-damage jump state, we note that the training errors are similar to the post-damage, post-technology jump state, but we cannot derive finite difference

method solutions for this setting. Therefore, we propose the DGM-PIA solution at a given point in the state space to the solution derived from another deep learning method proposed in Han and E (2016), in which the authors directly parameterize the control by neural nets and obtain the optimal parameters by maximizing the approximated utility. This method is known to be very stable, though significantly slower, because the solution is derived point-by-point via training neural nets across repeated simulations of the state variables. We provide these results in future work, though preliminary results show the solutions for the two methods are roughly consistent.



(a) HJB Error for post-technology, post-damage jump state solution



(b) Finite difference and DGM-PIA solution comparison

Figure 6: Model solution validation. The top figure plots the training error over epochs for the post-damage and post-technology jump state. The bottom figures plot the value function and optimal controls for the FD and NN solutions in the post-technology and post-damage jump state. Still to come is the Han and E (2016) solution for a given point in the state space.

7 Conclusion

Model uncertainty along multiple dimensions is a central issue when studying optimal climate policy for a carbon-neutral transition. We integrate 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 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 as the major policy mechanism.

Given the richness of our economic and geoscientific model components, we develop and implement a deep learning-based algorithm to solve the Hamilton-Jacobi-Bellman equations that characterize the planner's value function in our setting. The algorithm allows us to derive global solutions for models that have many endogenous and non-stationary state variables, maximize lifetime expected utility over an infinite horizon, are in continuous-time, and have potentially significant non-linearities. Because the computational method can handle such computationally difficult problems without being bogged down by "the curse of dimensionality" or other numerical issues, we see significant potential for our algorithm in addressing other economically and geoscientifically rich climate-economic problems. For the same reasons, our computational method also has significant promise for recursive, dynamic general equilibrium models used across economics and finances. 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.

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, Lars Peter Hansen, et al. 2021. Climate Change Uncertainty Spillover in the Macroeconomy. *Prepared for the 2021 Macoreconomics Annual* .

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.

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.

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.

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.

Eberly, Janice and Neng Wang. 2009. Reallocating and pricing illiquid capital. In *American Economic Review, Papers and Proceedings*.

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.

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.

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, NIPS, 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, 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.

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.

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* .

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.

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.

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.

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.

Way, Rupert, Matthew C Ives, Penny Mealy, and J Doyne Farmer. 2022. Empirically grounded technology forecasts and the energy transition. *Joule* 6 (9):2057–2082.

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

This appendix is provided in support of the paper "A Deep Learning Analysis of Climate Change, Innovation, and Uncertainty" by Michael Barnett, William Brock, Lars Peter Hansen, Ruimeng Hu, and Joseph Huang. Included here are details on theoretical derivations, alternative model settings, the model parameters, and the numerical methods used in the paper.

Appendix A Full HJB Equations

A.1 No Model Uncertainty Aversion

A.1.1 Post Damage and Technology Jumps HJB Equation

We provide this HJB equation in the main text.

A.1.2 Intermediate Jump State HJB Equations

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

$$
\begin{aligned}
\delta v^{(j)} = & \max_{i_g, i_d} \delta \log([A_d - i_d](1-R) + [A_g' - i_g]R) + \delta \log K \\
& + \left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \right) v^{(j)}_{\log K} \\
& + \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} v^{(j)}_{\log K, \log K} \\
& + \left([\alpha_g + \Gamma_g \log(1+\theta_g i_g)] - R\sigma_g^2 - [\alpha_d + \Gamma_d \log(1+\theta_d i_d)] + (1-R)\sigma_d^2 \right) R(1-R) v^{(j)}_R \\
& + \frac{1}{2} R^2 (1-R)^2 (\sigma_g^2 + \sigma_d^2) v^{(j)}_{RR} \\
& + \left[-R(1-R)^2 \sigma_d^2 + R^2(1-R)\sigma_g^2 \right] v^{(j)}_{\log K, R} \\
& + \beta_f \eta A_d (1-R) K v_y + \frac{|\varsigma|^2 (\eta A_d (1-R) K)^2}{2} v^{(j)}_{yy} \\
& - \left(\{\gamma_1 + \gamma_2 Y\} \beta_f \eta A_d (1-R) K + \frac{1}{2} \gamma_2 |\varsigma|^2 (\eta A_d (1-R) K)^2 \right) \\
& + \mathcal{I}_d(y) \sum_{m=1}^{M} \pi_d^m (v^{(m,j)} - v^{(j)}).
\end{aligned}
$$

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

$$
\begin{aligned}
\delta v^{(m)} = & \max_{i_g, i_d} \delta \log([A_d - i_d](1-R) + [A_g - i_g]R - i_\kappa) + \delta \log K \\
& + \left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \right) v^{(m)}_{\log K} \\
& + \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} v^{(m)}_{\log K, \log K} \\
& + \left(\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - R\sigma_g^2 - \left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + (1-R)\sigma_d^2 \right) R(1-R) v^{(m)}_R \\
& + \frac{1}{2} R^2 (1-R)^2 (\sigma_g^2 + \sigma_d^2) v^{(m)}_{RR} \\
& + \left[-R(1-R)^2 \sigma_d^2 + R^2 (1-R)\sigma_g^2 \right] v^{(m)}_{\log K, R} \\
& + \beta_f \eta A_d (1-R) K v^{(m)}_y + \frac{|\varsigma|^2 (\eta A_d (1-R) K)^2}{2} v^{(m)}_{yy} \\
& - \left(\{\gamma_1 + \gamma_2 \hat{Y} + \gamma_3^m (\hat{Y} - \underline{y})\} \beta_f \eta A_d (1-R) K + \frac{1}{2}(\gamma_2 + \gamma_3^m) |\varsigma|^2 (\eta A_d (1-R) K)^2 \right) \\
& + (-\zeta + \psi_0 i_\kappa^{\psi_1} \exp(\psi_1(\log K - \log \kappa)) - \frac{1}{2}|\sigma_\kappa|^2) v^{(m)}_{\log \kappa} + \frac{|\sigma_\kappa|^2}{2} v^{(m)}_{\log \kappa, \log \kappa} \\
& + \mathcal{I}_g(\kappa)(v^{(m,j)} - v^{(m)}).
\end{aligned}
$$

A.1.3 Pre Damage and Technology Jumps HJB Equation

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

$$
\begin{aligned}
\delta v = & \max_{i_g, i_d} \delta \log([A_d - i_d](1-R) + [A_g - i_g]R - i_\kappa) + \delta \log K \\
& + \left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \right) v_{\log K} \\
& + \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} v_{\log K, \log K} \\
& + \left(\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - R\sigma_g^2 - \left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + (1-R)\sigma_d^2 \right) R(1-R) v_R \\
& + \frac{1}{2} R^2 (1-R)^2 (\sigma_g^2 + \sigma_d^2) v_{RR} \\
& + \left[-R(1-R)^2 \sigma_d^2 + R^2 (1-R)\sigma_g^2 \right] v_{\log K, R} \\
& + \beta_f \eta A_d (1-R) K v_y + \frac{|\varsigma|^2 (\eta A_d (1-R) K)^2}{2} v_{yy} \\
& - \left(\{\gamma_1 + \gamma_2 Y\} \beta_f \eta A_d (1-R) K + \frac{1}{2} \gamma_2 |\varsigma|^2 (\eta A_d (1-R) K)^2 \right) \\
& + (-\zeta + \psi_0 i_\kappa^{\psi_1} \exp(\psi_1(\log K - \log \kappa)) - \frac{1}{2}|\sigma_\kappa|^2) v_{\log \kappa} + \frac{|\sigma_\kappa|^2}{2} v_{\log \kappa, \log \kappa} \\
& + \mathcal{I}_g(\kappa)(v^{(j)} - v) + \mathcal{I}_d(y) \sum_{m=1}^{M} \pi_d^m (v^{(m)} - v).
\end{aligned}
$$

A.2 Full Misspecification Concerns

A.2.1 Post Damage and Technology Jumps Setting

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

$$
\begin{aligned}
\delta v^{(m,j)} = & \max_{i_g, i_d} \min_{h} \delta \log([A_d - i_d](1-R) + [A_g^j - i_g]R) + \delta \log K \\
& + \left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \right) v_{\log K}^{(m,j)} \\
& + \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} v_{\log K, \log K}^{(m,j)} \\
& + \left([\alpha_g + \Gamma_g \log(1+\theta_g i_g)] - R\sigma_g^2 - [\alpha_d + \Gamma_d \log(1+\theta_d i_d)] + (1-R)\sigma_d^2 \right) R(1-R) v_R^{(m,j)} \\
& + \frac{1}{2} R^2 (1-R)^2 (\sigma_g^2 + \sigma_d^2) v_{RR}^{(m,j)} \\
& + \left(\left(v_{\log K}^{(m,j)} - R v_R^{(m,j)} \right)(1-R)\sigma_d + \left(v_{\log K}^{(m,j)} + (1-R) v_R^{(m,j)} \right) R\sigma_g \right) \cdot h \\
& + \left[-R(1-R)^2 \sigma_d^2 + R^2 (1-R)\sigma_g^2 \right] v_{\log K, R}^{(m,j)} \\
& + (\beta_f + \varsigma \cdot h)\eta A_d (1-R) K v_y^{(m,j)} + \frac{|\varsigma|^2 (\eta A_d (1-R) K)^2}{2} v_{yy}^{(m,j)} \\
& - \left(\{\gamma_1 + \gamma_2 \hat{Y} + \gamma_3^m (\hat{Y} - \bar{y})\}(\beta_f + \varsigma \cdot h)\eta A_d (1-R) K + \frac{1}{2}(\gamma_2 + \gamma_3^m)|\varsigma|^2 (\eta A_d (1-R) K)^2 \right) + \xi \frac{|h|^2}{2}.
\end{aligned}
$$

FOC

$$
\begin{aligned}
h = & -\frac{1}{\xi}\left[\left(v_{\log K}^{(m,j)} - R v_R^{(m,j)} \right)(1-R)\sigma_d + \left(v_{\log K}^{(m,j)} + (1-R) v_R^{(m,j)} \right) R\sigma_g \right] \\
& -\frac{1}{\xi} \varsigma \eta A_d (1-R) K \left(v_y^{(m,j)} - \{\gamma_1 + \gamma_2 \hat{Y} + \gamma_3^m (\hat{Y} - \bar{y})\} \right).
\end{aligned}
$$

A.2.2 Pre Damage and Post Technology Jumps Setting

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

$$
\begin{aligned}
\delta v^{(j)} = & \max_{i_g, i_d} \min_{f_m, h} \delta \log([A_d - i_d](1-R) + [A_g^j - i_g]R - i_\kappa) + \delta \log K \\
& + \left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \right) v^{(j)}_{\log K} \\
& + \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} v^{(j)}_{\log K, \log K} \\
& + \left(\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - R\sigma_g^2 - \left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + (1-R)\sigma_d^2 \right) R(1-R) v^{(j)}_R \\
& + \frac{1}{2} R^2 (1-R)^2 (\sigma_g^2 + \sigma_d^2) v^{(j)}_{RR} \\
& + \left(\left(v^{(j)}_{\log K} - R v^{(j)}_R \right)(1-R)\sigma_d + \left(v^{(j)}_{\log K} + (1-R) v^{(j)}_R \right) R \sigma_g \right) \cdot h \\
& + \left[-R(1-R)^2 \sigma_d^2 + R^2 (1-R) \sigma_g^2 \right] v^{(j)}_{\log K, R} \\
& + (\beta_f + \varsigma \cdot h) \eta A_d (1-R) K v^{(j)}_y + \frac{|\varsigma|^2 (\eta A_d (1-R) K)^2}{2} v^{(j)}_{yy} \\
& - \left(\{\gamma_1 + \gamma_2 Y\}(\beta_f + \varsigma \cdot h)\eta A_d (1-R) K + \frac{1}{2}\gamma_2 |\varsigma|^2 (\eta A_d (1-R) K)^2 \right) + \xi \frac{|h|^2}{2} \\
& + \mathcal{I}_d(y) \sum_{m=1}^{M} \pi_d^m f_m (v^{(m,j)} - v^{(j)}) + \xi \mathcal{I}_d(y) \sum \pi_d^m \left(1 - f_m + f_m \log f_m\right).
\end{aligned}
$$

FOC

$$
\begin{aligned}
f_m &= \exp\left(-\frac{1}{\xi_m} (v^{(m,j)} - v^{(j)}) \right), \\
h &= -\frac{1}{\xi} \left[\left(v^{(j)}_{\log K} - R v^{(j)}_R \right)(1-R)\sigma_d + \left(v^{(j)}_{\log K} + (1-R) v^{(j)}_R \right) R \sigma_g \right] \\
&\quad -\frac{1}{\xi} \varsigma \eta A_d (1-R) K \left(v^{(j)}_y - \{\gamma_1 + \gamma_2 Y\} \right).
\end{aligned}
$$

A.2.3 Post Damage and Pre Technology Jumps Setting

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

$$
\begin{aligned}
\delta v^{(m)} = & \max_{i_g, i_d, i_\kappa} \min_{g_j, h} \delta \log([A_d - i_d](1-R) + [A_g - i_g]R - i_\kappa) + \delta \log K \\
& + \left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \right) v^{(m)}_{\log K} \\
& + \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} v^{(m)}_{\log K, \log K} \\
& + \left(\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - R\sigma_g^2 - \left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + (1-R)\sigma_d^2 \right) R(1-R) v^{(m)}_R \\
& + \frac{1}{2} R^2 (1-R)^2 (\sigma_g^2 + \sigma_d^2) v^{(m)}_{RR} \\
& + \left(\left(v^{(m)}_{\log K} - R v^{(m)}_R \right)(1-R)\sigma_d + \left(v^{(m)}_{\log K} + (1-R) v^{(m)}_R \right) R \sigma_g \right) \cdot h \\
& + \left[-R(1-R)^2 \sigma_d^2 + R^2 (1-R) \sigma_g^2 \right] v^{(m)}_{\log K, R} \\
& + (\beta_f + \varsigma \cdot h) \eta A_d (1-R) K v^{(m)}_y + \frac{|\varsigma|^2 (\eta A_d (1-R) K)^2}{2} v^{(m)}_{yy} \\
& - \left(\{\gamma_1 + \gamma_2 \hat{Y} + \gamma_3^m (\hat{Y} - \bar{y})\} (\beta_f + \varsigma \cdot h) \eta A_d (1-R) K + \frac{1}{2} (\gamma_2 + \gamma_3^m) |\varsigma|^2 (\eta A_d (1-R) K)^2 \right) + \xi \frac{|h|^2}{2} \\
& + \left(-\zeta + \psi_0 i_\kappa^{\psi_1} \exp(\psi_1 (\log K - \log \kappa)) - \frac{1}{2} |\sigma_\kappa|^2 + \sigma_\kappa \cdot h \right) v^{(m)}_{\log \kappa} + \frac{|\sigma_\kappa|^2}{2} v^{(m)}_{\log \kappa, \log \kappa} \\
& + \mathcal{I}_g(\kappa) \sum_{j=1}^{J} \pi_g^j g_j \left(v^{(m,j)} - v^{(m)} \right) + \xi \mathcal{I}_g(\kappa) \sum_{j=1}^{J} \pi_g^j \left(1 - g_j + g_j \log g_j \right).
\end{aligned}
$$

FOC

$$
\begin{aligned}
g_j &= \exp\left(-\frac{1}{\xi_m} (v^{(m,j)} - v^{(m)}) \right), \\
h &= -\frac{1}{\xi} \left[\left(v^{(m)}_{\log K} - R v^{(m)}_R \right)(1-R)\sigma_d + \left(v^{(m)}_{\log K} + (1-R) v^{(m)}_R \right) R \sigma_g \right] \\
&\quad -\frac{1}{\xi} \varsigma \eta A_d (1-R) K \left(v^{(m)}_y - \{\gamma_1 + \gamma_2 \hat{Y} + \gamma_3^m (\hat{Y} - \bar{y})\} \right) \\
&\quad -\frac{1}{\xi} \sigma_\kappa v^{(m)}_{\log \kappa}.
\end{aligned}
$$

A.2.4 Pre Damage and Technology Jumps Setting

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

$$
\begin{aligned}
\delta v = & \max_{i_g, i_d, i_\kappa} \min_{g_j, f_m, h} \delta \log([A_d - i_d](1-R) + [A_g - i_g]R - i_\kappa) + \delta \log K \\
& + \left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \right) v_{\log K} \\
& + \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} v_{\log K, \log K} \\
& + \left(\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - R\sigma_g^2 - \left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + (1-R)\sigma_d^2 \right) R(1-R) v_R \\
& + \frac{1}{2} R^2 (1-R)^2 (\sigma_g^2 + \sigma_d^2) v_{RR} \\
& + \left((v_{\log K} - R v_R)(1-R)\sigma_d + (v_{\log K} + (1-R)v_R) R \sigma_g \right) \cdot h \\
& + \left[-R(1-R)^2 \sigma_d^2 + R^2 (1-R)\sigma_g^2 \right] v_{\log K, R} \\
& + (\beta_f + \varsigma \cdot h) \eta A_d (1-R) K v_y + \frac{|\varsigma|^2 (\eta A_d (1-R) K)^2}{2} v_{yy} \\
& - \left(\{\gamma_1 + \gamma_2 Y\} (\beta_f + \varsigma \cdot h) \eta A_d (1-R) K + \frac{1}{2} \gamma_2 |\varsigma|^2 (\eta A_d (1-R) K)^2 \right) \\
& + \left(-\zeta + \psi_0 i_\kappa^{\psi_1} \exp(\psi_1(\log K - \log \kappa)) - \frac{1}{2}|\sigma_\kappa|^2 + \sigma_\kappa \cdot h \right) v_{\log \kappa} + \frac{|\sigma_\kappa|^2}{2} v_{\log \kappa, \log \kappa} \\
& + \mathcal{I}_g(\kappa) \sum_{j=1}^{J} \pi_g^j g_j \left(v^{(j)} - v \right) + \mathcal{I}_d(y) \sum_{m=1}^{M} \pi_d^m f_m (v^{(m)} - v) \\
& + \xi \frac{|h|^2}{2} + \xi \mathcal{I}_g(\kappa) \sum_{j=1}^{J} \pi_g^j \left(1 - g_j + g_j \log g_j\right) + \xi \mathcal{I}_d(y) \sum \pi_d^m \left(1 - f_m + f_m \log f_m\right).
\end{aligned}
$$

FOC

$$
\begin{aligned}
g_j &= \exp\left(-\frac{1}{\xi_m} (v^{(j)} - v) \right), \\
f_m &= \exp\left(-\frac{1}{\xi_m} (v^{(m)} - v) \right), \\
h &= -\frac{1}{\xi} \left[(v_{\log K} - R v_R)(1-R)\sigma_d + (v_{\log K} + (1-R)v_R) R \sigma_g \right] \\
&\quad -\frac{1}{\xi} \varsigma \eta A_d (1-R) K \left(v_y - \{\gamma_1 + \gamma_2 Y\} \right) \\
&\quad -\frac{1}{\xi} \sigma_\kappa v_{\log \kappa}.
\end{aligned}
$$

Using the solutions for g_j, f_m, and h allows for an algebraic simplification of the form:

$$
\begin{aligned}
\delta v = & \max_{i_g, i_d, i_\kappa} \min_{g_j, f_m, h} \delta \log([A_d - i_d](1-R) + [A_g - i_g]R - i_\kappa) + \delta \log K \\
& + \left((1-R)\left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + R\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} \right) v_{\log K} \\
& + \frac{\sigma_d^2(1-R)^2 + \sigma_g^2 R^2}{2} v_{\log K, \log K} \\
& + \left(\left[\alpha_g + \Gamma_g \log(1+\theta_g i_g)\right] - R\sigma_g^2 - \left[\alpha_d + \Gamma_d \log(1+\theta_d i_d)\right] + (1-R)\sigma_d^2 \right) R(1-R) v_R \\
& + \frac{1}{2} R^2 (1-R)^2 (\sigma_g^2 + \sigma_d^2) v_{RR} \\
& + \left[-R(1-R)^2 \sigma_d^2 + R^2 (1-R)\sigma_g^2 \right] v_{\log K, R} \\
& + \beta_f \eta A_d (1-R) K v_y + \frac{|\varsigma|^2 (\eta A_d (1-R) K)^2}{2} v_{yy} \\
& - \left(\{\gamma_1 + \gamma_2 Y\}(\beta_f + \varsigma \cdot h)\eta A_d (1-R) K + \frac{1}{2}\gamma_2 |\varsigma|^2 (\eta A_d (1-R) K)^2 \right) \\
& + \left(-\zeta + \psi_0 i_\kappa^{\psi_1} \exp(\psi_1 (\log K - \log \kappa) D) - \frac{1}{2}|\sigma_\kappa|^2 \right) v_{\log \kappa} + \frac{|\sigma_\kappa|^2}{2} v_{\log \kappa, \log \kappa} \\
& - \frac{1}{2\xi} (v_{\log K} - R v_R)^2 (1-R)^2 \sigma_d^2 - \frac{1}{2\xi} (v_{\log K} + (1-R) v_R)^2 R^2 \sigma_g^2 \\
& - \frac{1}{2\xi} \varsigma^2 \eta^2 A_d^2 (1-R)^2 K^2 (v_y - \{\gamma_1 + \gamma_2 Y\})^2 - \frac{1}{2\xi} \sigma_\kappa^2 v_{\log \kappa}^2 \\
& + \xi \mathcal{I}_g(\kappa) \sum_{j=1}^{J} \pi_g^j (1 - g_j) + \xi \mathcal{I}_d(y) \sum \pi_d^m (1 - f_m) .
\end{aligned}
$$

Appendix B Abatement Model

Consider allowing for carbon abatement in the following way. First, emissions are given by

$$E_t = \lambda_t A_d K_d (1 - \iota_t),$$

where β_t is the emissions intensity of dirty output. Note that ι_t is the choice of abatement and can be solved for as a function of emissions:

$$\iota_t = 1 - \frac{E_t}{\lambda_t A_d K_d}.$$

The cost of abatement is given as a fraction of the dirty output, given by

$$J = [A_d K_d + A_g K_g] \phi_0 (\iota)^{\phi_1}.$$

Finally, market clearing gives us the output constraint

$$\begin{aligned} C &= A_d K_d + A_g K_g - I_d - I_g + I_\kappa - J \\ &= [A_d(1-\phi_0(1-\frac{E}{\lambda_t A_d K_d})^{\phi_1}) - i_d]K_d + [A_g(1-\phi_0(1-\frac{E}{\lambda_t A_d K_d})^{\phi_1}) - i_g]K_g - i_\kappa K. \end{aligned}$$

With that, we can construct the HJB equation for this setting, which is given as

$$\begin{aligned} \delta V &= \max_{i_g, i_d, i_\kappa, E} \delta \log([A_d K_d + A_g K_g](1-\phi_0(1-\frac{E}{\lambda_t A_d K_d})^{\phi_1}) - i_d K_d - i_g K_g - i_\kappa K) - \delta \log N_t \\ &\quad + \{\alpha_d + \Gamma_d \log(1+\theta_d i_d)\} V_d K_d + \{\alpha_g + \Gamma_g \log(1+\theta_g i_g)\} V_g K_g + \frac{|\sigma_d|^2 K_d^2}{2} V_{dd} + \frac{|\sigma_g|^2 K_g^2}{2} V_{gg} \\ &\quad + \beta_f E_t V_Y + \frac{|\varsigma|^2 (E_t)^2}{2} V_{YY} + [\{\gamma_1 + \gamma_2 Y_t\}\beta_f E_t + \frac{1}{2}\gamma_2 |\varsigma|^2 E_t^2] V_{\log N} + \frac{|\varsigma|^2 E_t^2}{2} V_{\log N \log N} \\ &\quad + (-\zeta + \psi_0 i_\kappa^{\psi_1} \exp(\psi_1(\log K - \log \kappa)) - \frac{1}{2}|\sigma_\kappa|^2) V_{\log \kappa} + \frac{|\sigma_\kappa|^2}{2} V_{\log \kappa, \log \kappa}. \end{aligned}$$

The FOC for investment and R&D are given by

$$\begin{aligned} 0 &= -\delta([A_d(1-\phi_0(1-\frac{E}{\lambda_t A_d K_d})^{\phi_1}) - i_d]K_d + [A_g(1-\phi_0(1-\frac{E}{\lambda_t A_d K_d})^{\phi_1}) - i_g]K_g - i_\kappa K)^{-1} \\ &\quad + \Gamma_d \theta_d (1+\theta_d i_d)^{-1} V_d \\ 0 &= -\delta([A_d(1-\phi_0(1-\frac{E}{\lambda_t A_d K_d})^{\phi_1}) - i_d]K_d + [A_g(1-\phi_0(1-\frac{E}{\lambda_t A_d K_d})^{\phi_1}) - i_g]K_g - i_\kappa K)^{-1} \\ &\quad + \Gamma_g \theta_g (1+\theta_g i_g)^{-1} V_g \\ 0 &= -\delta([A_d(1-\phi_0(1-\frac{E}{\lambda_t A_d K_d})^{\phi_1}) - i_d]K_d + [A_g(1-\phi_0(1-\frac{E}{\lambda_t A_d K_d})^{\phi_1}) - i_g]K_g - i_\kappa K)^{-1} K \\ &\quad + \psi_0 \psi_1 i_\kappa^{\psi_1 - 1} \exp(\psi_1(\log K - \log \kappa)) V_{\log \kappa} \\ 0 &= \delta([A_d(1-\phi_0(1-\frac{E}{\lambda_t A_d K_d})^{\phi_1}) - i_d]K_d + [A_g(1-\phi_0(1-\frac{E}{\lambda_t A_d K_d})^{\phi_1}) - i_g]K_g - i_\kappa K)^{-1} \\ &\quad \times \frac{\phi_0 \phi_1 (A_d K_d + A_g K_g)}{\lambda_t A_d K_d} (1-\frac{E}{\lambda_t A_d K_d})^{\phi_1 - 1} \\ &\quad + \beta_f V_Y + |\varsigma|^2 E_t V_{YY} + [\{\gamma_1 + \gamma_2 Y_t\}\beta_f + \gamma_2 |\varsigma|^2 E_t] V_{\log N} + \{\gamma_1 + \gamma_2 Y_t\}|\varsigma|^2 E_t V_{\log N \log N} \end{aligned}$$

Now we redefine the states to get it all correct. We use $\log K = \log(K_d + K_g)$ and $R = \frac{K_g}{K_d + K_g}$.

Define dW_g as the green capital shock and dW_d as the dirty capital shock. Having assumed the shocks are independent, then our cross-partial terms should drop out and we end up with the

following:

$$\begin{aligned}
d\log K &= (1-R)[\alpha_d + i_d - \frac{\phi_d}{2}i_d^2]dt + R[\alpha_g + i_g - \frac{\phi_g}{2}i_g^2]dt \\
&\quad -\frac{1}{2}|\sigma_d(1-R)|^2 - \frac{1}{2}|\sigma_g R|^2 dt + (1-R)\sigma_d dW_d + R\sigma_g dW_g \\
dR &= -[\alpha_d + i_d - \frac{\phi_d}{2}i_d^2]R(1-R)dt + R(1-R)[\alpha_g + i_g - \frac{\phi_g}{2}i_g^2]dt \\
&\quad +R(1-R)^2|\sigma_D|^2 dt - R^2(1-R)|\sigma_G|^2 dt - R(1-R)\sigma_d dW_d + R(1-R)\sigma_g dW_g
\end{aligned}$$

The new HJB equation is given by $V(K_d, K_g, Y, \log \mathcal{I}_g, \log N) = v(K, R, Y_t, \log \mathcal{I}_g) - \log N_t$:

$$\begin{aligned}
\delta v &= \max_{i_g, i_d, i_\kappa, E} \delta \log([A_d(1-R) + A_g R](1 - \phi_0(1 - \frac{E}{\lambda_t A_d K(1-R)})^{\phi_1}) - i_d(1-R) - i_g R - i_\kappa) + \delta \log K \\
&\quad +[\{\alpha_d + \Gamma_d \log(1+\theta_d i_d)\}(1-R) + \{\alpha_g + \Gamma_g \log(1+\theta_g i_g)\}R - (1-R)^2\frac{|\sigma_d|^2}{2} - R^2\frac{|\sigma_g|^2}{2}]v_{\log K} \\
&\quad +[\{\alpha_g + \Gamma_g \log(1+\theta_g i_g)\}R(1-R) - \{\alpha_d + \Gamma_d \log(1+\theta_d i_d)\}R(1-R) \\
&\quad +|\sigma_d|^2 R(1-R)^2 - |\sigma_g|^2 R^2(1-R)]v_R \\
&\quad +\{\frac{|\sigma_d|^2}{2} + \frac{|\sigma_g|^2}{2}\}R^2(1-R)^2 v_{RR} + \{\frac{|\sigma_g|^2}{2}R^2 + \frac{|\sigma_d|^2}{2}(1-R)^2\}v_{\log K, \log K} \\
&\quad -|\sigma_d|^2 v_{\log K, R} R(1-R)^2 + |\sigma_g|^2 v_{\log K, R} R^2(1-R) \\
&\quad +\beta_f E_t v_Y + \frac{|\varsigma|^2 (E_t)^2}{2} v_{YY} - [\{\gamma_1 + \gamma_2 Y_t\}\beta_f E_t + \frac{1}{2}\gamma_2 |\varsigma|^2 E_t^2] \\
&\quad +(-\zeta + \psi_0 i_\kappa^{\psi_1} \exp(\psi_1(\log K - \log \kappa)) - \frac{1}{2}|\sigma_\kappa|^2)v_{\log \kappa} + \frac{|\sigma_\kappa|^2}{2}v_{\log \kappa, \log \kappa}
\end{aligned}$$

The FOC for investment and R&D are given by

$$\begin{aligned}
0 &= -\delta([A_d(1-R) + A_g R](1 - \phi_0(1 - \frac{E}{\lambda_t A_d K(1-R)})^{\phi_1}) - i_d(1-R) - i_g R - i_\kappa)^{-1} \\
&\quad +\Gamma_d \theta_d (1+\theta_d i_d)^{-1}[v_k - (1-R)v_R] \\
0 &= -\delta([A_d(1-R) + A_g R](1 - \phi_0(1 - \frac{E}{\lambda_t A_d K(1-R)})^{\phi_1}) - i_d(1-R) - i_g R - i_\kappa)^{-1} \\
&\quad +\Gamma_g \theta_g (1+\theta_g i_g)^{-1}[v_k + (1-R)v_R] \\
0 &= -\delta([A_d(1-R) + A_g R](1 - \phi_0(1 - \frac{E}{\lambda_t A_d K(1-R)})^{\phi_1}) - i_d(1-R) - i_g R - i_\kappa)^{-1} \\
&\quad +\psi_0 \psi_1 i_{\mathcal{I}}^{\psi_1 - 1} \exp(\psi_1(\log K - \log \kappa))v_{\log \mathcal{I}_g} \\
0 &= \delta([A_d(1-R) + A_g R](1 - \phi_0(1 - \frac{E}{\lambda_t A_d K(1-R)})^{\phi_1}) - i_d(1-R) - i_g R - i_\kappa)^{-1} \\
&\quad \times \frac{\phi_0 \phi_1 (A_d(1-R) + A_g R)}{\lambda_t A_d (1-R)}(1 - \frac{E}{\lambda_t A_d K(1-R)})^{\phi_1 - 1} \\
&\quad +\beta_f(v_Y - \{\gamma_1 + \gamma_2 Y_t\}) + |\varsigma|^2 E_t(v_{YY} - \gamma_2)
\end{aligned}$$

So far, the model presented here abstracts from jumps and uncertainty. We could introduce jumps as before with different realizations of γ_3^m and A_g^j possible, a jump, or jumps, in the value of λ_t, and jump and diffusion misspecification concerns across the various different channels as outlined in the main text of the paper.

Appendix C Social Valuations

An important component of our analysis can be captured by social valuations, which are shadow prices for the marginal benefit or marginal cost of an additional unit of flows or stocks the various stocks in our model. In particular, the Social Cost of Carbon (SCC), Social Value of R&D (SVR), Social Value of Green Capital (SVG), and the Social Value of Dirty Capital (SVD), can each be computed from the FOC of the social planner's problem

$$\begin{aligned}
SCC_t &\propto 1000\eta\{-\beta_f v_Y - |\varsigma|^2 E_t v_{YY} + \{\gamma_1 + \gamma_2 Y_t\}\beta_f + \gamma_2|\varsigma|^2 E_t\} \\
SVR_t &\propto \psi_0\psi_1\left(\frac{I_t^j}{\kappa_t}\right)^{\psi_1 - 1} v_{\log\kappa} \\
SVG_t &\propto \Gamma_g\theta_g(1+\theta_g i_g)^{-1}[v_{\log K} + (1-R)v_R] \\
SVD_t &\propto \Gamma_d\theta_d(1+\theta_d i_d)^{-1}[v_{\log K} - Rv_R]
\end{aligned}$$

where the proportionality scaling for each social valuation term is the inverse of the marginal utility of consumption:

$$(MUC)^{-1} = \delta^{-1}([A_d - i_d](1-R) + [A_g - i_g]R - i_{\mathcal{I}}).$$

Following Barnett et al. (2020) and Barnett et al. (2021), we can decompose the contribution to these social valuations coming from uncertainty of different forms and from different sources. The decomposition requires solving Feynman-Kac equations that represent the expected discounted value of marginal contributions to each stock variable, where the expectation is varied across the different distributions of the potential models under consideration in our setting, e.g., the baseline prior distribution or various forms of the distorted, uncertainty-adjusted distribution of models. We plan to explore these important valuations in future work on this project.

Appendix D Alternative Model Parameterization

Economic Framework Parameters (Log Adjustment Costs - BBH RFS):

Parameters	values
δ	0.01
$(\alpha, \Gamma, \theta, \sigma)$	(-0.035, 0.060, 16.7, {0.01, 0.016, 0.02})
$(\alpha', \Gamma', \theta', \sigma')$	(-0.038, 0.0633, 15.7895, {0.01, 0.016, 0.02})
$(\zeta, \psi_0, \psi_1, \sigma_\kappa)$	(0, 0.10583, 0.5, 0.0078)
ϱ	1120
$(A_d, A_g; \{A_g^j\})$	$(0.12, 0.10; \{0.15, 0.20, 0.30\})$
(K_d, K_g, Y, κ)	$0.5 \times (85/0.11), 0.5 \times (85/0.11), 1.1, 11.2)$

- δ is consistent with the subjective discount rate used in macro-asset pricing literature.
- (Γ_d, θ_d) and (Γ_g, θ_g) are chosen so that the no-climate, one capital version of the model satisfies three conditions as in Barnett et al. (2020):

$$\begin{aligned} 1 &= \Gamma\theta = 1 \\ 0.02 &= \alpha + \Gamma \log(1+\theta i) = \mathbb{E}[dK/K] \\ 2.5 &= \frac{A-i}{\delta} v_{\log K} = \frac{1+\theta i}{\Gamma\theta} = q \end{aligned}$$

- σ_d and σ_g are chosen to match the values used in the World Bank database.
- A_d, A_g, A_g^j generate output consistent with World Bank World GDP values, and allow for a meaningful technological change upon realization of the technology jump.
- $(\zeta, \psi_0, \psi_1, \varrho)$ values are based on BLS Total US R&D stocks values (scaled to World values), Lucking et al. (2019) and Bloom et al. (2019) estimates for the returns to R&D investment, and consistent with an expected arrival time of potential carbon neutral technology innovation occurring between 30 and 80 years based on current knowledge levels.
- σ_κ is chosen to match the volatility of the other capital stocks.

Appendix E Neural Nets Implementations

Below we give the pseudo-code for the deep Galerkin method - policy improvement algorithms (DGM-PIA) described in Section 5.1, as well as implementation details, including network architectures and training hyperparameters.

For the DGM-PIA described in Section 5.1, we use feedforward neural networks with 4 hidden layers of width 32 and `tanh` activation function (except for the output layer) to approximate both the unknown value functions and the optimal controls. At the output layer, a customized hyperbolic tangent function is used for i_g, i_d, and the `sigmoid` function is used for i_κ.

To train the neural nets, we run 250000 epochs with a batch size of 32. 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 is done on Google Colab with a total runtime of 2.29 hours.

For the Han and E (2016) method solution, we use feedforward neural networks with 4 hidden layers of size 32 and `tanh` activation function to approximate the optimal controls. To train the neural nets, we run N = 100,000 epochs with a batch size of 2^7. The infinite time horizon is approximated by $[0, 1000]$, and the time step size is 0.1.

Algorithm 1 The DGM-PIA algorithm for solving the generic HJB (5.1)

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 lr_V, $\text{lr}_{\boldsymbol{\alpha}}$

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

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

3: Compute the value function loss in (5.1) 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 θ_{n+1}:

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

5: Compute the control loss in (5.1) 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 φ_{n+1}:

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

7: **end for**