Strategic interactions and uncertainty in decisions to curb greenhouse gas emissions

Margaret Insley* Tracy Snoddon† Peter A. Forsyth‡

September 2020

Abstract

This paper examines the strategic interactions of two large regions making choices about greenhouse gas emissions in the face of rising global temperatures. Three central features are highlighted: uncertainty, the incentive for free riding, and asymmetric characteristics of decision makers. Optimal decisions are modelled in a fully dynamic, feedback Stackelberg pollution game. Global average temperature is modelled as a mean reverting stochastic process. A numerical solution of a coupled system of Hamilton-Jacobi-Bellman equations is implemented and the probability distribution of outcomes is illustrated with Monte Carlo simulation. When players are identical, the outcome of the game is much worse than the social planner's outcome. An increase in temperature volatility reduces player utility, making cooperative action through a social planner more urgent. Asymmetric damages or asymmetric preferences for emissions reductions are shown to have important effects on the strategic interactions of players.

Keywords: climate change, dynamic game, feedback Stackelberg equilibrium, feedback Nash equilibrium, uncertainty, asymmetric players, HJB equation
JEL codes: C63, C73, Q52, Q54

*Department of Economics, University of Waterloo, Waterloo, Ontario, Canada. margaret.insley@uwaterloo.ca

†Department of Economics, Wilfrid Laurier University, Waterloo, Ontario, Canada. tsnoddon@wlu.ca

‡Cheriton School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada. paforsyt@uwaterloo.ca

1 Introduction

Climate change caused by human activity represents a particularly intractable tragedy of the commons, which calls for cooperative actions of individual decision makers at both national and regional levels. The likely success of cooperative actions is hampered by the large incentives for free riding by decision makers who may delay making deep cuts in carbon emissions in hopes that others will do the "heavy lifting". Further complicating the problem are the enormous uncertainties inherent in predicting climate responses to the buildup in atmospheric carbon stocks and resulting impacts on human welfare, including the prospects for adaptation and mitigation. These large uncertainties and the need for cooperative global action have been used by some as justification for delaying aggressive unilateral policy actions. Nevertheless, many nations and sub-national jurisdictions have acted on their own to adopt policies to reduce carbon emissions even without national agreements or legislation in place. As a prominent example, since the Trump administration has reneged on the Paris Climate Accord, several states have vowed to go it alone and continue with aggressive climate policies. Other examples of jurisdictions taking unilateral carbon pricing initiatives are given in Kossey et al. (2015).

The observation that national or regional governments implement environmental regulations sooner or more aggressively than required by international agreements or national legislation has been studied by various researchers.[1] Local circumstances, including voter preferences, local damages from emissions, and strategic considerations regarding the actions of other jurisdictions, may play a role. A nation or region may be motivated to act ahead of others if it experiences relatively more severe local damages from emissions. Differences in environmental preferences may prompt some jurisdictions to take early action (Bednar-Friedl 2012). California and British Columbia (B.C.) (a province in Canada), both early adopters of carbon pricing, appear to have residents who are more environmentally aware, implying these governments acted in accordance with the preferences of a large segment of their voters. A survey of stakeholders involved in the introduction of the B.C. carbon tax

[1] Urpelainen (2009) and Williams (2012) examine the puzzle at a sub-regional level.

concluded that a number of factors were at work. These factors include: (i) a high priority given to environmental stewardship by B.C. residents and (ii) the fact that several other regional jurisdictions appeared to be poised in 2008 to take climate change more seriously (Clean Energy Canada 2015). Governments may choose environmental policies strategically to gain a competitive advantage or to shift emissions to other regions (Barcena-Ruiz 2006).

This paper examines the strategic interactions of decision makers responding to climate change, focusing on three central features of the problem: uncertainty, the incentive for free riding, and asymmetric characteristics of decision makers. We develop a dynamic model of a Stackelberg game involving two regions and solve for a feedback equilibrium. Each region is a large emitter of greenhouse gases and benefits from their own emissions, but faces costs from the impact on global temperature of the cumulative emissions of both players. The modelling of the linkage between carbon emissions and global temperature is based on the assumptions of the well-known DICE model[2] (Nordhaus & Sztorc 2013). To capture uncertainty, average global temperature is modelled as a stochastic process. We solve the stochastic dynamic game using numerical techniques. We model temperature and carbon stock as evolving continuously in time as given by the solution of stochastic differential equations. Rather than assume continuously applied controls, we restrict the set of admissible controls to allow decisions only at fixed time intervals, which we view as a more realistic depiction of real world policy making. We allow for differing damages of climate change for each region as well as differing preferences for reducing greenhouse gas emissions. We explore the impact of these features on the optimal choice of emissions for each player and contrast with the choices made by a social planner. While our focus is on the outcome of a Stackelberg game, at each point in the state space, we can check if a feedback Nash equilibrium is possible, and if the feedback Stackelberg solution also represents a Nash equilibrium.

There is a significant prior literature which examines the tragedy of the commons caused by polluting emissions in a differential game setting. The relevant differential game literature is reviewed in Section 2, but we note here two papers most closely related to our paper in their focus on asymmetry of players' utilities. Both employ economic models in a deterministic

[2] Dynamic Integrated Model of Climate and the Economy

setting. Zagonari (1998) analyzes cooperative and non-cooperative games when the two players (countries) differ in the utility derived from a consumption good, the disutility caused by the pollution stock, and their concern for future generations as reflected in their discount rate. Interestingly, Zagonari finds equilibria for which the steady state pollution stock is lower than in the cooperative game. In particular, this result holds if the country with stronger environmental preferences (the "eco-country") has sufficiently large disutility from pollution and either a relatively strong concern for future generations or relatively small utility from consumption goods.

Wirl (2011) also examines whether differences in environmental sentiments can mitigate the tragedy of the commons associated with a problem such as global warming. The author characterizes a multi-player game with green and brown players. Green players are distinguished from brown players by a penalty term in their objective function which depends on the extent to which their emissions exceed the social optimum. In the examples chosen, the effect of green players on total emissions is modest, as their actions increase the free riding of brown players. Wirl notes the possibility of a type of green paradox in which the increasing numbers of green players causes increased emissions, because brown players increase their emissions and more than offset the impact of green players' decisions.

We also note Insley & Forsyth (2019) which explores alternate forms of games between symmetric players, including a leader-leader game as well as an interleaved game in which there is a significant delay between player decisions.

Our paper contributes to this literature is several ways. We develop a more general model which includes uncertainty and feedback strategies in a dynamic setting. The numerical results highlight the important influence of uncertainty in future temperature on optimal emissions choices and the evolution of the carbon stock. We study the effect of asymmetry in damages and environmental preferences on emissions choices, utility, and the implication for the evolution of global average temperature, contrasting the non-cooperative outcome with the outcome assuming a central planner empowered to make choices. Of interest is whether player asymmetries exacerbate free riding and the tragedy of the commons in a stochastic dynamic setting. Finally, we make a contribution in terms of the numerical methodology for solving a dynamic Stackelberg game under uncertainty with feedback strategies and path dependent variables. We describe the method used to determine the optimal Stackelberg solution (which always exists) and then show how to determine if a feedback Nash equilibrium exists.[3] Our numerical solution procedure involves use of a finite difference discretization of the system of Hamilton-Jacobi-Bellman (HJB) equations. In contrast to much of the previous literature, the choice of damage function can be any arbitrary function of state variables. In addition to providing the numerical solution of the HJB equations, which indicates optimal controls and expected utility at time zero for any chosen values of the state variables, we also undertake Monte Carlo simulation which allows us to depict the probability distribution of emissions, temperature and utility over the time frame of the analysis (150 years).

To preview our results, we highlight the crucial role of the damage function which specifies the harm from rising temperature, as has been noted by others (Weitzman 2012, Pindyck 2013). Very little reduction in carbon emissions occurs in the Stackelberg game or with the central planner using a conventional quadratic damage function. Exponentially increasing damages better reflect the catastrophic nature of damages anticipated if average global temperature should increase beyond 3°C above pre-industrial levels. We also find that temperature uncertainty plays a key role. With a larger temperature volatility, optimal emissions are reduced for the players in the game as well as for the social planner. The social planner's response is relatively large compared to the game for key values of the state variables (carbon stock and temperature), implying the benefit of cooperative action through a social planner increases at higher volatility. Monte Carlo analysis demonstrates the much higher risk of the game, relative to the social planner. Asymmetric costs are also found to have an important effect on strategic interactions of players. Higher damage experienced by one player may cause the other player to increase or decrease emissions relative to the symmetric

[3] It is well known that for differential games with feedback strategies, only special classes of models result in well-posed mathematical problems for which it is possible to characterize Nash equilibria. See Bressan (2011). These include linear-quadratic games where the feedback controls depend linearly on the state variable, as well as certain forms of stochastic differential games where the state evolves according to an Ito process.

case depending on the values of the state variables. As with increased volatility, we highlight the greater advantage provided by a social planner in this case. Finally, we observe that an increase in green preferences by one player has an impact on the optimal actions of the other player, but again, the direction of this effect varies depending on current values of state variables - and in particular the stock of atmospheric carbon. We identify both a green paradox and a green bandwagon effect.

The remainder of the paper proceeds as follows. In Section 2, we provide a more detailed literature review. The formulation of the climate change decision problem is described in Section 3. Section 4 provides a detailed description of the dynamic programming solution. Section 5 describes the detailed modelling assumptions and parameter values. Numerical results are described in Section 6, while Section 7 provides concluding comments.

2 Literature

This paper contributes to the literature on differential games dealing with trans-boundary pollution problems as well as to the developing literature on accounting for uncertainty in optimal policies to address climate change.

Economic models of climate change have long been criticized for arbitrary assumptions regarding functional forms and key parameter values as well as unsatisfactory treatment of key uncertainties including the possibility of catastrophic events.[4] Of course, this is not surprising given the intractable nature of the climate change problem. Policies to address climate change have been extensively studied using the DICE model, a deterministic model developed in the 1990s, which has been revised and updated several times since then (Nordhaus 2013). Initially uncertainty was addressed through sensitivities or Monte Carlo analysis, but there has since been a significant research effort to address uncertainty using more robust methodologies. We mention only a sample of that literature. Kelly & Kolstad (1999) and Leach (2007) embed a model of learning into the DICE model to examine active learning by a social planner about key climate change parameters. More recent papers which incor-

[4] See Pindyck (2013) for a harsh critique.

porate stochastic components into one or more state variables in the DICE model include Crost & Traeger (2014), Ackerman et al. (2013) and Traeger (2014). Lemoine & Traeger (2014) extend the work of Traeger (2014) by incorporating the possibility of sudden shifts in system dynamics once parameters cross certain thresholds. Policy makers learn about the thresholds by observing the evolution of the climate system over time. Hambel et al. (2017) present a stochastic equilibrium model for optimal carbon emissions with key state variables, including carbon concentration, temperature and GDP, modelled as stochastic differential equations. Chesney et al. (2017) examine optimal climate polices using a model in which global temperature is stochastic and assuming there is a known temperature threshold which will result in disastrous consequences if it is exceeded for a sustained period of time.

Differential game models have been used extensively to examine strategic interactions between players who benefit individually from polluting emissions but are also harmed by the cumulative emissions of all players. Key assumptions, such as the information known to each player, determine whether the game can be described by a closed form mathematical solution.[5] For example, open loop strategies, which depend solely on time, result when players know only the initial state of the system. Nash and Stackelberg equilibria for open loop strategies are well understood. In contrast, when players can directly observe the state of the system at every instant in time, feedback strategies (also called closed-loop or Markovian strategies) which depend on the state of the system may be employed. The resulting value functions satisfy a system of highly non-linear HJB partial differential equations (PDEs). From the theory of partial differential equations it is known that if the system is non-stochastic, it should be hyperbolic in order for it to be well posed, in that it admits a unique solution depending continuously on the initial data (Bressan & Shen 2004). Our system of HJB equations is degenerate parabolic, which further complicates matters.

In games with feedback strategies only special classes of models are known to result in well-posed mathematical problems. These include zero-sum games, as well as linear-quadratic games. Linear-quadratic games have been used extensively in the economics liter-

[5]See Bressan (2011) for a discussion of the challenges of finding appropriate mathematical models which result in closed form solutions.

ature to study pollution games, and some relevant papers, which admit closed form solutions, are detailed below. In this class of games, utility is a quadratic function of the state variable, while the state variable is linear in the control. Robust game models are also found with Nash feedback equilibria for stochastic differential games where the state evolves according to an Ito process such as

$$dx = f(t, x, u_1, u_2)dt + \sigma(t, x)d\mathcal{Z} \tag{1}$$

where x represents the state variable, t is time, u_1 and u_2 represent the controls of players 1 and 2, f and σ are known functions, and $d\mathcal{Z}$ is the increment of a Wiener process. As noted by Bressan (2011), for this case the value functions can be found by solving a Cauchy problem for a system of parabolic equations. The Cauchy problem is well posed if the diffusion tensor σ has full rank. In our case, the diffusion tensor is not of full rank (i.e. the system of partial differential equations is degenerate), hence we cannot expect that a Nash equilibrium will always exist. Additional discussion of the complexities of solving problems involving differential games can be found in Salo & Tahvonen (2001), Ludkovski & Sircar (2015), Harris et al. (2010), Cacace et al. (2013), Amarala (2015), and Ledvina & Sircar (2011).

Long (2010), Long (2011), Dockner et al. (2000) and Jorgensen et al. (2010) provide surveys of the sizable literature addressing strategic interactions in the optimal control of pollution or natural resource exploitation using games, much of it in a deterministic setting. This literature focuses on the questions: (i) are players are better off with cooperative behaviour and (ii) how do the steady state levels of pollution compare under cooperative versus non-cooperative games.

Examples of dynamic differential pollution games in a non-stochastic setting include Dockner & Long (1993), Zagonari (1998), Benchekroun & van Long (1998), List & Mason (2001), Wirl (2011), and Benchekroun & Chaudhuri (2014). Under certain conditions, closed-form solutions are found for linear and non-linear Nash strategies. In a recent paper, Colombo & Labrecciosa (2019) contrast Stackelberg and Cournot equilibria in a deterministic setting and derive a "feedback-generalized-Stackelberg-Nash-Cournot equilibrium" for the exploitation of a common pool renewable resource. A few papers derive analytical solutions to differential pollution models in stochastic settings. These include Xepapadeas (1998), Wirl (2008), and Nkuiya (2015).

There is a developing literature on the numerical solution of dynamic games in the context of non-renewable resource markets. Some earlier papers developed models where two or more players extract from a common stock of resource. Examples include van der Ploeg (1987) and Dockner et al. (1996). Salo & Tahvonen (2001) were among the first to explore oligopolistic natural resource markets in a differential Cournot game using feedback strategies. Prior to that, the focus had been on open-loop strategies, because of their tractability. Harris et al. (2010), Ludkovski & Sircar (2012), and Ludkovski & Yang (2015) study the extraction of an exhaustible resource as an N-player continuous time Cournot game when players have heterogeneous costs.

3 Problem Formulation

This section provides an overview of the climate change decision model. Details of functional forms and parameter values are provided in Section 5. A summary of variable names is given in Table 1. We model the optimal timing and stringency of environmental regulations (in terms of the reduction of greenhouse gas emissions) as a stochastic optimal control problem. Our two main cases are for a Stackelberg game and a social planner. In Appendix B we describe the controls for a Nash equilibrium, which is used to contrast with the Stackelberg game. The players in the Stackelberg game are two regions, each contributing to the atmospheric stock of greenhouse gases - which, for simplicity, we will refer to as the carbon stock. These regions may be thought of as single nations or groups of nations acting together, but each is a major contributor to the global carbon stock. Each region seeks to maximize discounted expected utility by making emission choices taking into account the optimal actions of the other region. The social planner chooses emission levels in each region so as to maximize the expected sum of utilities from both regions.

Table 1: List of Model Variables

Variable	Description
$E_p(t)$	Emissions in region p
$\bar{E}_p$	benchmark emissions for player p
e_1, e_2	Particular realizations of state variable $E_p(t)$
ω, ω_2	any possible control choice by players 1 and 2
e_1^+, e_2^+	particular controls chosen by players 1 and 2
$S(t)$	Stock of pollution at time t, a state variable
s	A realization of $S(t)$
$\bar{S}$	preindustrial level of carbon
$\rho(X, S, t)$	Rate of natural removal of the pollution stock
σ	temperature volatility
$\eta(t)$	speed of mean reversion in temperature equation
$X(t)$	Average global temperature, a state variable
x	A realization of $X(t)$
$\bar{X}$	long run equilibrium level of temperature, °C above pre-industrial levels
$B_p(E_p, t)$	Benefits from pollution
$C_p(X, t)$	Damages from pollution
$g_p(t)$	Emissions reduction in region p relative to a target
θ_p	Willingness to pay in region p for emissions reduction from a target
$A_p(g_p(t))$	Green reward benefits from emissions reductions
π_p	Flow of net benefits to region p
r	risk free interest rate

Regions emit carbon in order to generate income. For simplicity we assume that there is a one to one relation between emissions and regional income. The two regions are indexed by $p = 1, 2$ and E_p refers to carbon emissions from region p. The stock of atmospheric carbon, S, is augmented by the emissions of each player and is reduced by a natural cycle whereby carbon is removed from the atmosphere and absorbed into other carbon sinks. The removal of carbon from the atmosphere can be described by decay function, $\rho(X, S, t)$, which in theory may depend on the average surface temperature, $X(t)$, the stock of carbon, $S(t)$, and time, t. $\rho(X, S, t)$ is referred to as the removal rate. For simplicity, as described in

Section 5, we will later drop the dependence on X and S, assuming that ρ is a function only of time. However, our solution technique can easily accommodate more general functional forms for ρ. The evolution of the carbon stock over time is described by the deterministic differential equation:

$$\frac{dS(t)}{dt} = E_1 + E_2 + (\bar{S} - S(t))\rho(X, S, t);\; S(0) = s_0 \quad S \in [s_{min},\; s_{max}]\,. \tag{2}$$

$\bar{S}$ is the pre-industrial equilibrium level of atmospheric carbon.

The mean global increase in temperature above the pre-industrial level, denoted by X, is described by an Ornstein Uhlenbeck process:

$$dX(t) = \eta(t)\left[\bar{X}(S,t) - X(t)\right]dt + \sigma d\mathcal{Z}. \tag{3}$$

where $\eta(t)$ represents the speed of mean reversion and is a deterministic function of time. $\bar{X}$ represents the long run mean of global average temperature which depends on the stock of carbon and time. σ is the volatility parameter, assumed to be constant. The detailed specification of these functions and parameters is given in Section 5. $d\mathcal{Z}$ is the increment of a standard Wiener process, intended to capture the volatility in the earth's temperature due to random effects.

The net benefits from carbon emissions are represented as a general function $\pi_p = \pi_p(E_1, E_2, X, S, t)$. More specifically, π is composed of the benefits from emissions, $B_p(E_p, t)$, the damages from increasing temperature, $C_p(X, t)$, and a green reward that results from reducing emissions relative to a given target or baseline level, $A_p(g_p(t))$:

$$\pi_p = B_p(E_p, t) - C_p(X, t) + A_p(g_p(t)) \quad p = 1, 2; \tag{4}$$

where $g_p(t)$ refers to emissions reduction. The detailed specification of benefits, damages, and the green reward is left to Section 5

It is assumed that the control is applied at fixed decision times denoted by:

$$\mathcal{T} = \{t_0 = 0 < t_1 < ...t_m... < t_M = T\}. \tag{5}$$

We assume that $(t_m - t_{m-1})$ is constant (two years in our numerical example), reflecting the time lags in real world policy making. A sensitivity with one year intervals made little difference to our results.[6] We use the following short hand notation. Consider a function $f(t)$. We define

$$f(t^+) = \lim_{\epsilon \to 0^+} f(t + \epsilon) \quad ; \quad f(t^-) = \lim_{\epsilon \to 0^+} f(t - \epsilon). \tag{6}$$

Informally t^- and t^+ denote the instants immediately before and after t.

Let $e_1^+(E_1, E_2, X, S, t_m^+)$ and $e_2^+(E_1, E_2, X, S, t_m^+)$ denote the controls implemented by the players 1 and 2 respectively, which are contained within the set of admissible controls: $e_1^+ \in Z_1$ and $e_2^+ \in Z_2$. The controls act on the state variables, E_1 and E_2, either leaving them as is or changing to a new level. We can specify a control set which contains the optimal controls for all t_m.

$$K = \left\{ (e_1^+, e_2^+)_{t_0=0},\ (e_1^+, e_2^+)_{t_1=1},\ \dots, (e_1^+, e_2^+)_{t_M=T} \right\}. \tag{7}$$

In this paper we will consider three possibilities for selection of the controls (e_1^+, e_2^+) at $t \in \mathcal{T}$: Stackelberg, Nash, and social planner. We delay the precise specification of how the the Stackelberg and social planner controls are determined until Section 4.2, while the Nash controls are specified in Appendix B.

Regardless of the control strategy, the value function for player p, $V_p(e_1, e_2, x, s, t)$ is

[6] It is possible to let this time interval become vanishingly small, in which case this would become a classic impulse control problem. This would increase the computational cost of the numerical examples and is beyond the scope of the paper. The interval between decision times is currently exogenous. By making $(t_m - t_{m-1})$ very small we could examine the impact of endogenous decision times. In this case, it would make sense to add a cost for changing emissions to reflect administrative costs of applying a new policy. This would result in finite times between actual decision times, since the cost of continuous policy changes would be prohibitive.

defined as:

$$V_p(e_1, e_2, x, s, t) = \mathbb{E}_K \left[\int_{t'=t}^{T} e^{-r(t'-t)} \pi_p(E_1(t'), E_2(t'), X(t'), S(t'))\ dt' + \right.$$
$$\left. e^{-r(T-t)} V_p(E_1(T), E_2(T), X(T), S(T), T) \middle| E_1(t) = e_1, E_2(t) = e_2, X(t) = x,\ S(t) = s \right]. \quad (8)$$

$\mathbb{E}_K[\cdot]$ is the expectation under control set K. Note that lower case letters e_1, e_2, x, s have been used to denote realizations of the state variables E_1, E_2, X, S. The value in the final time period, T, is assumed to be the present value of a perpetual stream of expected net benefits at given carbon stock, S, and temperature levels, X, with emissions set to their maximum level. This is reflected in the term $V_p(E_1(T), E_2(T), X(T), S(T), T)$ and is described in Section 4.1 as a boundary condition. The justification is the assumption that the world has decarbonized by this time, and emissions still generate income but no longer add to the stock of carbon.

4 Dynamic Programming Solution

Using dynamic programming, we solve the problem represented by Equation (8) backwards in time, breaking the solution phases up into two components for $t \in (t_m^-, t_m^+)$ and (t_m^+, t_{m+1}^-), where $t_m \in \mathcal{T}$ are decision times (Equation (5)) and t_m^+ and t_m^- are defined in Equation (6) . In the interval (t_m^-, t_m^+), we determine the optimal controls, implying that for the Stackelberg game, the follower plays immediately after the leader. In the interval (t_m^+, t_{m+1}^-), we solve a system of partial differential equations. Recall it is assumed that $(t_{m+1} - t_m)$ is a fixed finite interval. As a visual aid, Equation (9) shows the noted time intervals going forward in time,

$$t_m^- \ \rightarrow\ t_m^+ \ \rightarrow\ t_{m+1}^- \ \rightarrow\ t_{m+1}^+ \ . \quad (9)$$

4.1 Advancing the solution backward in time from $t_{m+1}^- \rightarrow t_m^+$

The solution proceeds going backward in time from $t_{m+1}^- \rightarrow t_m^+$, which is a fixed finite interval where players take no actions, but temperature and carbon stock evolve. Consider at time interval $h < (t_{m+1} - t_m)$. For $t \in (t_m^+, t_{m+1}^- - h)$, the dynamic programming principle states that (for small h),

$$\begin{aligned} V(e_1, e_2, s, x, t) &= e^{-rh}\mathbb{E}\Big[V(E_1(t), E_2(t), S(t+h), X(t+h), t+h)\Big| \\ &\qquad S(t) = s, X(t) = x, E_1(t) = e_1, E_2(t) = e_2\Big] + \pi_p(e_1, e_2, s, x, t)h. \end{aligned} \tag{10}$$

The parameter r is the risk free interest rate. Note that for $t \in (t_m^+, t_m^-)$, the emission levels E_1 and E_2 are fixed. Letting $h \to 0$ and using Ito's Lemma,[7] the equation satisfied by the value function, V_p is expressed as:

$$\frac{\partial V_p}{\partial t} + \pi_p(e_1, e_2, x, s, t) + \mathcal{L}V_p = 0, \quad p = 1, 2\ . \tag{11}$$

where $\mathcal{L}$ is the differential operator for player p and is defined as follows:

$$\mathcal{L}V_p \equiv \frac{(\sigma)^2}{2}\frac{\partial^2 V_p}{\partial x^2} + \eta(\bar{X} - x)\frac{\partial V_p}{\partial x} + [(e_1 + e_2) + \rho(\bar{S} - s)]\frac{\partial V_p}{\partial s} - rV_p; \quad p = 1, 2\ . \tag{12}$$

The arguments in the V_p function, as well as in η and ρ, have been suppressed when there is no ambiguity.

The domain of Equation (11) is $(e_1, e_2, x, s, t) \in \Omega^\infty$, where $\Omega^\infty \equiv Z_1 \times Z_2 \times [x^0, \infty] \times [\bar{S}, \infty] \times [0, \infty]$. x^0 would be the lowest temperature possible on earth. For computational purposes, we truncate the domain Ω^∞ to Ω, where $\Omega \equiv Z_1 \times Z_2 \times [x_{min}, x_{\max}] \times [\bar{S}, s_{\max}] \times [0, T]$. T, $\bar{S}$, $s_{\max}$, Z_1, Z_2, x_{min}, and $x_{\max}$ are specified based on reasonable values for the climate change problem, and are given in Section 5.

Remark 1 (Admissible sets Z_1, Z_2)**.** *We will assume in the following that Z_1, Z_2 are compact discrete sets. Since e_1 and e_2 are the result of policy decisions about appropriate regional emissions levels, we argue that it is reasonable to consider these as discrete sets. We envision*

[7] Dixit & Pindyck (1994) provide an introductory treatment of optimal decisions under uncertainty characterized by an Ito process such as Equation (3). A more advanced treatment in a finance context is given by Bjork (2009). Note that we are applying Ito's Lemma to infinitely smooth test functions, as required by viscosity solution theory. This does not require that the value function be smooth. See Barles & Souganidis (1991).

governments being limited in their ability to finely tune emissions levels, but able to implement policies that change emissions to one of a range of possibilities. A sensitivity of different admissible sets is contained in Appendix D. Reisinger & Forsyth (2016) show that as the difference between elements in the discrete choice set go to zero, the solution converges to that of a continuous control space.

Boundary conditions for the PDEs are specified below.

- For fixed $\bar{X}$, as $x \to x_{max}$, it is assumed that $|\frac{\sigma^2}{2}\frac{\partial^2 V_p}{\partial x^2}|$ is small compared to $|\eta(\bar{X} - x)\partial V_p/\partial x|$. Intuitively this boundary condition implies that the impact of volatility at very high temperature levels is unimportant. At extreme temperature levels, the optimal emissions are zero. Assuming that $x_{max} > \bar{X}$, Equation (11) has outgoing characteristics (assuming $\frac{\sigma^2}{2}\frac{\partial^2 V_p}{\partial x^2}$ can be ignored at $x = x_{max}$) and hence no other boundary conditions are required.

- As $x \to x_{min}$, where x_{min} is below the pre-industrial temperature, the effect of volatility is small compared to the drift term. Hence we set $\sigma = 0$ at $x = x_{min}$. Assuming $x_{min} < \bar{X}$ then Equation (11) has outgoing characteristics at $x = x_{min}$ and no other boundary conditions are required. Note that we will show that $\pi_p \geq 0$ at $x = x_{min}$.

- As $s \to s_{max}$, it is assumed that emissions do not increase s beyond the limit of s_{max}. s_{max} is set to be a large enough value so that there is no impact on utility or optimal emission choices for s levels of interest. We have verified this in our computational experiments. This amounts to dropping the term $\frac{\partial V_P}{\partial S}(e_1 + e_2)$ from Equation (12). This can be justified by noting that if $s_{max} \gg \bar{S}$ then $\rho(\bar{S} - S) >> (e_1 + e_2)$ for reasonable values of e_1 and e_2.

- As $s \to \bar{S}$, no extra boundary condition is needed as we assume $e_1, e_2 \geq 0$ hence the Equation has outgoing characteristics at $s = \bar{S}$.

- At $t = T$, it is assumed that V_p is equal to the present value of the infinite stream of benefits associated with a given temperature when emissions are set to their maximum level. Essentially, it is assumed that players receive the costs associated with that temperature in perpetuity and T is large enough that we assume the world has decarbonized.

More details of the numerical solution of the system of PDEs are provided in Appendix A.

4.2 Advancing the solution backward in time from $t_m^+ \to t_m^-$

Going backward in time, the optimal control, is determined between $t_m^+ \to t_m^-$. We consider three possibilities for selection of the controls (e_1^+, e_2^+) at $t \in \mathcal{T}$: Stackelberg, social planner, and Nash. Below we describe the Stackelberg and social planner controls. We include the Nash case for reference only and the Nash controls are describe in Appendix B. We remind the reader that our controls are assumed to be feedback, i.e. a function of state. However, to avoid notational clutter in the following, we will fix (e_1, e_2, s, x, t_m), so that, if there is no ambiguity, we will write (e_1^+, e_2^+) which will be understood to mean $(e_1^+(e_1, e_2, s, x, t_m), e_2^+(e_1, e_2, s, x, t_m))$.

Remark 2. *In all cases the objective function for both players is given in Equation* (8)*. For each type of game there are constraints on the permitted controls which are apparent from the different best response functions defined below for the Stackelberg game and in Appendix B for the Nash game.*

4.2.1 Stackelberg Game

In the case of a Stackelberg game, suppose that, in forward time, player 1 goes first, and then player 2. Conceptually, we can then think of the time intervals (in forward time) as $(t_m^-, t_m]$, (t_m, t_m^+). Player 1 chooses control e_1^+ in $(t_m^-, t_m]$, then player 2 chooses control e_2^+ in (t_m, t_m^+).

We suppose at t_m^+, we have the value functions $V_1(e_1, e_2, s, x, t_m^+)$ and $V_2(e_1, e_2, s, x, t_m^+)$.

Definition 1 (Response set of player 2). *The best response set of player 2,* $R_2(\omega_1, e_1, e_2, s, x, t_m)$ *is defined to be the best response of player 2 to a control* ω_1 *of player 1.*

$$R_2(\omega_1, e_1, e_2, s, x, t_m) \;=\; \operatorname*{argmax}_{\omega_2' \in Z_2} V_2(\omega_1, \omega_2', s, x, t_m^+) \;;\; \omega_1 \in Z_1 \;. \tag{13}$$

Remark 3 (Tie breaking). *We break ties by (i) staying at the current emission level if possible, or (ii) choosing the lowest emission level. Rule (i) has priority over rule (ii). Note that rule (i) corresponds to an infinitesimal switching cost and rule (ii) to an infinitesimal green reward (see Section 5.3.3). Consequently there are no ties after applying either of these rules.*

Similarly, we define the best response set of player 1.

Definition 2 (Response set of player 1). *The best response set of player 1,* $R_1(\omega_2, e_1, e_2, s, x, t_m)$ *is defined to be the best response of player 1 to a control* ω_2 *of player 2.*

$$R_1(\omega_2, e_1, e_2, s, x, t_m) \;=\; \operatorname*{argmax}_{\omega_1' \in Z_1} V_1(\omega_1', \omega_2, s, x, t_m^+) \;;\; \omega_2 \in Z_2 \;. \tag{14}$$

Again, to avoid notational clutter, we will fix (e_1, e_2, s, x, t_m) so that we can write without ambiguity $R_1(\omega_2) = R_1(\omega_2, e_1, e_2, s, x, t_m)$ and $R_2(\omega_1) = R_2(\omega_1, e_1, e_2, s, x, t_m)$.

Remark 4 (Dependence on states e_1, e_2). *In Equations* (13) *and* (14) *the tie breaking rule induces dependence on the initial state,* e_1, e_2.

Definition 3 (Stackelberg Game: Player 1 first). *The optimal controls* (e_1^+, e_2^+) *assuming player 1 goes first are given by*

$$\begin{aligned} e_1^+ &= \operatorname*{argmax}_{\omega_1' \in Z_1} V_1(\omega_1', R_2(\omega_1'), s, x, t_m^+) \;, \\ e_2^+ &= R_2(e_1^+) \;. \end{aligned} \tag{15}$$

Since we use dynamic programming, we determine the optimal controls using the following algorithm.

Stackelberg Control: Player 1 first

Input: $V_1(e_1, e_2, s, x, t_m^+)$, $V_2(e_1, e_2, s, x, t_m^+)$.

Step 1: Compute the best response set for player 2 assuming player 1 chooses control ω_1 first, $\forall \omega_1 \in Z_1$, using Equation (13), giving $R_2(\omega_1)$.

Step 2: Determine an optimal pair (e_1^+, e_2^+) using Equation (15).

Determine solution at t_m^-

$$\begin{aligned} V_1(e_1, e_2, s, x, t_m^-) &= V_1(e_1^+(\cdot), e_2^+(\cdot), s, x, t_m^+) \ ; \\ V_2(e_1, e_2, s, x, t_m^-) &= V_2(e_1^+(\cdot), e_2^+(\cdot), s, x, t_m^+) \ . \end{aligned} \tag{16}$$

Output: $V_1(e_1, e_2, s, x, t_m^-), V_2(e_1, e_2, s, x, t_m^-)$

4.2.2 Social Planner

For the social planner case, we have that an optimal pair (e_1^+, e_2^+) is given by

$$(e_1^+, e_2^+) = \underset{\substack{\omega_1 \in Z_1 \\ \omega_2 \in Z_2}}{\text{argmax}} \left\{ V_1(\omega_1, \omega_2, s, x, t_m^+) + V_2(\omega_1, \omega_2, s, x, t_m^+) \right\} . \tag{17}$$

and as a result

$$V_1(e_1, e_2, s, x, t_m^-) = V_1(e_1^+, e_2^+, s, x, t_m^+) \quad ; \quad V_2(e_1, e_2, s, x, t_m^-) = V_2(e_1^+, e_2^+, s, x, t_m^+) \ . \tag{18}$$

Ties are broken by minimizing $|V_1(e_1^+, e_2^+, s, x, t_m^+) - V_2(e_1^+, e_2^+, s, x, t_m^+)|$. In other words, the social planner picks the emissions choices which give the most equal distribution of welfare across the two players.

5 Detailed model specification and parameter values

This section describes the functional forms and parameter values used in the numerical application. Assumed parameter values are summarized in Table 2.

5.1 Carbon stock details

The evolution of the carbon stock is described in Equation (2). In Integrated Assessment Models, there is typically a detailed specification of the exchange of carbon emissions between the various carbon reservoirs: the atmosphere, the terrestrial biosphere and different ocean layers (Nordhaus 2013, Lemoine & Traeger 2014, Traeger 2014, Golosov et al. 2014). In Equation (2) the removal function is given as $\rho(X, S, T)$. In our numerical application, we use a simplified specification, based on Traeger (2014), to avoid the creation of additional path dependent variables which increase computational complexity. We denote the rate at which carbon is removed from the atmosphere by $\rho(t)$ and assume it is a deterministic function of time which approximates the removal rates in the DICE 2016 model.

$$\rho(t) = \bar{\rho} + (\rho_0 - \bar{\rho})e^{-\rho^* t} \tag{19}$$

ρ_0 is the initial removal rate per year of atmospheric carbon, $\bar{\rho}$ is a long run equilibrium rate of removal, and ρ^* is the rate of change in the removal rate. Specific parameter assumptions for this Equation are given in Table 2. The resulting removal rate starts at 0.01 per year and falls to 0.0003 per year within 100 years.

The pre-industrial equilibrium level of carbon, $\bar{S}$ in Equation (2), is assumed to be 588 gigatonnes (GT) based on estimates used in the DICE (2016)[8] model for the year 1750. The allowable range of carbon stock is given by s_{min} = 588 GT and s_{max} = 10000 GT. s_{max} is

[8] The 2013 version of the DICE model is described in Nordhaus & Sztorc (2013). GAMS and Excel versions for the updated 2016 version are available from William Nordhaus's website: http://www.econ.yale.edu/ nordhaus/homepage/.

Table 2: Base Case Parameter Values

Parameter	Description	Equation Reference	Assigned Value
$\bar{S}$	Pre-industrial atmospheric carbon stock	(2)	588 GT carbon
s_{min}	Minimum carbon stock	(2)	588 GT carbon
s_{max}	Maximum carbon stock	(2)	10000 GT carbon
$\bar{\rho}$, ρ_0, ρ^*	Parameters for carbon removal equation	(19)	0.0003, 0.01, 0.01
ϕ_1, ϕ_2, ϕ_3	Parameters of temperature equation	(20)	0.02, 1.1817, 0.088
ϕ_4	Forcings at CO2 doubling	(22)	3.681
$F_{EX}(0)$ $F_{EX}(100)$	Parameters from forcing equation	(22)	0.5 1
α_1, α_2	Ratio of the deep ocean to surface temp, $\alpha(t) = \alpha_1 + \alpha_2 \times t$, t is time in years with 2015 set as year 0	(20)	0.008, 0.0021
σ	Temperature volatility	(20)	0.1
x_{min}, x_{max}	Upper and lower limits on average temperature, °C	(20)	-3, 20
a_p	Parameter in benefit function, player p	(24)	10
Z_1, Z_2	Admissible controls	(7)	0,1,2,...,10
$\bar{E}$	Baseline emissions	(27)	10
κ_1	Linear parameter in cost function for both players	(26)	0.75
κ_2	Exponent in cost function for both players	(26)	2 or 3
κ_3	Term in exponential cost function for both players	(25)	1
θ_P	WTP for emissions reduction by player p	(4)	0 or 3
T	terminal time	(5)	150 years
r	risk free rate	(12)	0.01
$(t_{m+1} - t_m)$	fixed time between decision dates	(5)	2 years

set well above the 6000 GT carbon in Nordhaus (2013) and will not be a binding constraint in the numerical examples.[9] A 2014 estimate of the atmospheric carbon level is 840 GT.[10]

5.2 Stochastic process temperature: details

Equation (3) specifies the stochastic differential equation which describes temperature $X(t)$ and includes the parameters $\eta(t)$ and $\bar{X}(t)$. To relate Equation (3) to common forms used in the climate change literature, we rewrite it in the following format:

$$dX = \phi_1 \Big[F(S,t) - \phi_2 X(t) - \phi_3 [1-\alpha(t)] X(t) \Big] dt + \sigma d\mathcal{Z} \tag{20}$$

where ϕ_1, ϕ_2, ϕ_3 and σ are constant parameters.[11] The drift term in Equation (20) is a simplified version of temperature models typical in Integrated Assessment Models, based on Lemoine & Traeger (2014). $\alpha(t)$ represents the ratio of the deep ocean temperature to the mean surface temperature and, for simplicity, is specified as a deterministic function of time.[12] Equation (20) is equivalent to Equation (3) with:

$$\begin{aligned} \eta(t) &\equiv \phi_1 \Big(\phi_2 + \phi_3 (1-\alpha(t)) \Big) \\ \bar{X}(t) &\equiv \tfrac{F(S,t)}{(\phi_2 + \phi_3 (1-\alpha(t)))}. \end{aligned} \tag{21}$$

$F(S,t)$ refers to radiative forcing, and it measures additional energy trapped at the earth's surface due to the accumulation of carbon in the atmosphere compared to preindustrial levels and also includes other greenhouse gases,

$$F(S,t) = \phi_4 \left(\frac{\ln(S(t)/\bar{S})}{ln(2)} \right) + F_{EX}(t) \ . \tag{22}$$

[9] Golosov et al. (2014) chose a maximum atmospheric carbon stock of 3000 GT which is intended to reflect the carbon stock that results if most of the predicted stocks of fossil fuels are burned in "a fairly short period of time" (page 67).

[10] According to the Global Carbon Project, 2014 global atmospheric CO2 concentration was 397.15 ± 0.10 ppm on average over 2014. At 2.21 GT carbon per 1 ppm CO2, this amounts to 840 GT carbon.(www.globalcarbonproject.org)

[11] ϕ_1, ϕ_2, ϕ_3 are denoted as ξ_1, ξ_2, and ξ_3 in Nordhaus (2013).

[12] We are able to get a good match to the DICE2016 results using a simple linear function of time.

ϕ_4 indicates the forcing from doubling atmospheric carbon.[13] $F_{EX}(t)$ is forcing from causes other than carbon and is modelled as an exogenous function of time as specified in Lemoine & Traeger (2014) as follows:

$$F_{EX}(t) = F_{EX}(0) + 0.01\big(F_{EX}(100) - F_{EX}(0)\big)\min\{t, 100\} \tag{23}$$

The values for the parameters in Equation (20) are taken from the DICE (2016) model. Note that $\phi_1 = 0.02$ which is the value reported in DICE (2016) divided by five to convert to an annual basis from the five year time steps used in the DICE (2016) model. $F_{EX}(0)$ and $F_{EX}(100)$ (Equation (22)) are also from the DICE (2016) model. The ratio of the deep ocean temperature to surface temperature, $\alpha(t)$, is modelled as a linear function of time. This function approximates the average values from the DICE (2016) base and optimal tax cases.

Useful intuition about the temperature model can be gleaned by substituting parameter values from Table 2 to determine implied values for the speed of mean reversion $\eta(t)$ and the long run temperature mean $\bar{X}(t)$ in Equation (3) for 2015. Using the definitions in Equation (21) it can be determined that $\eta(t) = 0.02$ and $\bar{X} = 1.9°\text{C}$. This value for η implies that, ignoring volatility, temperature would revert to its long run mean in about 50 years. The long run temperature of 1.9°C is above today's value of 1°C above preindustrial levels. This temperature model and assumed parameter values imply considerable momentum in the temperature trajectory.

Figure 1 shows the changes in global surface temperature relative to the 1951 to 1980 average.[14] Based on this data the volatility parameter was estimated using maximum likelihood techniques to be approximately $\sigma = 0.1/\sqrt{\text{year}}$. For the numerical solution we choose $x_{min} = -3$ and $x_{max} = 20$.

As time tends to infinity, the probability density of an Ornstein-Uhlenbeck process is Gaussian with mean $\bar{X}$ and variance $\sigma^2/2\eta$. Our assumed parameter values therefore give a

[13] ϕ_4 translates to Nordhaus's η (Nordhaus & Sztorc 2013).

[14] The data is from NASA's Goddard Institute for Space Studies and is available on NASA's web site Global Climate Change: *https://climate.nasa.gov/vital-signs/global-temperature/*.

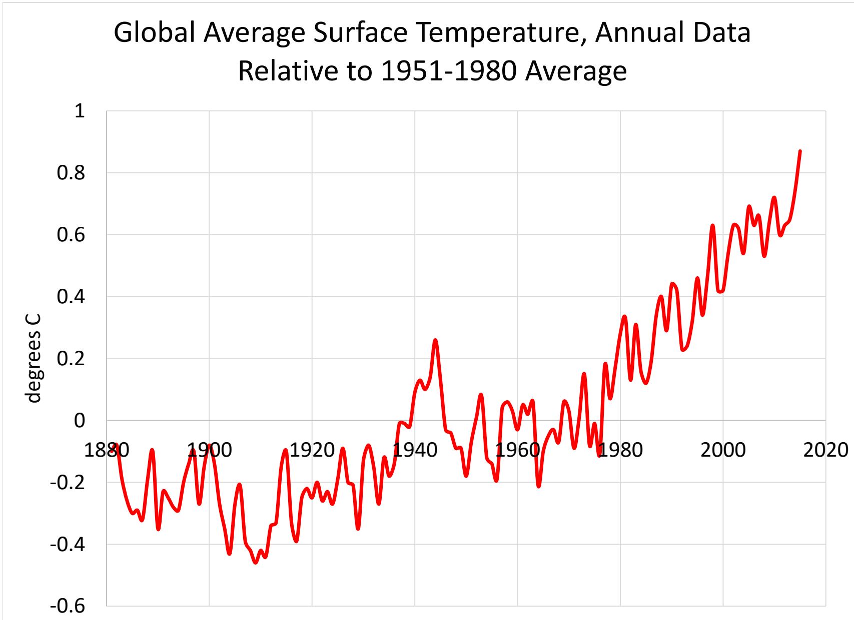


Figure 1: Global land-ocean temperature index, degrees C, annual averages since 1880 relative to the 1951-1980 average

long run standard deviation of 0.44°C and mean of 1.9°C. This implies there is a 2.3 percent probability that temperature could rise by 2 standard deviations (0.88 °C) due solely to randomness, independent of carbon emissions. We conclude that volatility should be an important consideration in any analysis of climate change policies.

5.3 Benefits, Damages and the Green Reward

The term π_p in Equation (4) comprises benefits and damages from emissions as well as the green reward. This section describes these components.

5.3.1 Benefits and Admissible Controls

As is common in the pollution game literature, the benefits of emissions are quadratic according the following utility function:

$$B_p(E_p) = a_p E_p(t) - E_p^2(t)/2,\ p = 1, 2 \tag{24}$$

a_p is a constant parameter which may be different for different players. As in List & Mason (2001), $E_p \in [0, a_p]$ so that the marginal benefit from emissions is always positive.

In the numerical example, there are eleven possible emissions levels for each player $E_p \in \{0, 1, 2, ..., 10\}$ in gigatonnes (GT) of carbon per year and we set $a_1 = a_2 = 10$. We argue that a discrete set of possible emission levels, rather than a continuous set, is more realistic from a policy making perspective. A sensitivity with a finer grid of possible emissions levels is reported in Appendix D.

The controls are applied at fixed time intervals which we set at two years apart. In other words, every two years the leader chooses their optimal control, and immediately thereafter the follower chooses their optimal control.[15]

5.3.2 Damages

Assumptions regarding damages from increasing temperatures are speculative, and this is a highly criticized element of climate change models. The DICE model (and others) specify damages as a multiple of GDP and a quadratic function of temperature, implying that damages never exceed 100 percent of GDP. This formulation ignores possible catastrophic effects. Damage function calibrations are generally based on estimates for the zero to 3°C range above pre-industrial temperatures.

A multiplicative formulation is not appropriate for the model used in this paper in which benefits are zero if emissions are zero (Equation (24)). This is because the multiplicative damage function implies that choosing zero emissions would reduce damages immediately to zero. For this analysis an additive damage function is adopted in which damages rise exponentially with temperature:

$$C_p(t) = \kappa_1 e^{\kappa_3 X(t)} \quad p = 1, 2., \tag{25}$$

where κ_3 is a constant and $p = 1, 2$ refers to the two players. We also explore results with

[15] A sensitivity using one year intervals between the application of controls did not change our results significantly. In (Insley & Forsyth 2019), the impact of increasing the interval between the leader and follower decision times (an interleaved game) is explored in some detail.

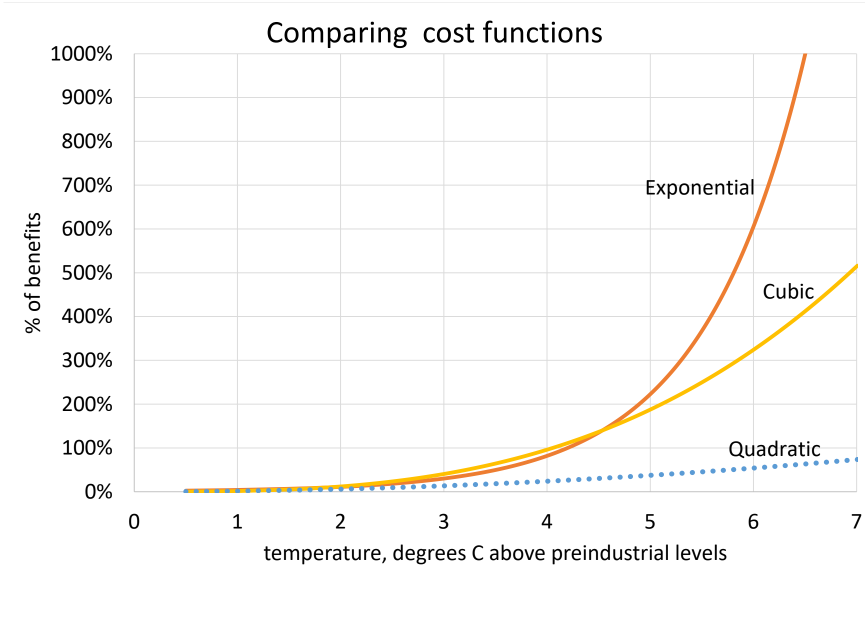


Figure 2: Comparing costs of increased temperatures as a percent of benefits for different cost functions. $\kappa_1 = 0.05$, $\kappa_3 = 1$, $\kappa_2 = 2$ or 3

quadratic or cubic forms of the cost function

$$C_p(X,t) = \kappa_1 X(t)^{\kappa_2} \quad p = 1, 2, . \tag{26}$$

where κ_1 and κ_2 are constants.

We choose the parameters in the damage functions (Equation (26) and (25)) so that damages represent a reasonable portion of benefits at current temperatures levels (i.e. at 0.86 degrees C over preindustrial levels). Base case values for κ_1, κ_2 and κ_3 imply damages of about 1 percent of benefits at current temperature levels. Figure 2 compares the three cost functions as a percentage of benefits. The comparison is for the exponential function and for the power damage function with the exponent set to 2 or 3 in the latter. We observe that the three cost functions are virtually indistinguishable up to 3 °C above pre-industrial levels. After 3 °C the cost functions diverge dramatically. We choose the exponential cost function for our base case as it implies that for temperature increases above 3 °C, damages from climate change would be disastrous, which seems a reasonable supposition. We report on sensitivities with quadratic and cubic damage functions in Section 6.5.

5.3.3 Green Reward

We define emissions reduction, $g_p(t)$, relative to a baseline level of emissions level, $\bar{E}$, for each region.

$$g_p(t) = \max(\bar{E}_p - E_p(t), 0), \quad p = 1, 2 \tag{27}$$

Citizens of each region are assumed to value emissions reduction as contributing to the public good. We denote the degree of environmental awareness in a region by θ_p which represents a willingness to pay for emissions reduction because of a desire to be good environmental citizens, distinct from the expressions for the benefits and costs of emissions as defined in Equations (24) and (25).

The benefit from emissions reduction, called the green reward, A_p, depends on environmental awareness as well as emissions reduction in both regions:

$$A_p(t) = \theta_p g_p(t), \quad p = 1, 2. \tag{28}$$

In our base case, $\theta_p = 0$ for both players initially. We then explore differential green preferences by setting $\theta_p = 3$ for one of the players. In future work, we will explore the possibility that environmental preferences may evolve randomly over time and may depend on environmental actions taken in the other region.

6 Numerical results

In this section we analyze results for four different cases of interest. In the base case, players are identical, the willingness to pay for emissions reduction due to the green reward is zero, and assumed parameter values are as described as in Table 2. In the second case, players are also identical but temperature is much more volatile than in the base case. In the third and fourth cases, players are asymmetric, differing either in terms of damages from increased temperature or in terms of preferences for emissions reduction (i.e. green preferences). In all cases the damage function is assumed to be exponential as in Equation (25), but we report sensitivity analysis for quadratic and cubic damage functions in Section 6.5.

The numerical results are depicted in two different ways. Firstly, the optimal controls, (e_p^+) and expected utilities, (V_p), of the players are shown at time zero for particular values of state variables. Secondly we undertake Monte Carlo simulations of the stochastic state variables and apply the previously determined optimal controls to simulate possible paths, going forward in time, for temperature, atmospheric carbon stock, player emissions and utilities, given assumptions about starting values for the state variables. The Monte Carlo analysis allows us to compute percentiles for variables of interest. In the results discussion, player 1 always refers to the leader in the Stackelberg game, and player 2 refers to the follower.

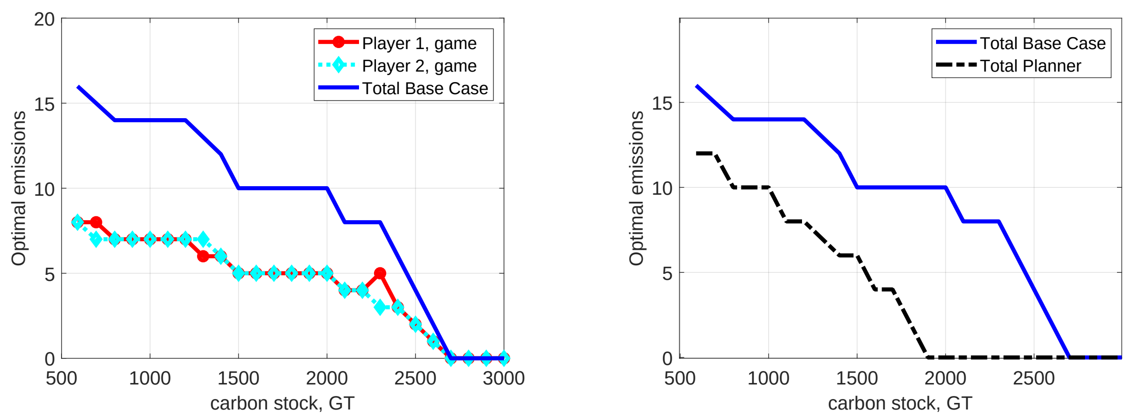
6.1 Base case: identical players

Figure 3 depicts the optimal controls for the game and the social planner versus the stock of carbon at time zero, conditional on a temperature of 1.0 °C (close to the current value), and starting emissions for both players at 10 GT. For reference, recall that the stock of carbon in 2017 was about 870 GT.[16] The expected time path of optimal controls is captured in the Monte Carlo analysis below.

Figure 3(a) shows the optimal controls for individual players and the resulting total for the game, while Figure 3(b) compares total emissions choices under the game (repeated from Figure 3(a)) versus the social planner. The individual players' choices of emissions are below the initial value of 10 GT for all carbon stock levels, starting at 7 GT for low levels of S and then falling as S increases, reaching zero at about 2700 GT of carbon. (Note that the jump up and then down for Player 1's emissions at $S = 2300$ GT is the result of a very flat value surface around this point, so that there is little difference in value between a choice of

[16] Note that we can also show similar graphs for any time period between $t = 0$ and $t = T$. The optimal controls for other time periods will be the same as at time zero, until the boundary condition at $t = T$ begins to have an effect.

4 GT versus 5 GT for the optimal control.) The social planner chooses lower total emissions at every level of carbon stock, compared to the game, with emissions of zero if the stock of carbon is at 1800 GT or above. (Recall that the social planner maximizes total utility, which implies equalizing emissions between the two players, since players are symmetric in the benefits received from emissions.) Similar graphs can be drawn showing optimal controls versus temperature for given values of the carbon stock. These graphs (not shown) indicate that the optimal choice of emissions falls with increasing temperature.



(a) Stackelberg base game, total and player emissions

(b) Social planner and Stackelberg base game, total emissions

Figure 3: Optimal control versus carbon stock, at time zero. Contrasting game and social planner, base case. State variables: temperature = 1 degrees C above pre-industrial levels, and initial emissions of 10 GT for both players.

Figure 4 plots expected utilities, V_p, at time zero for the game and the social planner versus various initial temperature levels at $S = 800$ GT, consistent with these optimal controls. We observe that utility declines with the initial temperature, as expected. Under the game, player 1 has slightly higher utility than player 2. Recall that this is a repeated game which is played (i.e. optimal control applied) every two years over 150 years. Since the leader is able to choose an optimal control first, with knowledge of how the other player will react, this imparts some advantage to the leader depending on the values of the state variables. Player utilities are identical under the social planner, and hence are not shown. The social planner choices result in significantly higher utility than under the game, indicating a tragedy of the commons whereby strategic interactions of the two decision makers leave both worse off than when decisions are made by a planner. Similar plots of utility could be drawn for different starting values of S. For higher S values, the utility curves shift inward.

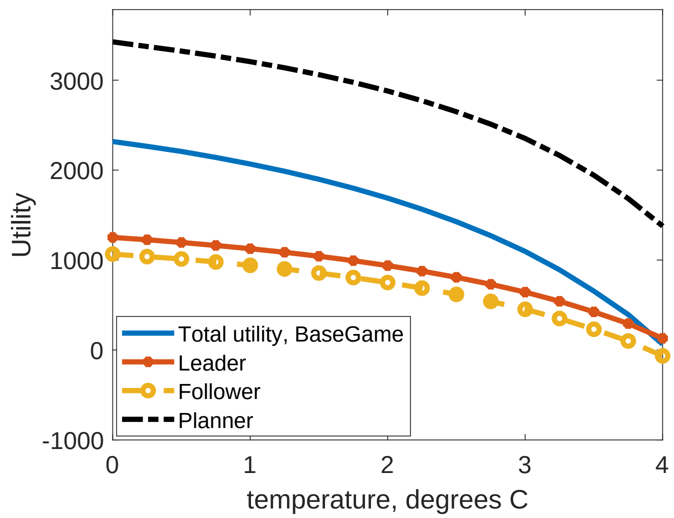
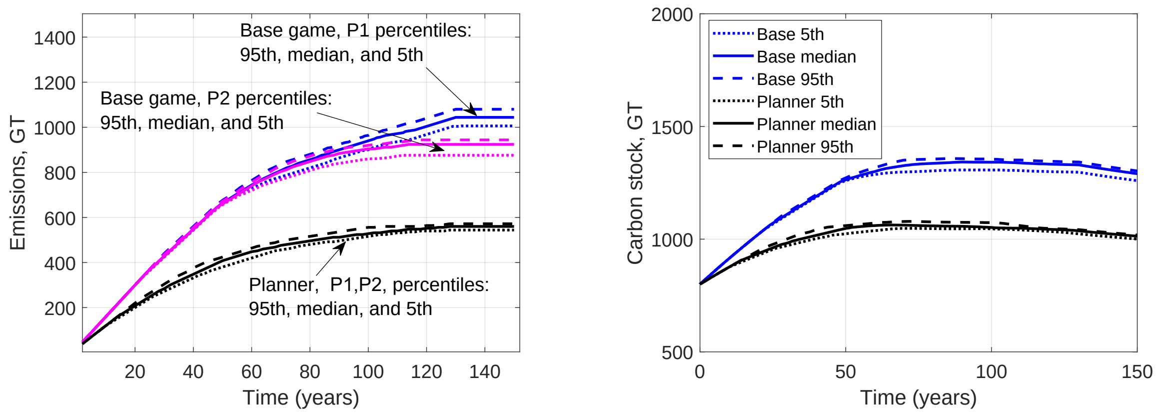


Figure 4: Utility versus temperature, comparing base case game (total, player 1 and player 2) and social planner (total). Utility refers to V_p defined in Equation (8) for player p. In the planner case, the sum of player utilities is shown. Time zero. Exponential damage function. Stock of carbon at 800 GT. Initial emissions at $E_1 = E_2 = 10$

We use Monte Carlo simulation to illustrate the evolution of cumulative emissions, the carbon stock, average global temperature, and utility assuming players follow optimal strategies, as previously computed through the numerical solution of the optimal control problem (Equation (8)), and given assumed starting values at time zero. Figure 5(a) depicts median, 5th and 95th percentiles for cumulative emissions for the players in the base case game contrasted with cumulative emissions given the social planner's choices. This is based on 10,000 Monte Carlo simulations in which players choose the optimal control and state variables evolve accordingly. We observe that median cumulative emissions for the social planner are much lower than in the game over the entire 150 years. In addition, player 1 in the base game has higher median emissions than player 2 beginning at about year 60. Recall that players had similar optimal controls at time zero in Figure 3(a), but from the Monte Carlo simulation depicted in 5(a), it is clear that optimal controls diverge as time goes forward with the leader able to benefit by choosing higher emissions. In contrast, the social planner chooses the equal emissions for players 1 and 2. Figure 5(b) depicts percentiles for the carbon stock. Median carbon stock for both the planner and the game initially rises, and then eventually starts to drop as emissions go to zero and natural processes gradually remove some carbon from the atmosphere. Consistent with the paths shown for cumulative emissions, the carbon stock under the social planner is lower than under the game and also starts to decline sooner.

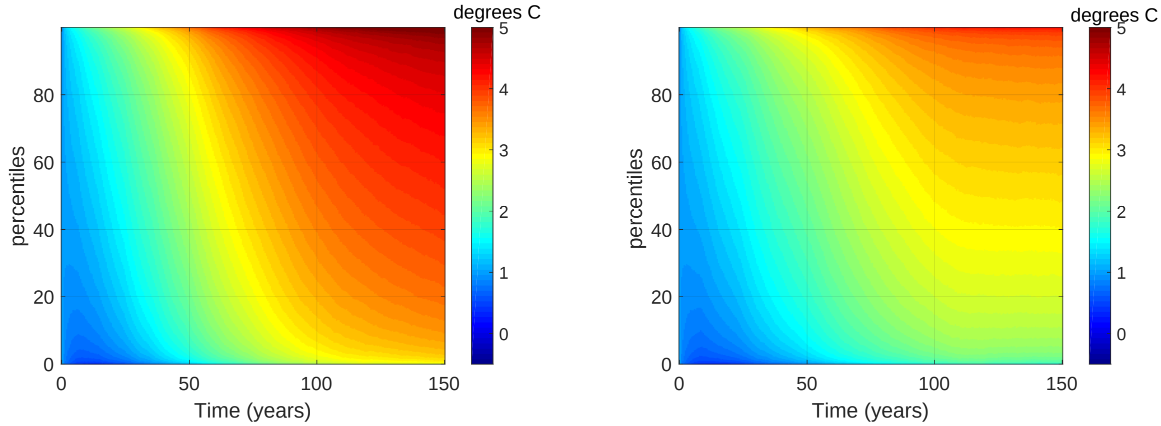


(a) Cumulative emissions percentiles, base and planner

(b) Carbon stock percentiles, Base Game and Planner

Figure 5: Cumulative emission percentiles versus time, base game and social planner, $X(0) = 1$, $S(0) = 800$, , $E_1(0) = E_2(0) = 10$. Dashed lines = 95th percentiles, solid lines = medians, dotted lines = 5th percentiles. 10,000 simulations.

The cumulative emissions and carbon stock affect the expected path of temperature over time. One way to view possible future temperature paths is via a heat map. Figure 6(a) shows 10,000 possible realizations of the path of temperature with temperature values represented by colours according to the legend given on the right of the graph. Blue represents cooler temperatures while red represents hotter temperatures. The graph shows the distribution of temperature in terms of percentiles (y-axis) going forward in time (x-axis). Figure 6(b) is a similar plot for the social planner case. The differences in these two graphs become most apparent after 50 years, from which point the hotter colours of 3 C°(above pre-industrial levels) and greater are much more in evidence for the game. By year 75, the 25th percentile



(a) Temperature map, base game

(b) Temperature map, planner

Figure 6: Temperature maps for base case game and social planner. $X(0) = 1$ C°, $S(0) = 800$ GT, , $E_1(0) = E_2(0) = 10$ GT. 10,000 simulations.

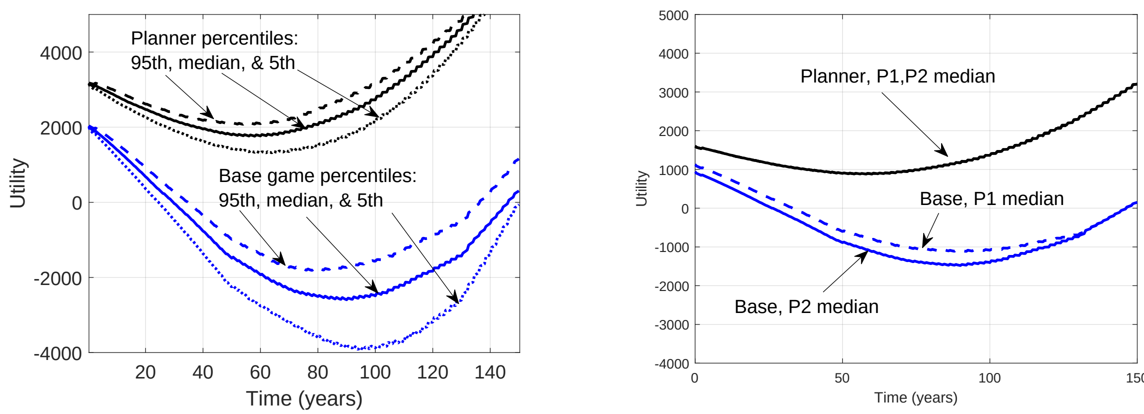
	50 years		**100 years**	
	Base Game	**Planner**	**Base Game**	**Planner**
25th percentile	1.79	1.45	2.96	2.25
median	2.50	2.12	3.67	2.96
95th percentile	3.18	2.81	4.36	3.62

Table 3: Selected temperature percentiles, C°, for base case game and social planner

in the game is at about 3°C whereas the social planner is below 2.5°C. Table 3 highlights some other key percentiles from the graphs.

Figure 7 depicts expected utility, $V_p(e_1, e_2, s, x, t)$, from $t = 0$ to $t = 150$ for the game and the social planner. At $t = 0$, the values shown match those in Figure 4. As t increases, V_p evolves over time showing the expected present value of starting the game at a given $t_m > 0$. For example, at 80 years the combined expected value for the two players in the game (left graph) is approximately -2500 utils. This means that for players starting this game in year 80, the present value of total combined expected utility from year 80 to year 150 is -2500 utils. The left hand graph shows that median utility under the planner is much higher than under the game. Total median utility initially declines for both cases, but eventually rises as the boundary condition at time $T = 150$ has an effect. Recall that at time T it is assumed that the economy is decarbonized and emissions no longer add to the stock of carbon. At $T = 150$ the economy benefits from carbon emissions, but faces damages depending on the long term equilibrium temperature implied by the carbon stock in that year. These net benefits are received as a perpetual annuity.[17]

We also observe in Figure 7 that the 95th and 5th percentiles are much more spread apart in the game than under the social planner, indicating that the game is more risky in terms of the variability of possible outcomes. The right hand graph shows that both players have the same median utility under the planner (as expected since players are symmetric), while under the game, player 1 has slightly higher median utility over most of the 150 years.



(a) Total utility percentiles, Base Game and Planner (b) Individual player utility percentiles, Base Game and Planner

Figure 7: Utility percentiles over time for base case game and social planner. Utility refers to V_p defined in Equation (8) for player p. In the planner case, the sum of player utilities is shown. $X(0) = 1$ C°, $S(0) = 800$ GT, $E_1(0) = E_2(0) = 10$ GT. Dashed lines = 95th percentiles, solid lines = medians, dotted lines = 5th percentiles. 10,000 simulations.

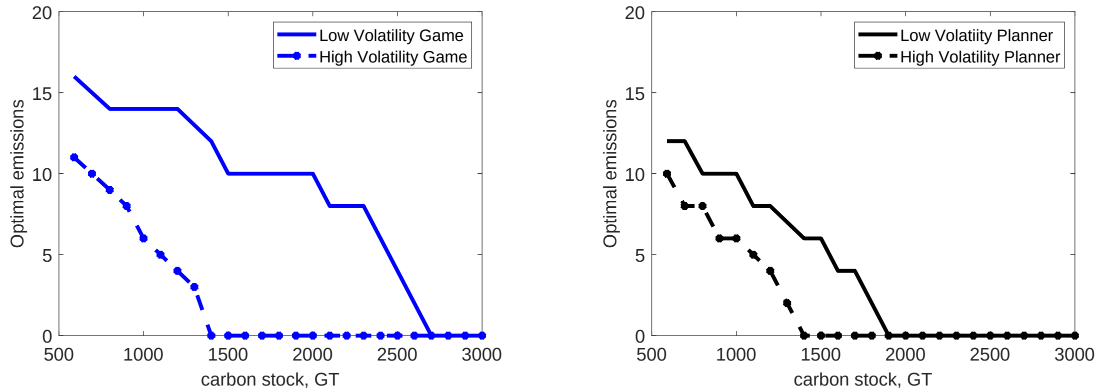
[17] A sensitivity was carried out with $T = 200$. The median temperature path matched the base case closely for the first 75 years, and then for the next 75 years temperature in the sensitivity ($T = 200$) case was slightly above the base case. The utility profiles for the sensitivity case have the same shape, but are consistently lower than those in Figure 7, reflecting the longer time until decarbonization.

6.2 Importance of temperature volatility

As noted in Section 5.2, average global temperature exhibits significant volatility and in this section, we analyze its impact on the outcome of the game. Figure 8 compares the optimal controls at time zero for the base (low volatility) case where $\sigma = 0.1$ and a high volatility case where $\sigma = 0.3$. In Figure 8(a), we observe that a higher volatility reduces total optimal emissions significantly in the game. (Individual player emissions are not shown as they are quite similar to each other at time zero.) The same is true for the social planner (Figure 8(b)), however the relative reduction is much larger for the game. Median cumulative emissions over time are compared in Figure 9. In Figure 9(a) we observe that in the high volatility case, median cumulative emissions for player 1 exceed those of player 2 beginning around year 50, whereas for the low volatility case, player 1 exceeds player 2 median cumulative emissions closer to year 100. As already noted, at time zero the optimal controls are similar for the leader and follower in both high and low volatility cases, but diverge over time as indicated by the Monte Carlo analysis. The results in Figure 9(a) indicate that the follower makes the greater relative sacrifice in emissions reduction when volatility is higher. This observation is confirmed in the comparison of player utilities plots (Figure 11(c)) discussed below.

Figure 9(b) contrasts the two players' emissions under the social planner for the high and low volatility cases, showing that the planner curbs emissions significantly along the median path with more volatile temperatures. These results make sense given that damages are highly convex in temperature, causing both the social planner and the players of the game to react accordingly. Figure 9(c) shows the median path for atmospheric carbon stock is highest for the low volatility game over the entire 150 years, followed by the high volatility game, then the low volatility planner then the high volatility planner.

Figure 10(a) shows heat maps for the game and social planner cases in the high volatility scenario. These heat maps produce more optimistic forecasts compared to those shown in Figure 6 in that they indicate a higher probability of lower temperatures throughout the 150 year time frame. From the optimal control discussed above, we know that both the social



(a) Game, combined player emissions (b) Social planner, combined player emissions

Figure 8: High Volatility: Optimal control versus carbon stock for high volatility ($\sigma = 0.3$) and low volatility ($\sigma = 0.1$) cases, game (symmetric players) and social planner, exponential damage. Only total combined emissions are shown.

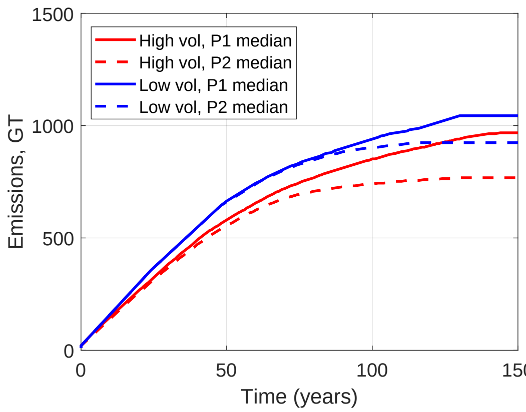
	Game	Planner	ratio planner/game
Base case	2068	3206	1.6
High volatility	558	2081	3.7

Table 4: Total expected utility comparison ($V_1 + V_2$) at time zero. $X(0) = 1$, $S(0) = 800$, $E_1(0) = E_2(0) = 10$. 10,000 simulations.

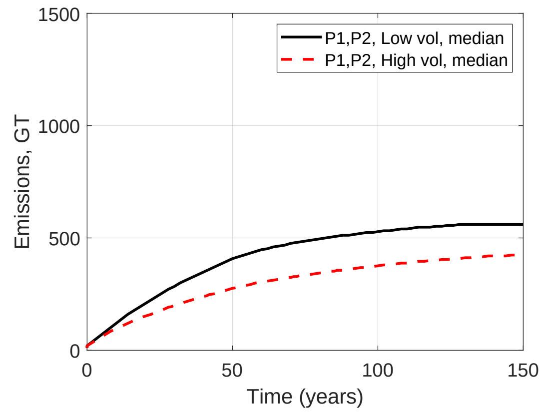
planner and the decision makers in the game reduce emissions when volatility is high to avoid the most damaging temperatures. The 95th percentiles show very high temperatures - over 4.5 C°for both the planner and the game - indicating the high risk of this case whereby even the social planner may not be able to avoid a very negative outcome.

Figures 11(a) and 11(b) show total utility percentiles over time for the game and social planner. For ease of comparison, Table 4 shows numerical values at time zero. The difference in total expected utility between the planner and the game is much larger under the high volatility case. Clearly the benefit of cooperative action, as provided by the social planner, is higher in the high volatility scenario.

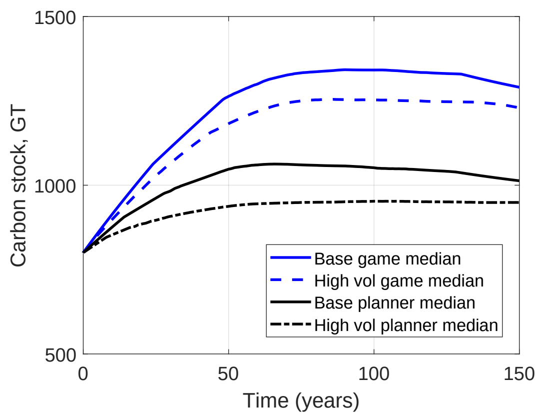
Figure 11(c) compares individual player utilities for the high and low volatility games. We observed previously that player 1 emissions begin to exceed player 2 emissions earlier



(a) Game Cumulative Emissions, Low and high volatility



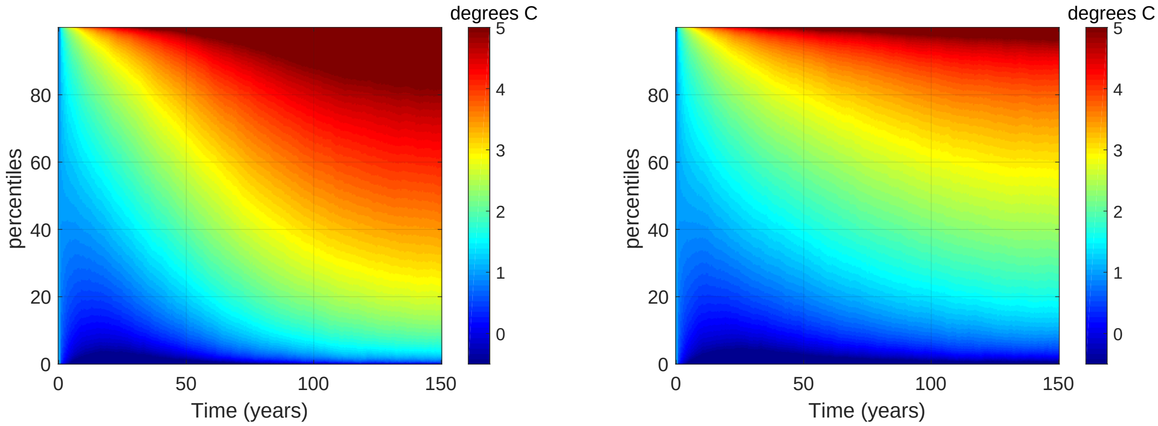
(b) Planner Cumulative Emissions, Low and high volatility



(c) Carbon stock medians, low and high volatility, game and planner

Figure 9: High Volatility: Cumulative player emissions, median values over time, comparing high and low volatility cases. $X(0) = 1$, $S(0) = 800$, , $E_1(0) = E_2(0) = 10$. 10,000 simulations.

in the game in the high volatility case. Consistent with this, we observe that the relative difference between player 1 and player 2 utilities is larger in the high volatility case. So although the value of the game at time zero to both players is less in the high volatility case, the relative advantage of being the first player has increased.



(a) Temperature map, high volatility game

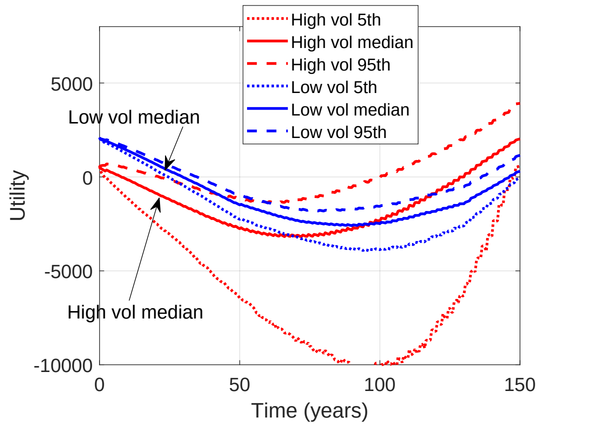
(b) Temperature map,high volatility planner

Figure 10: Temperature maps for high volatility games and social planner. $X(0) = 1$, $S(0) = 800$, , $E_1(0) = E_2(0) = 10$. 10,000 simulations.

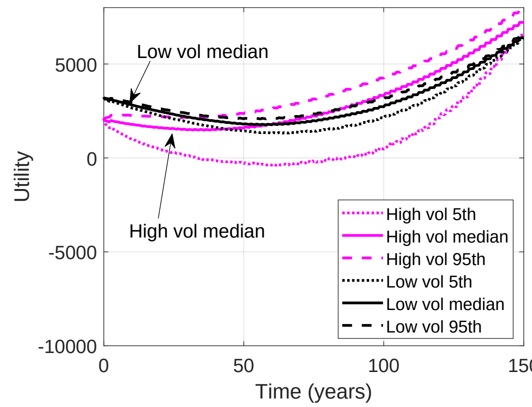
6.3 Asymmetric damages

An important feature of global warming is the distribution of damages across nations, with some of the world's poorer regions suffering disproportionately. In this section we explore the effect of asymmetric damages on strategic interactions by considering a case in which the follower has much higher sensitivity to increasing temperatures than the leader. Specifically, we compare the base case where $\kappa_3 = 1.0$ in Equation (25) for both players to one where $\kappa_3 = 1$ for player 1 and $\kappa_3 = 1.15$ for player 2. We refer to the latter as the asymmetric damages case and the former as the base or symmetric damages case.

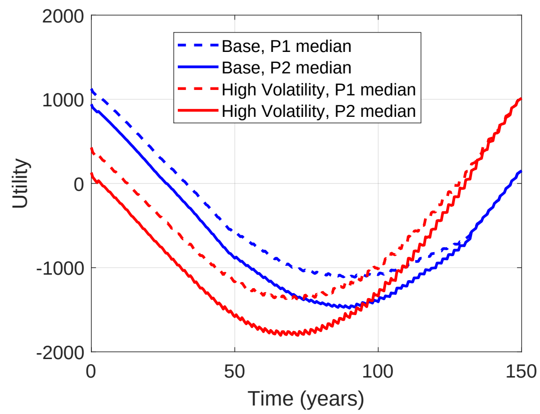
The optimal controls in these cases for various carbon stocks and at a temperature of 1°C are shown in Figure 12. In Figure 12(a) we observe that the follower facing higher damages (red line) starkly curtails emissions, compared to the base case (blue line). Similarly the planner chooses lower emissions for player 2 in the asymmetric case (magenta line) compared to the symmetric case (black line). Figure 12(b) depicts the leader's optimal controls. Comparing the blue (symmetric case) and red (asymmetric case) line, we observe that for lower levels of the carbon stock, the leader chooses higher emissions under the asymmetric case. The fact that the follower experiences higher damages allows the leader



(a) Total utility for the game: High vs low volatility



(b) Total utility for the planner: High vs low volatility



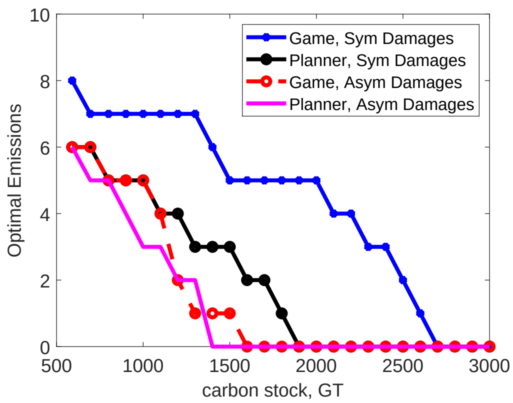
(c) Individual player utility for the game: High vs low volatility

Figure 11: High Volatility: Utility percentiles over time, comparing high and low volatility cases, game and social planner. Utility refers to V_p defined in Equation (8) for player p. In the planner case, the sum of player utilities is shown. $X(0) = 1$, $S(0) = 800$, $E_1(0) = E_2(0) = 10$. 10,000 simulations.

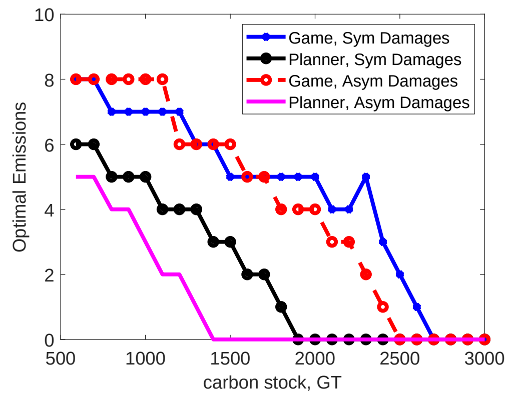
to take advantage and increase their own emissions. However, this result does not hold for higher levels of the carbon stock. For $S > 1700$ the leader chooses lower emissions in the asymmetric damages case. This is an interesting interaction of the two players. In effect for these large levels of the carbon stock, the asymmetry in damages reduces the tragedy of the commons compared to the symmetric case. The leader knows that the follower will curtail their emissions due to the higher damages it experiences. Therefore the leader is able to reduce emissions, knowing that the follower will not fill in the gap. While Figure 12 is drawn for a current temperature of 1 C°, this same phenomenon is observed when other temperatures (such as 2, 3 or 4 C°) are chosen as the reference point. Note that these results also hold when the leader has the higher damages. In this case (not shown) the follower takes advantage and increases their own emissions at low carbon stock levels, but curtails their emissions at high carbon levels (all relative to the symmetric case). Looking at total emissions in Figure 12(c) we observe that emission choices are highest in the symmetric game, followed by the asymmetric game, then the symmetric planner case and then the asymmetric planner case.

Player utility at time zero for different carbon stock levels is depicted in Figure 13(a) for the asymmetric damages case and in Figure 13(b) for the symmetric damages case. We observe that in the asymmetric damages case, player 1's utility is everywhere above that of player 2's and as well the slope $\partial V/\partial S$ is less negative than for player 2. This compares with the right hand graph where the two player are much closer together and the slopes $\partial V/\partial S$ are similar. (Note that the player 1's utility is above that of player 2 in the right hand graph - by 20 percent at $S = 800$ - but this is not visible due to the scale of the graph.) The main point is that when player 2 experiences higher damages, player 1's utility at the beginning of the game is less affected by the current carbon stock (as indicated by $\partial V/\partial S$) because player 2 is motivated to take on a larger share of emissions reduction compared to the case of symmetric damages.

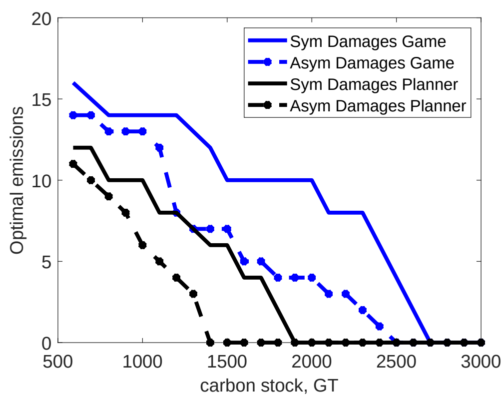
A comparison of median emissions over time is shown in Figure 15 in Appendix C. Along the median path, player 1 increases their cumulative emissions relative to the base case while player 2 reduces their cumulative emissions. Median carbon emissions (Figure 15(c)) remain below 1200 GT, which is also consistent with the range of S values when the leader's actions partially offset those of the follower relative to the base case (as observed from the optimal controls in Figure 12). In contrast, when the planner chooses emissions (Figure 15(b)), player 2 receives larger emissions than player 1. The planner makes up for some of the added damage to player 2 from rising temperatures by allotting more economic activity (as indicated by emissions) to player 2. Median total emissions (Figure 15(c)) are lower in the asymmetric damages case compared to the symmetric case over the entire 150 years, whether for the game or the planner.



(a) Follower optimal emissions



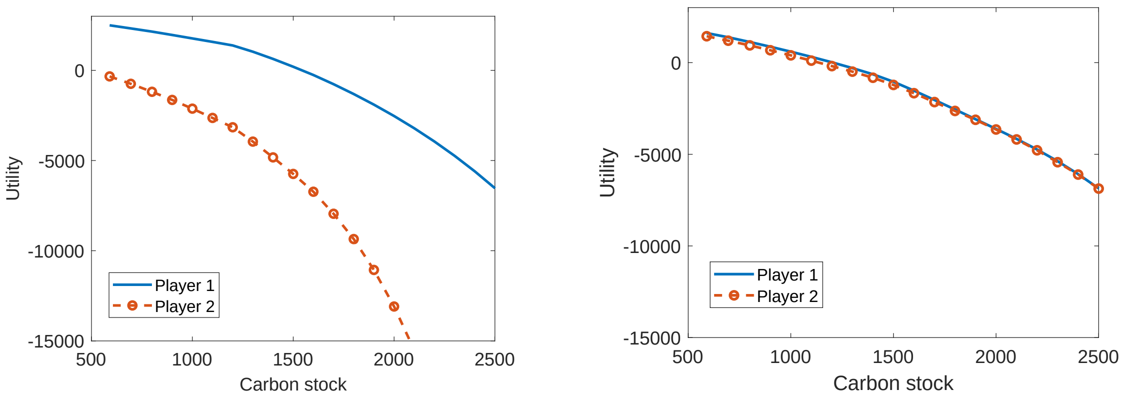
(b) Leader optimal emissions



(c) Total emissions

Figure 12: Asymmetric Versus Symmetric Damages: Optimal controls versus carbon stock in GT, temperature = 1 °C, Symmetric case: $\kappa_3 = 1$ for both players; Asymmetric case: $\kappa_3 = 1$ for player 1 and $\kappa_3 = 1.15$ for player 2

Time paths for other variables of interest in the asymmetric damages case are shown in Figure 16 in Appendix C. Figure 16(a) shows that median temperature is kept lower for the asymmetric cases (game and planner) compared to the symmetric counterparts. Total utility



(a) Asymmetric damages

(b) Symmetric damages

Figure 13: Asymmetric Versus Symmetric Damages: Utility at time zero for follow, leader and total versus stock of carbon in GT. Utility refers to V_p defined in Equation (8) for player p. Current temperature = 1 °C, Symmetric case: $\kappa_3 = 1$ for both players; Asymmetric case: $\kappa_3 = 1$ for player 1 and $\kappa_3 = 1.15$ for player 2

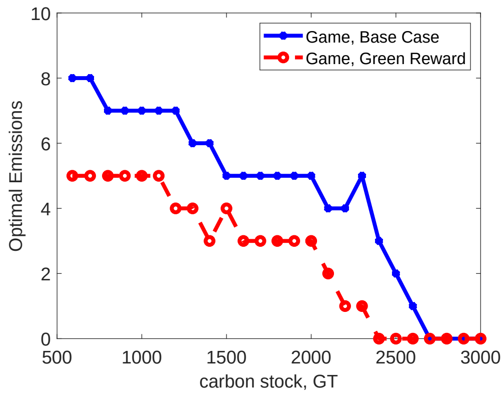
is lower for the asymmetric game over most of the time frame (Figure 16(b). For the planner (also Figure 16(b)), total utility is lower under the asymmetric game until a bit beyond year 50, after which we observe higher utility under the planner asymmetric damages case. This is a reflection of the strong emissions reduction from the first 50 years paying off in utility terms for the last 100 years. An interesting observation (Figure 16(c)) is that player 1 has the highest expected utility in the asymmetric game until year 50, compared with the other three cases (symmetric game, asymmetric planner, symmetric planner), but after that does better under the planner. From the perspective of time zero, player 1 benefits from the fact that player 2 experiences higher damages from climate change. Not surprisingly, player 2 does worse under the game throughout the entire 150 years compared to the other three cases (Figure 16(d)).

6.4 Asymmetric preferences

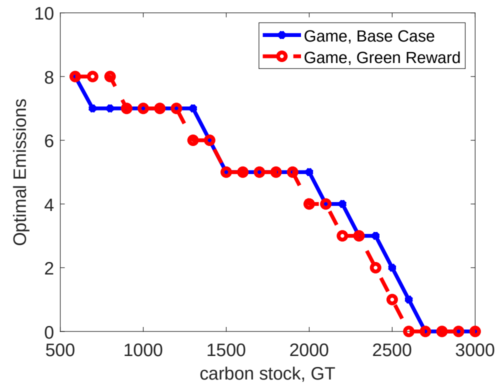
This section examines the impact of asymmetric preferences by considering a case in which one of the players gains a psychic benefit for reducing emissions relative to a given benchmark, which we refer to as the green reward (GR). We report only the outcome of the GR for the Stackelberg game, and not for the social planner case.[18] We assume that the environmentally friendly player (player 1 in this example) is willing to pay 3 utility units, ($\theta_p = 3$) for reductions in emissions below the benchmark $\bar{E}$. The results are shown in Figure 14 which depicts optimal emissions choices at time 0, for different carbon stock levels conditional on a temperature of 1 °C. We observe from Figure 14(a) that, as expected, when the leader has greener sentiments, it chooses a lower level of emissions than in the base case over most carbon stock levels, and chooses the same emissions for $S > 2700$ GT. In Figure 14(b) we observe that the follower increases emissions compared to the base case for low carbon stock levels. However for higher carbon stock levels ($\geq 1400GT$), player 2 has either the same or less carbon emissions than in the base case game. At low carbon stock levels this can be explained as form of green paradox, characterized elsewhere in the literature, whereby increased green sentiments of one player causes the other player to free ride by increasing emissions. At higher carbon stock levels, rather than a green paradox, we have a sort of green bandwagon effect. An explanation is that at high carbon levels, the environmentally friendly policies of the leader make it worthwhile for the follower to also choose environmentally friendly policies, because the follower knows this choice will help avert highly damaging consequences. In other words, green sentiments on the part of player 1, give player 2 more agency to affect future outcomes. This is similar to our observations in the asymmetric damages case above. In Figure 14(c) we observe that total emissions are lower in the green reward case over most carbon stock levels.

Median cumulative emissions are depicted in Figure 17(a) in Appendix C. Player 2's median cumulative emissions are shown to be significantly higher in the GR case compared to the base game. In Figure 17(b) we observe that all carbon stock percentiles are lower in the GR case compared to the base case. Consistent with the these carbon stock paths, Figure 18(a) in Appendix C shows temperature percentiles for the GR case are below those of the base case. Figure 18(b) indicates that both players have substantially higher median

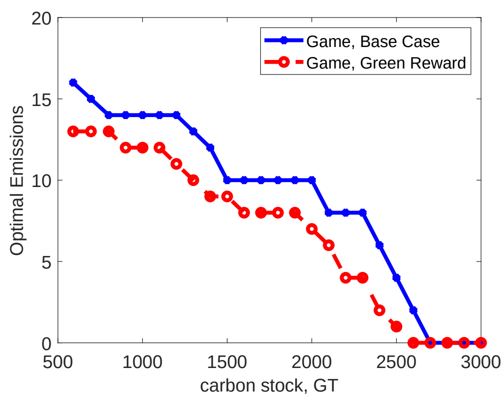
[18]The case of a planner maximizing total utility including one player's green reward seems of less interest than the social planner actions in previous cases.



(a) Leader emissions



(b) Follower emissions



(c) Total emissions

Figure 14: Green reward: Optimal control versus pollution stock when leader receives a green reward for emissions reductions. $X(0) = 1$, $S(0) = 800$, $E_1(0) = E_2(0) = 10$. 10,000 simulations.

utility along the entire 150 year time path in the GR case.

6.5 Alternate damage functions

Sensitivities were conducted using the alternate damage function given by Equation (26) in the case of symmetric players. Using a quadratic function, ($\kappa_2 = 2$), the optimal choice of emissions is near the maximum possible (9 GT for each player, compared to a maximum of 10 GT) in both the game and the social planner. In contrast a cubic damage function,

($\kappa_2 = 3$) results in some curtailment of emissions, but emissions are still at higher levels than with exponential damages and never go to zero. We consider the exponential damage function to be the most reasonable as the damages quickly become very large at temperature above 3°C.

6.6 Checking for Nash equilibria

Recall that we are solving for a repeated series of Stackelberg games which happen every 2 years over the 150 year time span of the analysis. It is of interest to note whether Nash equilibria exist for these repeated games. In each of the cases described above, we check for the existence of Nash equilibria across all state variables and at each of the 75 decision times. We find that at each decision time about 25 percent of the nodes (representing carbon stock, temperature and emissions levels) satisfy the Nash equilibrium criterion. Further, we determine that 8 to 9 percent of the Stackelberg equilibria are also Nash equilibria. See Appendix B for details.

7 Concluding comments

In this paper we have examined the strategic interactions of large regions making choices about greenhouse gas emissions in the face of rising global temperatures. We have modelled optimal decisions of players in a fully dynamic, feedback, repeated Stackelberg game and have demonstrated its numerical solution. The results indicate a classic tragedy of the commons whereby regions acting in their own self interest in a non-cooperative game choose higher levels of emissions and have lower total utility than would be chosen by a social planner. As expected, the leader in the Stackelberg game was found to have an advantage over the follower. However, unexpectedly, this advantage is small relative to the reduction in total utility compared to a cooperative solution as represented by the social planner. We examined the effects of temperature volatility, asymmetric damages and asymmetric preferences on the strategic interactions of players and considered their effects on carbon emissions choices and utilities.

Volatility is found to have an important effect on optimal choices of players in the game as well as the social planner. An increase in volatility increases the likelihood of high temperatures and resulting high damages. This causes players in the game and to social planner to choose lower levels of emissions. The difference in total expected utility between the social planner and the game (the social planner advantage) at time zero is larger for higher volatility, implying that the tragedy of the commons is exacerbated by higher volatility, or, in other words, the need for cooperative action is increased. This conclusion is reinforced by observing percentiles for total utility over 150 years showing that possible outcomes are much more variable in the high volatility Stackelberg game compared to the social planner. In effect, the game becomes more risky relative to the social planner. Although the drift in long run temperature is key in climate change policy, the impact of volatility on strategic interactions of decision makers is significant.

Asymmetric damages are also found to affect the outcome of the game. When one player experiences greater harm with rising temperatures, we find that over lower carbon stock levels, the player with higher damages is made worse off by the response of the other player. While the player with higher damages cuts back on their emissions more aggressively, the low damage player takes advantage of this by increasing their own own emissions relative to the symmetric damage case. At higher carbon stock levels, we find a contrasting interaction of the two players in that the player with lower damages actually reduces their emissions compared to the symmetric damage case. In effect, emissions reduction by the high damage player are reinforced by those of the low damage player, all relative to the symmetric damages case. However, the median path of emissions over 150 years shows the low damage player with higher cumulative emissions compared to the symmetric case. The benefit of cooperative action via a social planner is higher in the case of asymmetric damages versus symmetric damages, as the social planner optimally distributes emissions across the two players, allowing the player experiencing higher damages from climate change to emit more carbon. The impact of the stock of carbon on player interactions in the asymmetric damages case is an interesting conclusion of this paper.

We also examined a case where one of the players receives a psychic benefit from emissions reductions compared to a benchmark, an effect we labelled the green reward. A green reward for one player causes that player to cut back their emissions from what would have otherwise been the case. There are various responses by the player with no green reward (the brown player) ranging from no response, to increasing or decreasing emissions depending on the values of the state variables. At low carbon stock levels the brown player increases their emissions relative to the case of symmetric preferences. This is similar to the green paradox effects observed by Wirl (2011) in a deterministic game whereby an increase in green sentiments increases the free riding of brown players. We also observed a contrary effect, which we call the green bandwagon effect, whereby for some high values of the carbon stock, the presence of a green reward for one player causes the brown player to reduce their own emissions (relative to the case with no green reward). Our interpretation is that at high carbon stocks where disaster is on the horizon, the brown player can be assured that the green player will cut back emissions, making it worthwhile for the brown player to also reduce emissions. The green preferences of the green player give the brown player more agency to effect a change in climate outcomes. While this green band wagon effect is a possibility, we find that along the median path of emissions, the cumulative emissions of the brown player exceed those of the case of symmetric preferences.

Appendices

A Numerical methods

A.1 Advancing the solution from $t_{m+1}^- \to t_m^+$

This section elaborates further on the description of the numerical solution in Section 4.1 which describes the solution of the relevant PDEs that hold between decision dates.

Since we solve the PDEs backwards in time, it is convenient to define $\tau = T - t$ and use the definition

$$
\begin{aligned}
\hat{V}_p(e_1, e_2, x_i, s, \tau) &= V_p(e_1, e_2, x_i, s, T - \tau) \\
\hat{\pi}_p(e_1, e_2, x_i, s, \tau) &= \pi_p(e_1, e_2, x_i, s, T - \tau) \ .
\end{aligned} \tag{29}
$$

We rewrite Equation ((11)) in terms of backwards time $\tau = T - t$

$$
\begin{aligned}
\frac{\partial \hat{V}_p}{\partial \tau} &= \hat{\mathcal{L}}\hat{V}_p + \hat{\pi}_p + [(e_1 + e_2) + \rho(\bar{S} - s)]\frac{\partial \hat{V}_p}{\partial s} \\
&\qquad \hat{\mathcal{L}}\hat{V}_p \equiv \frac{(\sigma)^2}{2}\frac{\partial^2 \hat{V}_p}{\partial x^2} + \eta(\bar{X} - x)\frac{\partial \hat{V}_p}{\partial x} - r\hat{V}_p \ .
\end{aligned} \tag{30}
$$

Defining the Lagrangian derivative

$$
\frac{D\hat{V}_p}{D\tau} \equiv \frac{\partial \hat{V}_p}{\partial \tau} + \left(\frac{ds}{d\tau}\right)\frac{\partial \hat{V}_p}{\partial s} \ , \tag{31}
$$

then Equation (30) becomes

$$
\frac{D\hat{V}_p}{D\tau} = \hat{\mathcal{L}}\hat{V}_p + \pi_p \tag{32}
$$

$$
\frac{ds}{d\tau} = -[(e_1 + e_2) + \rho(\bar{S} - s)] \ . \tag{33}
$$

Integrating Equation (33) from τ to $\tau - \Delta\tau$ gives

$$
s_{\tau - \Delta\tau} = s_\tau \exp(-\rho\Delta\tau) + \bar{S}(1 - \exp(-\rho\Delta\tau)) + \big(\frac{e_1 + e_2}{\rho}\big)(1 - \exp(-\rho\Delta\tau)) \ . \tag{34}
$$

We now use a semi-Lagrangian timestepping method to discretize Equation (30) in backwards time τ. We use a fully implicit method as described in Chen & Forsyth (2007).

$$
\begin{aligned}
\hat{V}_p(e_1, e_2, x, s_\tau, \tau) &= (\Delta\tau)\hat{\mathcal{L}}\hat{V}_p(e_1, e_2, x, s_\tau, \tau) \\
&\qquad + (\Delta\tau)\pi_p(e_1, e_2, x, s_\tau, \tau) + \hat{V}_p(e_1, e_2, x, s_{\tau - \Delta\tau}, \tau - \Delta\tau) \ .
\end{aligned} \tag{35}
$$

Equation (35) now represents a set of decoupled one-dimensional PDEs in the variable x, with (e_1, e_2, s) as parameters. We use a finite difference method with forward, backward, central differencing to discretize the $\hat{\mathcal{L}}$ operator, to ensure a positive coefficient method. (See Forsyth & Labahn (2007) for details.) Linear interpolation is used to determine $\hat{V}_p(e_1, e_2, x, s_{\tau-\Delta\tau}, \tau - \Delta\tau)$. We discretize in the x direction using an unequally spaced grid with n_x nodes and in the S direction using n_s nodes. Between the time interval t^-_{m+1}, t^+_m we use n_τ equally spaced time steps. We use a coarse grid with $(n_\tau, n_x, n_s) = (2, 27, 21)$. We repeated the computations with a fine grid doubling the number of nodes in each direction to verify that the results are sufficiently accurate for our purposes.

A.2 Advancing the solution from $t^+_m \to t^-_m$

This section elaborates on the solution of the game at fixed decision dates as described in Section 4.2.1.

We model the possible emission levels as ten discrete states for each of e_1, e_2, which gives 100 possible combinations of (e_1, e_2). We then determine the optimal controls using the methods described in Section 4.2.1. We use exhaustive search (among the finite number of possible states for (e_1, e_2)) to determine the optimal policies. This is, of course, guaranteed to obtain the optimal solution.

B Nash Equilibrium

Section 4.2.1 describes choice of controls for the Stackelberg game and the social planner. In this appendix we describe how we determine whether a particular choice of controls at a given decision time is a Nash Equilibrium.

We again fix (e_1, e_2, s, x, t_m), so that we understand that $e^+_p = e^+_p(e_1, e_2, s, x, t_m)$, $R_p(\omega) = R_p(\omega, e_1, e_2.s, x, t_m)$.

Definition 4 (Nash Equilibrium)**.** *Given the best response sets $R_2(\omega_1)$, $R_1(\omega_2)$ defined in*

Equations (13)-(14), then the pair (e_1^+, e_2^+) *is a Nash equilibrium point if and only if*

$$e_1^+ = R_1(e_2^+) \quad ; \quad e_2^+ = R_2(e_1^+) \ . \tag{36}$$

From Definition 3 of a Stackelberg game, if player 1 goes first, we have the optimal pair $(\hat{e}_1^+, \hat{e}_2^+)$

$$\begin{aligned} \hat{e}_1^+ &= \operatorname*{argmax}_{\omega_1' \in Z_1} V_1(\omega_1', R_2(\omega_1'), s, x, t_m^+) \ , \\ \hat{e}_2^+ &= R_2(\hat{e}_1^+) \ . \end{aligned} \tag{37}$$

Similarly, we have the pair $(\bar{e}_1^+, \bar{e}_2^+)$ if player 2 goes first

$$\begin{aligned} \bar{e}_2^+ &= \operatorname*{argmax}_{\omega_2' \in Z_2} V_2(R_1(\omega_2'), \omega_2', s, x, t_m^+) \ , \\ \bar{e}_1^+ &= R_1(\bar{e}_2^+) \ . \end{aligned} \tag{38}$$

Suppose $(\hat{e}_1^+, \hat{e}_2^+) = (\bar{e}_1^+, \bar{e}_2^+)$. Consequently, we have $(e_1^+, e_2^+) = (\hat{e}_1^+, \hat{e}_2^+) = (\bar{e}_1^+, \bar{e}_2^+)$ and we replace the $\hat{e}_p^+$ by e_p^+ and $\bar{e_p}^+$ by ${e_p}^+$ in Equations (37) - (38) giving

$$\begin{aligned} e_1^+ &= \operatorname*{argmax}_{\omega_1' \in Z_1} V_1(\omega_1', R_2(\omega_1'), s, x, t_m^+) \ , \\ e_2^+ &= \operatorname*{argmax}_{\omega_2' \in Z_2} V_2(R_1(\omega_2'), \omega_2', s, x, t_m^+) \ , \\ & e_1^+ = R_1(e_2^+) \ ; \ e_2^+ = R_2(e_1^+) \ , \end{aligned} \tag{39}$$

which is a Nash equilibrium from Definition 4. We can summarize this result in the following

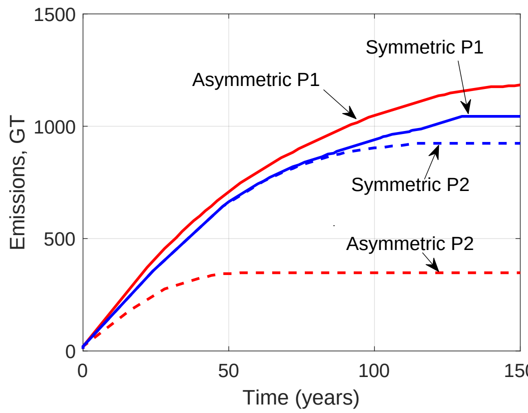
Proposition 1 (Sufficient condition for a Nash Equilibrium)**.** *A Nash equilibrium exists at a point* (e_1, e_2, s, x, t_m) *if* $(\hat{e}_1^+, \hat{e}_2^+) = (\bar{e}_1^+, \bar{e}_2^+)$.

Remark 5 (Checking for a Nash equilibrium)**.** *A necessary and sufficient condition for a Nash Equilibrium is given by condition* (36)*. However a sufficient condition for a Nash equilibrium in the Stackelberg game is that the optimal control of either player is independent* *of who goes first.*

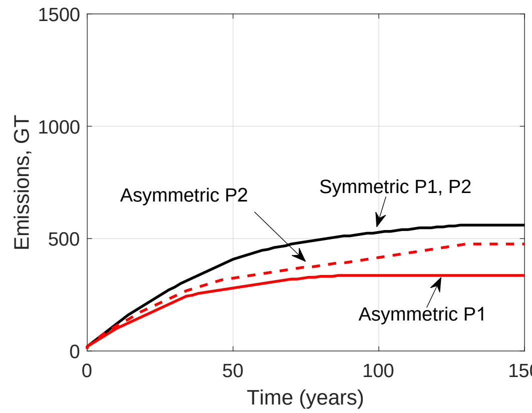
In our numerical experiments we find Nash equilibria exist only at some points (not all) over the state space. This is, of course, not surprising since the system of PDEs is degenerate. Insley & Forsyth (2019) examine this issue, along with other possible games, such as leader-leader, follower-follower games, and interleaved games.

C Additional figures displaying results

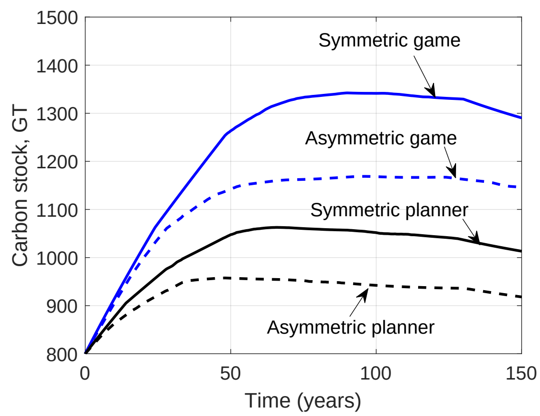
This section displays figures that depict numerical results described in Section 6.



(a) Game cumulative emissions medians, asymmetric and symmetric

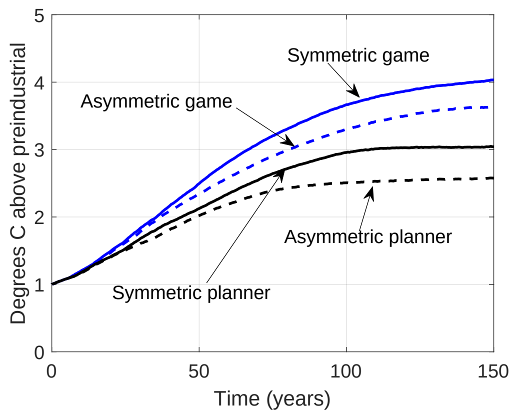


(b) Planner cumulative emissions medians, asymmetric and symmetric

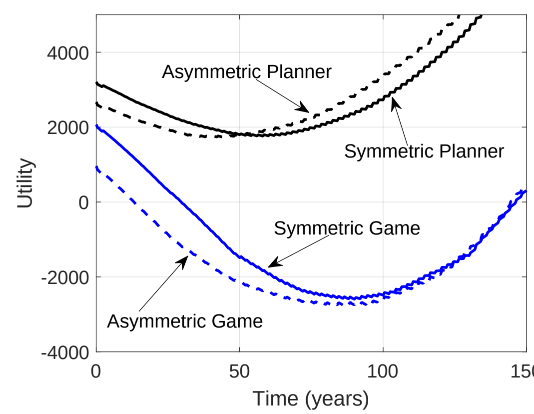


(c) Carbon stock medians, Asymmetric and Symmetric, game and planner

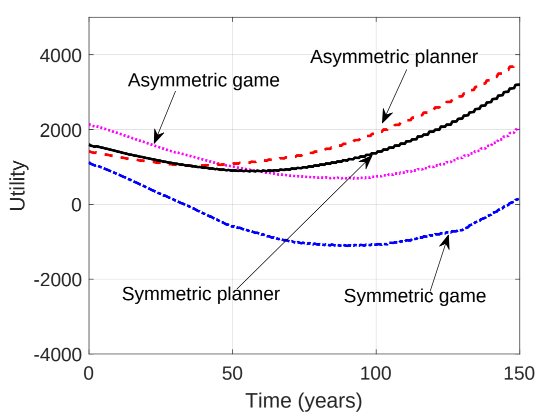
Figure 15: Asymmetric and Symmetric Damages: Cumulative player emissions and carbon stock, median values over time. $X(0) = 1$, $S(0) = 800$, , $E_1(0) = E_2(0) = 10$. 10,000 simulations.



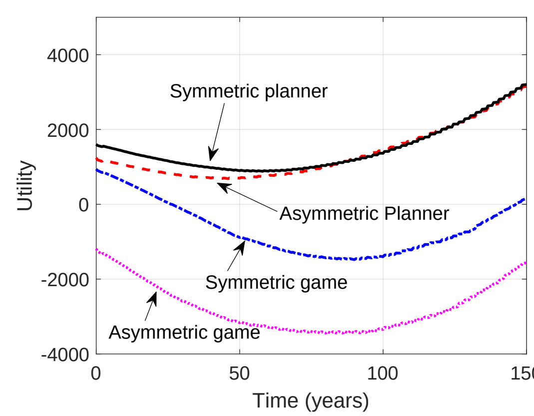
(a) Median temperature paths, asymmetric and symmetric



(b) Total utility medians: asymmetric and symmetric

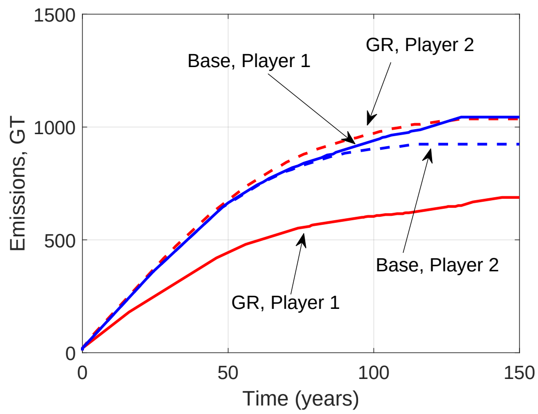


(c) Player 1 utility medians, asymmetric and symmetric, game and planner

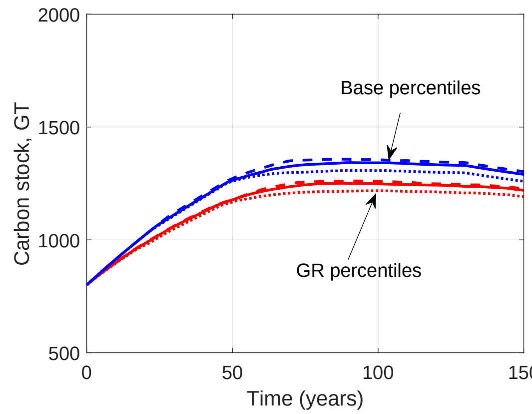


(d) Player 2 utility medians, asymmetric and aymmetric, game and planner

Figure 16: Asymmetric and Symmetric Damages: Temperature and utility, median values over time. $X(0) = 1$, $S(0) = 800$, , $E_1(0) = E_2(0) = 10$. 10,000 simulations.

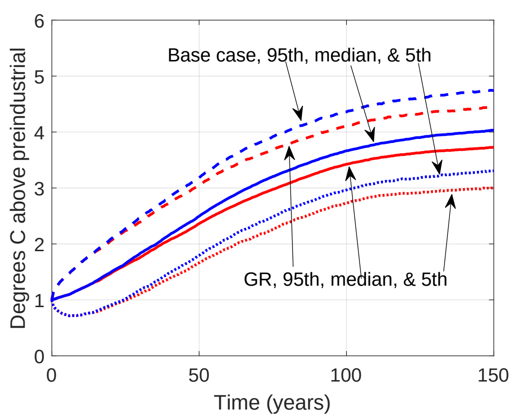


(a) Cumulative median player emissions, Base Game and Green Reward

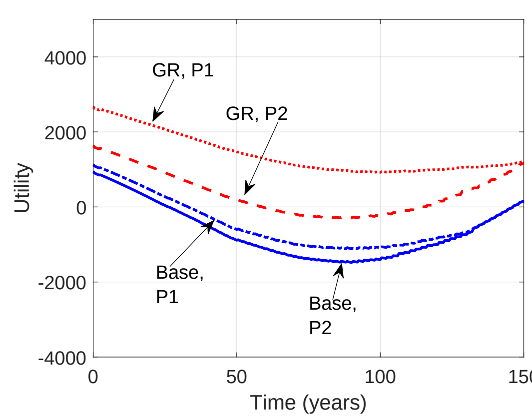


(b) Carbon stock percentiles, Base Game and Green Reward

Figure 17: Green Reward: Cumulative player median emissions and carbon stock percentiles for base case game and green reward. $X(0) = 1$, $S(0) = 800$, , $E_1(0) = E_2(0) = 10$. 10,000 simulations. In right hand figure, solid lines are medians, dashed are 95th percentiles and dotted are 5th percentiles.



(a) Temperature percentiles, Base Game and Green Reward



(b) Median Player Utility (V), Base Game and Green Reward

Figure 18: Green reward: Temperature percentiles median player utilities for the base case game and green reward. Utility refers to V_p defined in Equation (8) for player p. $X(0) = 1$, $S(0) = 800$, , $E_1(0) = E_2(0) = 10$. 10,000 simulations.

D Sensitivity to the admissible set for emissions

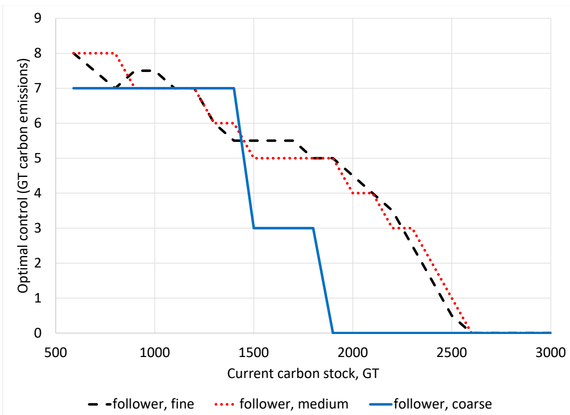
Decision makers in the model represent two large nations or groups of nations with a significant impact on total world emissions. We have assumed optimal emissions are chosen from a discrete set of possibilities, which we believe is a more logical assumption than assuming policy makers choose from a continuous set. The latter assumption seems to imply that decision makers have a greater ability to fine tune policy choices than is likely to hold in reality. However this naturally raises the question of what is the impact of varying the admissible choice set. We explored this question through several sensitivities. In particular we compared results between three different admissible sets.

$$\begin{aligned} \text{Coarse:} \quad & Z_p = \{0, 3, 7, 10\} \\ \text{Medium:} \quad & Z_p = \{0, 1, 2, ..., 9, 10\} \\ \text{Fine:} \quad & Z_p = \{0, 0.5, 1, ..., 9.5, 10\}, \quad p = 1, 2 \end{aligned}$$

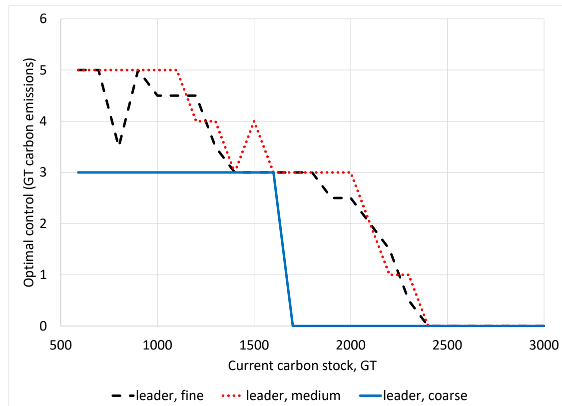
The Medium admissible set is the one adopted in the current paper. The Coarse admissible set is used to explore different games in Insley & Forsyth (2019).

For the base case game we find that expected utility is quite close whether we use the coarse or medium admissible sets. A larger difference in results emerges for cases with asymmetric players. In this appendix we present the results for the Green Reward case in which the leader gets positive utility from reducing emissions due a more the more environmentally friendly preferences of its citizens. Figure 19 compares optimal controls for the three admissible sets. We observe the greatest difference between the coarse versus the other cases. With the coarse admissible set, the optimal choice of emissions is lower for most values of the carbon stock compared to the admissible sets with finer grids. As noted in the main text, the point where the optimal controls jump down and then back up again are points where the value function is quite flat, so that there is little difference in value over the particular range of emissions where these "blips" occur. The different choices for emissions in the Coarse admissible set, translate into lower utility levels as is shown in Figure 20. Utility for the medium and fine admissible sets are very close.

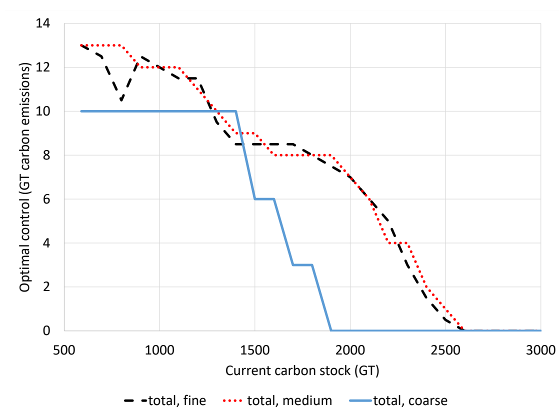
We conclude that the admissible set for the optimal control does affect the strategic interactions of players, particularly when players are asymmetric. We have argued that a discrete set of choices is a more realistic representation of available policy choices. However, the further refinement of the admissible set from the Medium to the Fine case does not make a large difference for the analysis in this paper.



(a) Follower optimal emissions

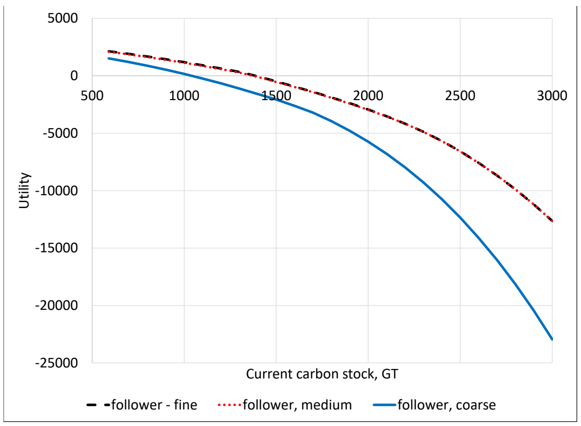


(b) Leader optimal emissions

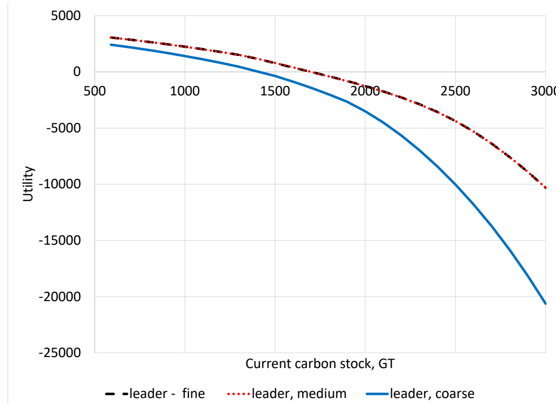


(c) Total emissions

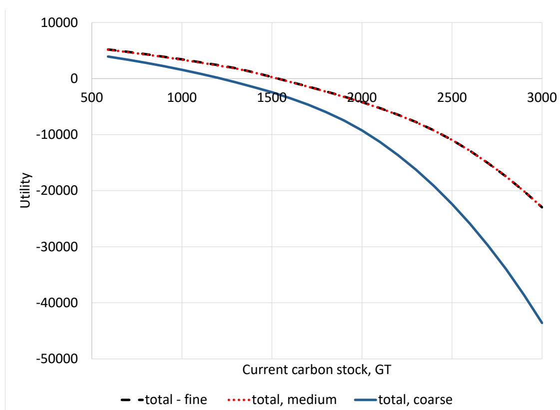
Figure 19: Comparing coarse, medium and fine admissible sets for optimal emissions: Optimal controls versus carbon stock in GT, carbon stock = 800 GT, Base case is labeled as medium.



(a) Follower utility



(b) Leader utility



(c) Total utility

Figure 20: Comparing coarse, medium and fine admissible sets for optimal emissions: versus carbon stock in GT. Utility refers to V_p defined in Equation (8) for player p. Temperature $= 1$ °C, Base case is labeled as medium.

References

Ackerman, F., Stanton, E. A. & Bueno, R. (2013), 'Epstein-Zin Utility in DICE: Is Risk Aversion Irrelevant to Climate Policy?', *Environmental and Resource Economics* **56**(1), 73–84.

Amarala, S. (2015), Monotone numerical methods for nonlinear systems and second order partial differential equations, PhD thesis, University of Waterloo, Waterloo, Ontario, Canada.

Barcena-Ruiz, J. C. (2006), 'Environmental taxes and first-mover advantages', *Environmental and Resource Economics* **35**, 19–39.

Barles, G. & Souganidis, P. (1991), 'Convergence of approximation schemes for fully nonliner second order equations', *Asymptotic Analysis* **4**, 271–283.

Bednar-Friedl, B. (2012), 'Climate policy targets in emerging and industrialized economies: the influence of technological differences, environmental preferences and propensity to save', *Empirica* **39**, 191–215.

Benchekroun, H. & Chaudhuri, A. R. (2014), 'Transboundary pollution and clean technologies', *Resource and Energy Economics* **36**(2), 601–619.

Benchekroun, H. & van Long, N. (1998), 'Efficiency inducing taxation for polluting oligopolists', *Journal of Public Economics* **70**(2), 325–342.

Bjork, T. (2009), *Arbitrage Theory in Continuous Time*, Oxford University Press.

Bressan, A. (2011), 'Noncooperative Differential Games', *Milan Journal of Mathematics* **79**(2), 357–427.

Bressan, A. & Shen, W. (2004), 'Semi-cooperative strategies for differential games', *International Journal of Game Theory* **32**(4), 561–593.

Cacace, S., Cristiani, E. & Falcone, M. (2013), Numerical approximation of Nash equilibria for a class of non-cooperative differential games, *in* L. Petrosjan & V. Mazalov, eds, 'Game Theory and Applications', Vol. 16, Nova Science Publishers.

Chen, Z. & Forsyth, P. (2007), 'A semi-Lagrangian approach for natural gas storage valuation and optimal operation', *SIAM Journal on Scientific Computing* **30**, 339–368.

Chesney, M., Lasserre, P. & Troja, B. (2017), 'Mitigating global warming: a real options approach', *Annals of operations research* **255**(1-2), 465–506.

Clean Energy Canada (2015), How to adopt a winning carbon price. Initiative of the Centre for Dialogue, Simon Fraser University, Vancouver Canada; Retrieved November 27, 2015 at `http://cleanenergycanada.org/wp-content/uploads/2015/02/Clean-Energy-Canada-How-to-Adopt-a-Winning-Carbon-Price-2015.pdf`.

Colombo, L. & Labrecciosa, P. (2019), 'Stackelberg versus Cournot: A differential game approach', *Journal of Economic Dynamics & Control* **101**, 239–261.

Crost, B. & Traeger, C. P. (2014), 'Optimal CO2 mitigation under damage risk valuation', *Nature Climate Change* **4**(7), 631–636.

Dixit, A. & Pindyck, R. (1994), *Investment Under Uncertainty*, Princeton University Press.

Dockner, E. J., Jorgensen, S., Long, N. V. & Sorger, G. (2000), *Differential games in economics and management science*, Cambridge University Press.

Dockner, E. J. & Long, N. V. (1993), 'International pollution control: Cooperative versus noncooperative strategies', *Journal of Environmental Economics and Management* **25**, 13–29.

Dockner, E., Long, N. V. & Sorger, G. (1996), 'Analysis of Nash equilibria in a class of capital accumulation games', *Journal of Economic Dynamics & Control* **20**(6-7), 1209–1235.

Forsyth, P. & Labahn, G. (2007), 'Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance', *Journal of Computational Finance* **11**(2), 1–44.

Golosov, M., Hassler, J., Krusell, P. & Tsyvinski, A. (2014), 'Optimal taxes on fossil fuel in general equilibrium', *Econometrica* **82**(1), 41–88.

Hambel, C., Kraft, H. & Schwartz, E. (2017), Optimal carbon abatement in a stochastic equilibrium model with climate change, Technical report. NBER Working Paper No. 21044.

Harris, C., Howison, S. & Sircar, R. (2010), 'Games with exhaustible resourses', *SIAM Journal of Applied Mathematics* **70**(7), 2556–2581.

Insley, M. & Forsyth, P. (2019), 'Climate Games: Who's on first? What's on second?', *L'Actualité économique* **95**.

Jorgensen, S., Martin-Herran, G. & Zaccour, G. (2010), 'Dynamic Games in the Economics and Management of Pollution', *Environmental Modeling & Assessment* **15**(6), 433–467.

Kelly, D. & Kolstad, C. (1999), 'Bayesian learning, growth, and pollution', *Journal of Economic Dynamics & control* **23**(4), 491–518.

Kossey, A., Peszko, G., Oppermann, K., Prytz, N., Klein, N., Blok, K., Lam, L., Wong, L. & Borkent, B. (2015), 'State and trends of carbon pricing 2015'. retrieved from `http://documents.worldbank.org/curated/en/636161467995665933/State-and-trends-of-carbon-pricing-2015`.

Leach, A. (2007), 'The climate change learning curve', *Journal of Economic Dynamics and Control* **31**, 1728–1752.

Ledvina, A. & Sircar, R. (2011), 'Dynamic Bertrand oligopoly', *Applied Mathematics and Optimization* **63**(1), 11–44.

Lemoine, D. & Traeger, C. (2014), 'Watch your step: optimal policy in a tipping climate', *American Economic Journal: Economic Policy* **6**(2), 137–166.

List, J. A. & Mason, C. F. (2001), ‘Optimal institutional arrangements for transboundary pollutants in a second-best world: Evidence from a differential game with asymmetric players’, *Journal of Environmental Economics and Management* **42**, 277–296.

Long, N. V. (2010), *A Survey of Dynamic Games in Economics*, World Scientific Publishing Company.

Long, N. V. (2011), ‘Dynamic Games in the Economics of Natural Resources: A Survey’, *Dynamic Games and Applications* **1**(1), 115–148.

Ludkovski, M. & Sircar, R. (2012), ‘Exploration and exhaustibility in dynamic Cournot games’, *European Journal of Applied Mathematics* **23**(3), 343–372.

Ludkovski, M. & Sircar, R. (2015), Game theoretic models for energy production, *in* R. A´id, M. Ludkovski & R. Sircar, eds, ‘Commodities, Energy and Environmental Finance’, Springer, Berlin.

Ludkovski, M. & Yang, X. (2015), Dynamic cournot models for production of exhaustible commodities under stochastic demand, *in* R. A´id, M. Ludkovski & R. Sircar, eds, ‘Commodities, Energy and Environmental Finance’, Springer.

Nkuiya, B. (2015), ‘Transboundary pollution game with potential shift in damages’, *Journal of Environmental Economics and Management* **72**, 1–14.

Nordhaus, W. (2013), Integrated economic and climate modeling, *in* P. B. Dixon & D. W. Jorgenson, eds, ‘Handbook of Computable General Equilibrium Modeling, First Edition’, Vol. 1, Elsevier, chapter 16, pp. 1069–1131.

Nordhaus, W. & Sztorc, P. (2013), Dice 2013r: Introduction and user’s manual, Technical report.

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

Reisinger, C. & Forsyth, P. A. (2016), 'Piecewise constant policy approximations to Hamilton-Jacobi-Bellman Equations', *Applied Numerical Mathematics* **103**, 27–47.

Salo, S. & Tahvonen, O. (2001), 'Oligopoly equilibria in nonrenewable resource markets', *Jounal of Economic Dynamics & Control* **25**(5), 671–702.

Traeger, C. (2014), 'A 4-stated dice: Quantitatively addressing uncertainty effects in climate change', *Environmental and Resource Economics* **59**(2), 1–37.

Urpelainen, J. (2009), 'Explaining the Schwarzenegger phenomenon: Local frontrunners in climate policy', *Global Environmental Politics* **9**, 82–105.

van der Ploeg, F. (1987), 'Inefficiency of credible strategies in oligopolistic resource markets with uncertainty', *Jounal of Economic Dynamics & Control* **11**(1), 123–145.

Weitzman, M. L. (2012), 'GHG targets as insurance against catastrophic climate damages', *Journal of Public Economic Theory* **14**, 221–244.

Williams, R. C. (2012), 'Growing state-federal conflicts in environmental policy: The role of market-based regulation', *Journal of Public Economics* **96**, 1092–1099.

Wirl, F. (2008), 'Tragedy of the commons in a stochastic game of a stock externality', *Journal of Public Economic Theory* **10**(1), 99–124.

Wirl, F. (2011), 'Global warming with green and brown consumers', *Scandinavian Journal of Economics* **113**(4, SI), 866–884.

Xepapadeas, A. (1998), 'Policy adoption rules and global warming - theoretical and empirical considerations', *Enviromental & Resource Economics* **11**(3-4), 635–646. 1st World Congress of Environmental and Resource Economists, Venice, Italy, June 25-27, 1998.

Zagonari, F. (1998), 'International pollution problems: Unilateral initiatives by environmental groups in one country', *Journal of Environmental Economics and Management* **36**(1), 46–69.