Provision of a public good with multiple dynasties*

Larry Karp†

First version: December 2011. This version: March 2016

Abstract

Because carbon emissions create externalities across countries and generations, climate policy requires international cooperation and intergenerational altruism. A differential game using overlapping generations with intergenerational altruism shows how altruism and cooperation interact, and provides estimates of their relative importance in determining equilibrium steady state carbon levels. A small increase in cooperation has a larger equilibrium effect than a small increase in altruism, beginning at empirically plausible levels. A large increase in altruism may have a larger equilibrium effect, compared to a large increase in cooperation. Climate investments may be dynamic strategic complements, reducing but not eliminating incentives to free ride.

Keywords: Overlapping generations, altruism, time consistency, Markov Perfection, differential games, climate policy

JEL, classification numbers: C73, D62, D63, D64, H41, Q54

---

*An earlier version of this paper was entitled "Provision of a public good with altruistic overlapping generations and many tribes". The various versions benefitted from comments by Geir Asheim, Ivar Ekeland, Reyer Gerlagh, Bard Harstad, Antii Iho, Terry Iverson, Jacek Krawczyk, Derek Lemoine, Armon Rezai, John Roemer, Leo Simon, Christian Traeger, seminar participants at Berkeley, Santa Barbara, Paris, Vienna, SURED 2012, and the TIGER conference, Toulouse 2013, two anonymous referees and an Editor (who suggested a way of simplifying the proof of Proposition 1). The usual disclaimer holds.

†Department of Agricultural and Resource Economics, University of California, Berkeley, email: karp@berkeley.edu

1 Introduction

The reduction of greenhouse gas (GHG) emissions is a global public good whose payoff may occur in the distant future. Nations, or coalitions of nations, choose their climate policy primarily with a view to their constituents' welfare. These constituents care about their own current and future utility flows, and about their successors. At a point in time, the social planner in each nation (or coalition) aggregates their constituents' preferences and chooses current policy. Two factors impede meaningful climate policy: people have limited altruism with respect to their successors, and they have limited ability to cooperate with their contemporaries.

Distinct literatures examine separately the effect on climate policy of international cooperation or attitudes toward future generations. By including both of these features in a tractable model, I am able to study the interaction between them, and to illustrate how each influences equilibrium outcomes. Holding fixed either international cooperation or intergenerational altruism, how does an increase in the other characteristic alter the equilibrium?

I describe the contours of the model and then summarize the findings. I view climate policy as the equilibrium of a game, not (except in limiting cases) the solution to an optimization problem. The world consists of a fixed population, divided into n symmetric coalitions. At each point in time, the decisionmaker in each coalition chooses a level of emissions or a carbon tax to maximize her constituents' welfare, ignoring welfare in other coalitions. Each coalition contains $\frac{1}{n}$'th of the world and therefore internalizes $\frac{1}{n}$'th of their effect on the climate. A decrease in n increases internalization across countries, and represents an increase in international cooperation; $n = 1$ maximizes international cooperation.

Each coalition contains many generations; over time, people die and new generations replace them. People care about their own current and future flow of utility, which they discount at the constant pure rate of time preference; this parameter measures agents' impatience for their own future utility. People discount the utility (or welfare) of unborn generations in their coalition at a constant rate, an inverse measure of intergenerational altruism. No one cares about people in other coalitions.

A costly current emissions reduction (abatement) may benefit people currently alive late in their life, but future generations likely obtain most of the benefit. Abatement therefore involves transfers from a person to her future self, and to people who have not yet been born. An altruism parameter that differs from the pure rate of time preference disentangles the welfare effect of these two types of transfers, and results in time inconsistent preferences for the planner who aggregates the preferences of currently living coalition citizens.

The parameters measuring international cooperation and intergenerational altruism correspond to mutable features of the real world. For example, the European countries' delegating their climate policy to the European Union, or developing nations following the lead of Brazil-Russia-India-China, encourage countries to replace national interest with the interest of a larger bloc. These moves correspond to a decrease in n, i.e. an increase in international cooperation. Components of the Kyoto Agreement (e.g. the Clean Development Mechanism and Joint Implementation) and of the Copenhagen Accord (e.g. the global funding scheme to finance adaptation to climate change) can be construed as attempts to increase international cooperation, and thus correspond to a reduction in n.

Economists agree that discounting is important, but disagree about how discounting should be used to formulate climate policy (Stern, 2006; Nordhaus, 2007; Weitzman, 2007; Roemer, 2011; Arrow et al, 2013; Drupp et al, 2014). The disagreement may arise from differences in preferences, and possibly also from a conflation (inherent in the infinitely lived agent model) of intra- and intergenerational transfers.[1] Many preference characteristics (e.g., associated with racism, sexism, and homophobia) have changed over time, at least partly as a consequence of efforts to change them. Educating people about the potential long run effects of carbon emissions might alter their views on intergenerational altruism.

I examine the relative importance of altruism and cooperation on the equilibrium steady state atmospheric carbon stock (equivalently, the carbon tax that supports

[1] France, the UK and the US use lower social discount rates to evaluate climate policy. Changing views about the proper way to take into account the welfare of distant generations may have contributed to the use of a lower long run discount rate.

this stock). This comparison uses altruism and cooperation indices that range from 0 to 1. At empirically plausible levels, a small increase in cooperation has a much larger equilibrium effect, compared to a small increase in altruism. Beginning at empirically plausible levels, a large increase in altruism can have a much greater effect on the steady state, compared to a large increase in cooperation. I also find that an increase in altruism (respectively, cooperation) has a larger effect on the steady state when cooperation (respectively, altruism) is low. In addition, climate investments may be dynamic strategic complements, thus reducing incentives to free ride.

This paper bridges two large literatures. The first category consists of models (some cited above) that calculate optimal climate policy for an infinitely lived agent, or equivalently for finitely lived agents whose altruism parameter equals their pure rate of time preference. The second category, using game theory to examine equilibrium policy, underpins suggestions for designing an effective climate agreement (Aldy and Stavins, 2007, Guesnerie and Tulkens, 2008). One strand studies coalition formation (e.g. Barrett 2005); with few exceptions (e.g. Harstad 2012), most papers in this strand use static models. A second strand uses differential games, in which infinitely lived agents have a constant discount rate, to model international externalities (van der Ploeg and de Zeeuw, 1992; Wirl, 1994; Long, 2010; Haurie, Krawczyk and Zaccour, 2012).[2]

Section 2 describes the overlapping generations (OLG) model, and explains how altruism affects the discount rate of the planner who aggregates the preferences of agents alive in a coalition at a point in time. Section 3 describes the game and the equilibrium conditions. The model nests a single-agent problem of non-

[2] Calvo and Obstfeld, 1988 and Schneider, Traeger and Winkler, 2012 study OLG models in which agents discount their own and their successors utility at different rates. However, the social planner discounts the old generations' future utility from the time of their birth, not the current time, giving older people less weight in evaluating current policy, eliminating the time inconsistency. Many papers use OLG models to study environmental and resource problems (Kemp and Long 1979; John, Pecchenino and Schhimmelfennig, 1995; Kosekla, Ollikainen and Puhakka, 2002), and a growing number use OLG models to study climate policy (Howarth, 1998; Gerlagh and van der Zwaan, 2001; Rasmussen, 2003; Laurent-Lucchetti and Leach, 2011). Those papers do not include the strategic elements that arise with non-constant discounting, which is central to my paper.

constant discounting (Strotz, 1956; Laibson, 1997), a differential game with constant discounting, and a standard optimal control problem. Section 4 describes the climate model and obtains the results summarized above.

2 Discounting

This section describes the model of time preferences, altruism, and dynasties. A social planner who aggregates her constituents' preferences has a non-constant utility discount rate, a function of agents' impatience, altruism, and longevity. Ekeland and Lazrak (2010) obtain this discount rate for paternalistically altruistic agents; I extend their result by also considering purely altruistic agents. Saez-Marti and Weibull (2005) establish an isomorphism between paternalistic and pure altruism for a general sequence of pure rates of time preference, in a setting where a sequence of agents each lives a single period. I identify a different isomorphism in an OLG setting, for a less general model of discounting. There is no coalition index in this section, because I consider a representative coalition here.

Agents' lifetime is exponentially distributed with mortality rate θ, giving expected lifetime $\frac{1}{\theta}$; with constant population, the birth rate is also θ. Due to the exponential distribution's memoryless property, currently living agents' random times of death do not depend on their current ages. Agents have the pure rate of time preference r, so their risk-adjusted pure rate of time preference is $r+\theta$. For a utility stream $\{u_\tau\}_{\tau=0}^{\infty}$ with utility flow $u_\tau < \infty$, the expected present discounted value of lifetime utility for an agent alive at time t is $U(t) = \int_{\tau=t}^{\infty} e^{-(r+\theta)(\tau-t)} u(\tau)\, d\tau$; this integral is the "selfish" component of the agent's welfare.

An agent at t with paternalistic altruism cares about the lifetime utility of her successors, all those born at $s > t$. But she does not take into account the fact that those born at s', with $s > s' > t$, also care about the agents born at time s. In contrast, an agent with pure altruism does take into account the fact that her successors care about their own successors' welfare, not just their utility streams. The paternalistic agent discounts her successors' utility at rate λ, and the agent with pure altruism discounts her successor' welfare at rate $\tilde{\lambda}$.

To obtain a tractable model, I adopt

Assumption 1 *(i) All agents have the same utility function, which depends only on the global public good and the agent's investment in the public good. (ii) In each period, agents in a coalition share equally their coalition's cost of investment in the public good. (iii) Agents might care about current and future members of their own coalition, but they do not care about the citizens of other coalitions.*

Assumption 1.i means that there are no privately owned assets. Due to the exponential distribution and Assumption 1.i, any two currently living citizens of a coalition are identical, rendering Assumption 1.ii innocuous. Assumption 1.i&ii imply that any currently living coalition member can be chosen as the social planner who decides current (but not future) investment levels. Assumption 1.iii makes it possible to consider discounting within a coalition independently of events in other coalitions.

The welfare of an agent with paternalistic altruism and utility stream $\{u_\tau\}_{\tau=0}^{\infty}$, is

$$W(t) \equiv U(t) + \theta \int_t^\infty e^{-\lambda(\tau - t)} U(\tau)\, d\tau = \int_{\tau=t}^\infty D(\tau - t)\, u(\tau)\, d\tau. \qquad (1)$$

Her welfare consists of two components: her own lifetime utility (the "selfish" component, $U(t)$) and an altruistic component. Over the interval of time $(\tau, \tau + d\tau)$, approximately $\theta d\tau$ new agents are born, accounting for the θ in front of the first integral in equation 1. Each future agent has her own lifetime utility $U(\tau)$, which the agent at t discounts at rate λ. The equality implicitly defines the discount factor, $D(\tau - t)$, under paternalistic altruism. Using the definition of $U(\tau)$ in equation 1 and simplifying by changing the order of integration, produces Ekeland and Lazrak's (2010) discount factor under paternalistic altruism, $D^{EL}(t)$:[3]

$$D^{EL}(t) = \left(\frac{\lambda - r}{\lambda - (r+\theta)}\right) e^{-(r+\theta)t} - \frac{\theta}{\lambda - (r+\theta)} e^{-\lambda t}. \qquad (2)$$

[3]Several papers use a convex combination of exponentials to represent non-constant discounting for a single infinitely lived agent (Li and Lofgren, 2000; Gollier and Weitzman, 2010; Zuber, 2010; and Jackson and Yariv, 2015). In Ekeland and Lazrak's OLG model, the discount factor is a weighted combination of exponentials; it is a convex combination only if $\lambda \leq r$.

If $\theta = \infty$, a coalition consists of a succession of agents, each of whom lives for a single instant, implying a constant social discount rate λ. At the other extreme, $\theta = 0$, a coalition consists of an infinitely lived agent, with a constant discount rate r. For these two limiting cases, there is no time consistency problem.

Given the utility stream $\{u_{t+\tau}\}_{\tau=0}^{\infty}$, welfare at t for the agent with pure altruism, $V(t)$, satisfies the recursion

$$V(t) \equiv U(t) + \theta \int_t^\infty e^{-\tilde{\lambda}(\tau-t)} V(\tau) d\tau = \int_t^\infty D^{pure}(\tau - t)\, u(\tau)\, d\tau. \tag{3}$$

The agent's welfare consists of the discounted stream of her own utility, plus the stream of successors' welfare, discounted using the altruism parameter $\tilde{\lambda}$. The equality implicitly defines the discount factor $D^{pure}(\tau - t)$.

The two discount factors are related in a simple way:

Proposition 1 *Agents have mortality rate θ and pure rate of time preference r; agents with pure altruism discount future agents' welfare at rate $\tilde{\lambda} > \theta$, and agents with paternalistic altruism discount future agents' utility at rate $\lambda > 0$. (i) The two types of agents, and thus the planners who represent them, have the same preferences if and only if $\tilde{\lambda} = \lambda + \theta$. (ii) If $\tilde{\lambda} < \lambda + \theta$, the planner under paternalistic altruism discounts the future flow of utility more heavily than the planner with pure altruism.*

In view of the isomorphism described in Proposition 1.i, I hereafter consider only the case of paternalistic altruism, and drop the superscript on the discount factor. Using $\gamma \equiv r + \theta$ to denote the risk-adjusted pure rate of time preference, the discount rate, $\eta(t)$, corresponding to equation 2, is

$$\eta(t) \equiv -\frac{dD}{dt}\frac{1}{D} = \frac{-\gamma\lambda + \gamma r + \theta\lambda e^{-(\lambda-\gamma)t}}{-\lambda + r + \theta e^{-(\lambda-\gamma)t}}, \tag{4}$$

with

$$\begin{gathered} \text{sign}\tfrac{d\eta(t)}{dt} = \text{sign}\ \lambda - r; \quad \eta(0) = r \text{ for } \lambda < \infty; \\ \lim_{t\to\infty} \eta(t) = \lambda \text{ for } \lambda \leq r; \quad \text{and } \lim_{t\to\infty} \eta(t) = \gamma \text{ for } \lambda > r. \end{gathered} \tag{5}$$

Constant discounting corresponds to $\lambda = r$; hyperbolic discounting (a declining discount rate) corresponds to $\lambda < r$; $\lambda > r$ means that the discount rate used to evaluate future utility increases with distance.

It is important to agree on the meaning, but not on the "correct" value, of λ. For $\lambda = \infty$, currently living agents do not care about those born in the future. For $\lambda = r$, people make no distinction between a utility exchange from a person to her older self, and from a person to a different person born in the future.

For $\lambda = 0$, people put the same value on the lifetime expected utility stream, $U(t)$ of all agents, regardless of their date of birth. "Brute luck" is the outcome of an involuntary and uninsurable lottery. The school of "luck egalitarians" claims that it is morally wrong to disadvantage others as a consequence of brute luck (Roemer, 2009). Because the date of a person's birth is a matter of brute luck, this school regards $\lambda = 0$ as the ethical choice. With $\lambda = 0$, the weight put on the lifetime expected utility stream of a person does not depend on their date of birth.[4]

3 The game

This section describes the game, defines the equilibrium, presents the necessary conditions, and then discusses the generic multiplicity of equilibria. A final subsection considers a particular equilibrium for λ close to 0.

3.1 Description of the game

The vector of state variables at t, common to all coalitions, is $\mathbf{S}_t$. This vector possibly includes average temperatures and carbon stocks in different reservoirs, e.g. the atmosphere and ocean. At time t coalition i takes an action x_{it}, such as a carbon tax or a ceiling on carbon emissions. The vector of these actions for the n coalitions is $\mathbf{x}_t \in \mathbb{R}^n$, with i'th element x_{it}. The evolution of the state variable depends on the state variable and coalitions' actions. Coalition i's flow payoff depends on the state

[4] The integrals in equation 1 fail to converge in general, if $\lambda = 0$. Adding a small positive constant to λ to take into account that our species might vanish, accommodates the case $\lambda = 0$.

variable and coalition i's actions (Assumption 1).

The equation of motion for the state variable, the utility flow for coalition i, and the payoff for the planner in coalition i at time t are, respectively:

$$\begin{array}{c} \text{constraint: } \frac{d\mathbf{S}}{dt} = f\left(\mathbf{S}_t, \mathbf{x}_t; n\right); \text{ utility: } u_{it} = u\left(\mathbf{S}_t, x_{it}; n\right); \\ \text{and payoff: } \int_t^\infty D(\tau - t) u(\mathbf{S}_\tau, x_{i\tau}; n) d\tau. \end{array} \tag{6}$$

Coalition i's payoff uses the discount function in equation 2.

Section 4 shows how n enters $f\left(\mathbf{S}_t, \mathbf{x}_t; n\right)$ and $u\left(\mathbf{S}_t, x_{it}; n\right)$. A larger n represents greater fragmentation, not a larger population. A change in n alters equilibrium decisions, changing the evolution of the state variable and the flow payoff, without altering the set of feasible paths for the state variable or aggregate utility flows.[5]

3.2 Equilibrium

At time t the state variable, $\mathbf{S}_t$, is predetermined; it is the initial condition for the subgame that begins at t. There are many subgame perfect Nash equilibria to the game defined by this initial condition and the constraint, utility flow, and payoff in equation 6. I consider only stationary symmetric Markov Perfect equilibria (hereafter, "MPE"). In a MPE, agents' actions depend only on the directly payoff-relevant state, here $\mathbf{S}_t$. I denote the mapping from the state variable at t to i's action at t as $x_{it} = \chi\left(\mathbf{S}_t\right)$; χ does not depend explicitly on time or the coalition index, because of the assumption of stationarity and symmetry (over coalitions).

Planner i, t (the planner in coalition i at time t) plays a game involving both present and future planners in all other coalitions. Due to the time inconsistency of preferences arising from the nonconstant discount rate, a subgame perfect equilibrium requires that planner i, t also play strategically with respect to future planners in her coalition, agents i, s for $s > t$. Planner i, t chooses the current action for her coalition, and understands that future coalition actions depend on the future value

[5] If the world consists of N countries, and each coalition controls m countries, then $n = \frac{N}{m}$. A coalition planner internalizes the effect of her action on residents in all m of the countries in her coalition. A smaller n means that there are fewer coalitions: that each internalizes a greater fraction of the effect of its emissions.

of the state variable. The function $\chi(\mathbf{S})$ is a MPE if and only if $x_{it} = \chi(\mathbf{S}_t)$ is the best response, for all feasible $\mathbf{S}$, for planner i, t when all other planners (including future planners in coalition i) use the decision rule $x_{j\tau} = \chi(\mathbf{S}_\tau)$.

Symmetry and stationarity make it straightforward to write the necessary conditions for a MPE. Denote $\mathbf{i}_{n-1} \in \mathbb{R}^{n-1}$ as the vector consisting of 1's, and denote $F(\mathbf{S}, x_i) \equiv f(\mathbf{S}, \mathbf{i}_{n-1}\chi(\mathbf{S}), x_i; n)$. This function is the time derivative of $\mathbf{S}(t)$ when the current value of the state variable is $\mathbf{S}$, all other coalitions use $\chi(\mathbf{S})$, and coalition i uses x_i. When all other coalitions use $\chi(\mathbf{S})$, coalition i's payoff and constraint are

$$\int_\tau^\infty D(t-\tau)u(\mathbf{S}_t, x_{it}; n)dt \quad \text{and} \quad \dot{\mathbf{S}}_t \equiv \frac{d\mathbf{S}_t}{dt} = F(\mathbf{S}, x_i). \tag{7}$$

$F(\mathbf{S}, x_i)$ is a functional, depending on the endogenous $\chi(\mathbf{S})$. Apart from this fact, the game defined by the payoff and constraint in equation 7 is identical to the games studied by Karp (2007) and Ekeland and Lazrak (2010); both papers find the necessary conditions, and Ekeland and Lazrak (2010) establish sufficiency.

The limiting values, as $t \to \infty$, of the discount rate, $\eta(t)$, differ in the two cases corresponding to $\lambda < r$ and $\lambda > r$ (equation 5). The equilibrium conditions also differ in these two cases. I provide details for $0 < \lambda \leq r$ (where $\lim_{t\to\infty} \eta(t) = \lambda$), relegating the other case to Appendix B.2.1. Define $J(\mathbf{S})$ as the equilibrium value of i's payoff (the integral in equation 7), when all other coalitions use the decision rule $\chi(\mathbf{S})$. Denote $\mathbf{S}^*_{t+\tau}(\mathbf{S}_t)$ as the equilibrium value of $\mathbf{S}_{t+\tau}$, the solution to the differential equation in the first line of equation 6, given initial condition $\mathbf{S}_t$, when all players use the equilibrium decision rule $\chi(\mathbf{S})$. The coalition's utility flow on the equilibrium path is $u\left(\mathbf{S}^*_{t+\tau}(\mathbf{S}_t), \chi\left(\mathbf{S}^*_{t+\tau}(\mathbf{S}_t)\right); n\right)$.

Proposition 2 *Suppose that $J(\mathbf{S})$ is differentiable and $0 < \lambda \leq r$. A (symmetric stationary) MPE $\chi(\mathbf{S})$ satisfies the necessary condition to the following "auxiliary" optimal control problem with constant discount rate λ:*

$$J(\mathbf{S}_0) = \max \int_0^\infty e^{-\lambda t}\left(u(\mathbf{S}_t, x_t; n) - K(\mathbf{S}_t)\right)dt \quad \textit{subject to } \dot{\mathbf{S}} = F(\mathbf{S}, x), \tag{8}$$

with the side condition (a definition):

$$K(\mathbf{S}_t) \equiv (r-\lambda) \int_0^\infty e^{-\gamma\tau} u\left(\mathbf{S}^*_{t+\tau}(\mathbf{S}_t), \chi\left(\mathbf{S}^*_{t+\tau}(\mathbf{S}_t)\right); n\right) d\tau. \tag{9}$$

The integral in equation 9 equals, in equilibrium, the function previously defined as $U(t)$, the "selfish component" of welfare. The quantity $r-\lambda$ can be interpreted as an "altruism weight", with a limiting value r corresponding to $\lambda = 0$. The function $K(\mathbf{S}_t)$ equals selfish component of welfare times the "altruism weight". The integrand in equation 8 equals the current flow of utility minus $K(\mathbf{S}_t)$, discounted at the rate used to evaluate intergenerational transfers.

This model includes familiar special cases. For $n > 1$, the endogenous function $F(\mathbf{S}, x) = f(\mathbf{S}, \mathbf{i}_{n-1}\chi(\mathbf{S}), x; n)$ depends on the policies of the other $n-1$ agents. Those agents do not exist if $n = 1$, in which case, $F(\mathbf{S}, x) = f(\mathbf{S}, x; 1)$, an exogenous function; there, the model collapses to a sequential game with a single agent at each point in time. For $\lambda = r$, $K = 0$ and the model collapses to a standard (constant discounting) differential game for $n > 1$ or a control problem for $n = 1$.

In general, the equilibrium to this game is not unique.[6] Tsutsui and Mino (1990) note the existence of an open interval of stable steady states in the game with constant discounting. For each point in this interval, there exists a different equilibrium policy function. The economic explanation for this multiplicity is that the decision whether to remain in a particular steady state depends on an agent's beliefs regarding the actions that rivals would take if a single agent were to drive the state away from that steady state. The MPE conditions do not pin down these beliefs. The same consideration applies for $n = 1$ under non-constant discounting. Thus, when $n > 1$ and the discount rate is non-constant, two sources of multiplicity create a coordination problem across coalitions and generations.

[6] Ekeland, Karp, and Sumaila (2015) study a model in which the equilibrium is unique, within the class that induce differentiable value functions. Dropping the differentiability assumption leads to many other MPE (Dutta and Sundaram 1993). Agents might "behave well" if the state variable is in a certain region, but follow a "bad" MPE if the state variable leaves that region. This kind of MPE has the flavor of trigger strategies in repeated games. There are many types of equilibria, apart from the MPE with differentiable value functions, studied here. Krusell and Smith (2003) and Vieille and Weibull (2009) discuss multiplicity in different settings.

3.3 The Green Golden Rule

For $n = 1$, there is a single coalition, and thus no conflict amongst contemporaneous agents; however, unless $\lambda = r$ or $\lambda = \infty$, the time inconsistency of preferences results in a game across generations, not a standard optimization problem. Here I assume that $n = 1$ and that the state is a scalar, S.

The "*G*reen *G*olden *R*ule" ("GGR") is the steady state chosen by an infinitely patient planner (Chichilnisky, Heal, and Beltratti 1995): $GGR \equiv \arg\max_S u(S, x; 1)$ subject to $f(S, x; 1) = 0$. I assume that this static optimization problem is concave, so steady state utility increases as the state variable approaches the GGR. There is no presumption that the GGR is an ethically attractive steady state, but it provides an obvious benchmark against which to compare any other steady state.

For small positive λ and bounded u_{it}, the payoff is well defined and is asymptotic to $\frac{u_\infty}{\lambda}$, the steady state utility flow divided by λ. For small λ, the payoff in the steady state therefore determines the evaluation of welfare. (Lemma 2 in Appendix B.1.) This fact and the assumed concavity of the problem that defines the GGR implies that MPE are Pareto ranked for λ sufficiently small: if a particular MPE supports a steady state not equal to the GGR, then all generations would prefer a deviation that causes the state to move closer to the GGR. Moreover, there exists a MPE that supports a steady state arbitrarily close to the GGR:[7]

Proposition 3 *With the class of differentiable MPE policy rules and $n = 1$, and for arbitrarily small positive ε, there exists a MPE steady state within ε of the GGR, provided that λ is sufficiently small (but positive).*

4 Climate policy

I present the climate model, discuss some of its features, and then explain why two of the many MPE have a special claim to our attention. The next subsections use the

[7] Karp (2007), Proposition 2, and Ekeland and Lazrak (2010), Theorem 8, establish similar results for cases in which the utility flow is independent of the state. A climate model requires that utility depend on the state variable, as in Proposition 3.

climate model to study these two equilibria. I then discuss the relative importance of altruism and cooperation.[8]

4.1 The linear-in-state model

In the "linear-in-state" (LIS) model: (i) the utility function is linear in the state variable and additively separable in the state and the control variables; and (ii) the equation of motion is linear in the state and in the control. I first describe the utility function and then the equation of motion, suppressing time indices.

Utility flow The state variable, $\mathbf{S}$, contains all climate-related stocks, such as temperatures and carbon stocks in the carbon reservoirs. Denote the first element as s (e.g. atmospheric temperature) and assume that climate-related damages depend only on s.[9] Define X as aggregate (= world-wide) emissions. The LIS structure means that for $n = 1$ the aggregate utility flow at a point in time, $u\left(X, s, t; n\right)_{|n=1}$, can be written as $v\left(X, t; n\right)_{|n=1} - \kappa s$, where κ is a parameter; to obtain a stationary equilibrium I assume that κ is a constant. In a symmetric equilibrium, $X = nx$, where x is emissions in a particular coalition. In order that n represent only an increase in fragmentation, the aggregate utility given the state s and aggregate emissions X must equal the sum of coalitions' utility if each coalition emits $x = \frac{X}{n}$: $nu\left(\frac{X}{n}, s, t; n\right) \equiv u\left(X, s, t; 1\right)$. This identity and the LIS structure require $u\left(\frac{X}{n}, s, t; n\right) = \frac{v(X,t;1) - \kappa s}{n}$. The function $v\left(X, t; 1\right)$ and the parameter κ thus determine the function $u\left(x, s, t; n\right)$.

One can take the function $v\left(X, t; 1\right)$ as primitive, but I use an alternative in which the aggregate utility flow, $u\left(X, s, t; 1\right)$, depends only on consumption and exogenously changing variables captured by t. Moreover, aggregate consumption depends only on aggregate emissions, the state s, and exogenously changing variables. For example, increased emissions increase consumption by making it possible to

[8] Supplementary material B.2.3 ("Robustness") discusses (i) an OLG model in which agents live a known finite amount of time, and (ii) a climate model that is not linear-in-state.

[9] This restriction is not plausible because, for example, damages might also depend directly on oceanic temperature. It is easy to dispense with this assumption in the LIS framework, but the generalization requires more demanding calibration.

avoid costly abatement; a larger climate-related stock, s, creates damage, decreasing output and thereby decreasing equilibrium consumption. I assume that utility is logarithmic in consumption. Denoting $C(X, t; 1)$ as aggregate consumption in the absence of environmental damage, and $e^{-\kappa s}$ as the multiplicative damage function, actual aggregate consumption is $C(X, t; 1) e^{-\kappa s}$ and the utility of consumption is $u(X, s, t; 1) = \ln C(X, t; 1) - \kappa s$. Defining $v(X, t; 1) = \ln C(X, t; 1)$ gives the desired form. The argument t allows for the possibility of exogenous changes, including those associated with changes in technology (e.g., carbon intensity) or capital stocks.[10]

Two examples of $C(X, t; 1)$ illustrate this formulation. In the first, C is Cobb Douglas in X, leading to a simplified version of Golosov et al (2014), the GeaS ("Golosov et al. Simplified") model. In the second ("Quadratic") example, C is the exponential of a quadratic in X, causing v to be quadratic.

Example 1 *(GeaS model): $C(X, t; 1) = A_t X_t^{\alpha_t}$. Aggregate utility is $u(X, s, t; 1) = \ln A_t + \alpha_t \ln X_t - \kappa s_t$, i.e. $v(X, t; 1) = \ln A_t + \alpha_t \ln X_t$. Setting $v(x, t; n) = \frac{\ln A_t}{n} + \frac{\alpha_t}{n} \ln(nx)$ means that in a symmetric equilibrium (where $X = nx$) aggregate utility is $n\left(\frac{\ln A_t}{n} + \frac{\alpha_t}{n} \ln(nx) - \frac{\kappa}{n} s\right) = u(X, s, t; 1)$.*

Example 2 *(Quadratic model): $C(X, t; 1) = \exp\left(a_{0,t} + a_t X - \frac{d_t}{2} X^2\right)$, so $v(X, t; 1) = a_{0,t} + a_t X - \frac{d_t}{2} X^2$. Here, utility for a coalition emitting x is $u(x, s, t; n) = v(x, t; n) - \frac{\kappa}{n} s$, with $v(x, t; n) = \frac{a_{0,t}}{n} + a_t x - n\frac{d_t}{2} x^2$. Aggregate utility in a symmetric equilibrium is $nu(x, s, t; n) = v(nx, t; 1) - \kappa s$, where $v(nx, t; 1) = v(X, t; 1) = a_{0,t} + a_t X - \frac{d_t}{2} X^2$.*

The equation of motion Carbon emissions enter the atmosphere and disperse amongst the different sinks, influencing temperature and altering the variable s, and thus altering the utility flow. With constant matrix B and vector b, LIS requires

$$\dot{\mathbf{S}} = B\mathbf{S} + bX. \tag{10}$$

[10] For example, if ω_t equals average carbon intensity of energy, and if all anthropogenic emissions were caused by energy consumption, then $X_t = \omega_t \times$ energy consumption. This formulation provides one of many ways to link emissions to economic variables.

Discussion of this model Integrated assessment models such as DICE treat capital as endogenous, although they typically treat other time-varying features such as technology as exogenous. My "stripped down" model treats everything except for the climate-related variables as exogenous.

Failure to treat capital as endogenous might not matter much. Golosov et al. (2014) use a discrete time model with endogenous investment, logarithmic utility, and Cobb Douglas production; capital depreciates 100% in a single period. In that setting (with $n = 1$ and constant discounting), the optimal savings rate is a constant that is independent of climate parameters. Gerlagh and Liski (2012) and Iverson (2013) study that discrete time model under more general discounting (with $n = 1$) and again find that the savings rate (in one equilibrium) is a constant, independent of climate parameters. These models decouple the investment and climate components.

In the continuous time setting, there is no analog to "100% depreciation in a period", so the savings rate in the continuous time setting, extended to include endogenous capital, would not be constant. I avoid this complication by taking the capital stock, in addition to technology, as exogenous. The functional assumptions in Golosov et al. (2014) produce an exact decoupling between investment and the climate. There, and in other models where the investment decision is insensitive to climate considerations (Hwang, Reynes, and Tol 2013), studying the climate problem in isolation from the investment decision has little effect on climate policy.

Technological progress and capital accumulation might make distant generations so much richer than us, that climate-induced reductions in their consumption are unimportant. Reductions in future carbon intensity might make future abatement cheap. In these cases, we should not sacrifice much today to reduce our carbon emissions. These policy conclusions are driven by assumptions about technology.

There are at least three reasons why we might want a model in which policy is not driven by the assumption that we will grow our way out of the climate problem. First, the familiar relation between high expected growth and a high consumption discount rate arises in the standard model with time-additive expected utility; making growth uncertain leads to only a second-order correction. However, in a model that disentangles risk aversion from the elasticity of intertemporal substitution, Traeger (2014) shows that *stochastic* growth (compared to zero growth) might have little effect on the certainty equivalent discount rate. Second, the assumption, adopted by most integrated assessment models, that natural and man-made capital are highly substitutable, may be incorrect (Guesnerie 2004, Hoel and Sterner 2007, Traeger 2011). In that case, we may want to protect natural capital even if future generations have much more man-made capital than we do. Third, most integrated assessment models identify growth with increased GDP, leading to increased consumption. The limitations of GDP as the sole index of well-being are well understood; alternatives or supplements include the Genuine Progress Indicator (GPI), Human Development Index (HDI) and Ecological Footprint. Kubiszewski et al. (2013) discuss these, and note that over the past 25 years GPI has been flat, while GDP has continued to grow: the indices might be only weakly correlated.

The modeling dilemma is that we can anticipate large changes in technology, but we might want to avoid having today's climate policy driven by beliefs about future technological improvements. The LIS model provides one solution to this dilemma. The parameters in the equation of motion, $\dot{\mathbf{S}}$, are determined by natural processes, and thus independent of technology. The damage parameter, κ, could be altered by technology, but given the model's level of abstraction, treating κ as a constant parameter is defensible. With this assumption, only the function v depends explicitly on changing technology or capital. For one equilibrium studied below, the tax in utility units is independent of v, and thus independent of technology. However, both emissions and the tax in monetary units do depend on v, and thus on current technology. They do not, however, depend on beliefs about *future* technology.

Equilibrium selection Section 3.2 notes the generic multiplicity of MPE. The most natural equilibrium candidate, the "limit equilibrium", is the limit of the sequence of equilibria of finite horizon models, as the horizon goes to infinity. For the LIS model, this equilibrium is unique, independent of the state variable, and dominant. Gerlagh and Liski (2012) and Iverson (2013) use this equilibrium to study LIS climate models with $n = 1$ (although Gerlagh and Liski (2012) describe the equilibrium differently).

The infinite horizon model also has many other differentiable MPE. The life of our planet is finite, but insisting on a finite horizon model (or its limit) implies that there is some generation that *knows* it is the last generation. Equilibria that rely on an infinite horizon are used throughout economics, and can be motivated as ε-equilibria to a finite horizon game (Fudenberg and Levine 1983).

For a scalar specialization of the LIS model, I use the infimum of the set of states that can be supported as a MPE steady state as a means of describing the set of "non-limit" equilibria . All other MPE steady states lie above this infimum, and thus are further from the GGR and have lower steady state utility. The limit and the non-limit equilibria have different properties: the latter, unlike the former, are functions of the state variable and are not dominant. Consideration of non-limit equilibria provides both a different perspective on the climate policy game and a robustness check for conclusions obtained using the limit equilibrium.

Equilibrium representation It is convenient to present results in terms of a tax, measured in units of utility, instead (as is more common) in monetary units. The utility-denominated tax, τ, that supports aggregate emissions X, in a decentralized *aggregate* economy is

$$\tau(\mathbf{S},t) = \frac{1}{C(X,t;1)\,e^{-\kappa s}} \frac{\partial C(X,t;1)}{\partial X} e^{-\kappa s} = \frac{1}{C(X,t;1)} \frac{\partial C(X,t;1)}{\partial X}, \qquad (11)$$

which equals the marginal utility of consumption times the marginal increase in consumption due to an extra unit of emissions. This tax has units of utility/emissions. Dividing by the marginal utility of consumption (multiplying by $C(X,t;1)\,e^{-\kappa s}$) converts the tax from utility to monetary units.

4.2 The limit equilibrium

Here I assume that $\lambda \in (0,r]$, so I use Proposition 2. In a finite horizon model, backward induction yields a unique equilibrium. The utility-denominated tax in the limit equilibrium is independent of the state variable and time, t. If other agents use state-independent decision rules, then the shadow value of the state, for an arbitrary coalition at an arbitrary point in time, depends on model parameters but not on the level of the state. Therefore, the agent's optimal action is independent of the state. The independence with respect to time is then a consequence of the fact that the climate-related parameters (B and κ) do not depend on time. I characterize the limit equilibrium for the general model, and then specialize to the climate setting.

General results Denote the diagonal matrix of eigenvalues of B (in the equation of motion 10) as Λ, with i'th diagonal element Λ_i, and the matrix of eigenvectors as P. I assume that Λ_i are non-positive real numbers and P is of full rank.

Proposition 4 *For the LIS model, if other agents use state-independent (but possibly time- and coalition-dependent) emissions policies, the functions $K(\mathbf{S}, t)$ and $J(\mathbf{S}, t)$ in the auxiliary control problem (Proposition 2) are linear in the state; $J(\mathbf{S}; t) = g_{0t} + g'\mathbf{S}$ with:*

$$g' = \frac{\kappa}{n}\left(i_1' - (r - \lambda)\,\tilde{q}'\right)(B - \lambda I)^{-1}, \text{ with } \tilde{q}' = \int_0^\infty i_1' P e^{-(\gamma I - \Lambda)\tau} P^{-1} d\tau, \tag{12}$$

where I is the m dimensional identity matrix and i_1 the first unit vector.

(ii) A planner's best response to rivals' state-independent policies is independent of the state, and is a dominant strategy: it does not depend on her beliefs about the state-independent emissions of any future planner, or about the actions of other current planners. Equilibrium emissions are also independent of beliefs about future technology. Within the class of state-independent policies, the unique equilibrium is

$$\chi = \arg\max_x v(x, t; n) + g'bx. \tag{13}$$

Iverson (2013) demonstrates uniqueness of the limit equilibrium for the discrete-time log-linear model with 100% depreciation in a period and $n = 1$. He also shows that the first period action of a planner who can commit equals the action in a state-independent MPE. Phelps and Pollack (1968) obtain this result for a simpler model. These results are consistent with Proposition 4.ii, which holds for arbitrary n and concave function v. State-independence is a consequence of the LIS structure. State independence, plus Assumption 1.iii, imply dominance.

The utility-denominated tax depends on climate-related parameters, n, and discounting parameters, but not on $v(\cdot)$ or the state or t:

Corollary 1 *(i) The utility-denominated tax that in the aggregate economy supports the equilibrium level of emissions, $\tau \equiv \frac{\partial v(X(n),t;1)}{\partial X}$, is*

$$\tau = -\frac{\kappa}{n}\left(i_1' - (r-\lambda)\,\tilde{q}'\right)(B-\lambda I)^{-1} b, \tag{14}$$

and thus independent of the state variable, time, and the payoff function $v(\cdot)$. (ii) The absolute value of the elasticity of this tax with respect to n is 1, and the elasticity of the tax with respect to λ is

$$\varepsilon \equiv -\frac{d\tau}{d\lambda}\frac{\lambda}{\tau} = -\lambda \frac{\left[\tilde{q}' + \left(i_1' - (r-\lambda)\,\tilde{q}'\right)(B-\lambda I)^{-1}\right](B-\lambda I)^{-1} b}{\left(i_1' - (r-\lambda)\,\tilde{q}'\right)(B-\lambda I)^{-1} b}. \tag{15}$$

Equilibrium emissions, unlike the utility-denominated tax, depend on v, and thus on current (but not future) technology (equation 13). Examples 1 and 2 illustrate the relation between $v(\cdot)$ and equilibrium emissions. For the GeaS model, if each of n coalitions fragments into two, aggregate equilibrium emissions double, despite no fundamental (non-strategic) change in the economy. The Quadratic model does not have this extreme feature:

Corollary 2 *In the GeaS model, equilibrium emissions per coalition are independent of n, so aggregate emissions are increasing in and proportional to n. In the Quadratic model, aggregate emissions are a strictly concave increasing function of n, so emissions per coalition fall with n.*

The climate application Temperature change likely adjusts to GHG stocks with a lag.[11] This lag causes the marginal damage of emissions to rise over time (as

[11] In DICE and other policy models, the delay between emissions and maximum temperature change is several or many decades. Ricke and Caldeirar (2014) challenge this feature, claiming that most of the temperature change due to current emissions occurs within a decade.

temperature slowly responds to the increased GHG stock) and eventually to decrease (as the stock dissipates, and temperature slowly adjusts). A two-dimensional model captures this non-monotonicity with respect to time, and still leads to an explicit expression for the equilibrium tax, τ. I also consider a scalar model, which eliminates the lag, forcing the marginal damage of current emissions to fall over time.

For the two-dimensional model, I choose units of s so that $\kappa = -1$, making the fractional consumption loss at t due to climate change equal $1 - e^{s(t)}$. I assume that emissions enter the atmospheric stock and decay at a constant rate. Rezai (2010) reports estimates of dissipation rates for CO_2 that imply half-lives between 126 and 276 years, for a midpoint of 200 years, the value that I use.[12] To calibrate the two remaining climate parameters, I rely on Gerlagh and Liski (2012), who use DICE. I adopt their assumptions that: (i) doubling atmospheric stocks (relative to preindustrial levels) reduces output (in my setting, consumption) by 2.6%, once s has adjusted, and (ii) following a pulse increase in atmospheric CO_2, the loss rises during the first 60 years, and then falls slowly. Appendix B.2.2 provides the formula for the equilibrium tax, explains how I use the calibration assumptions, and discusses the relation between my two-dimensional model and Gerlagh and Liski (2012) and Nordhaus (2008).

The dashed curve in Figure 1 shows the tax elasticity with respect to λ, and the solid curve shows the tax expressed in \$100/ton of CO_2. The tax increases by a factor of 4.9 as λ falls from $r = 0.02$ to its lower bound 0, ranging from \$17/ton to \$83/ton. (I multiply τ by annual gross world consumption to obtain a tax measured in dollars. I set gross world output to \$63 Tr and the investment rate to 25%, giving consumption at \$47.25 Tr. $C0_2$ is in units of Teratons, so the units of $\tau \times \$Tr$ are dollars/ton of $C0_2$.)

Remark 1 *There is "observational equivalence" between the model with discount parameters* r, λ *and a second model with a constant discount rate* $R(\lambda; r)$*: the two*

[12] No single half-life provides a close approximation of the carbon cycle, although a higher dimensional linear model, with different sinks, can approximate that cycle (Forster 2007, note a, page 213). The two-dimensional linear model captures important aspects of the non-monotonic dynamics contained in earlier models.

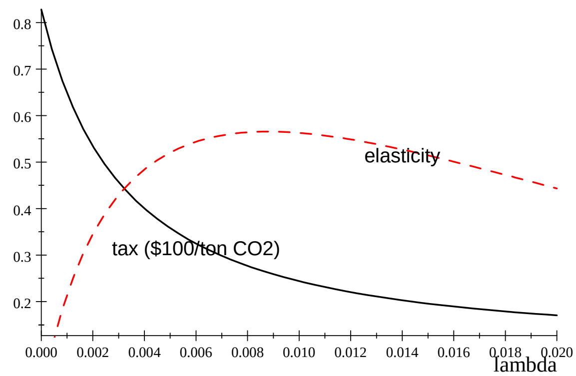

Figure 1: Tax elasticity with respect to λ, and the tax (\$100/ton CO_2); $n = 1$, baseline climate parameters, $r = \theta = 0.02$ and world consumption = \$47.25 trillion.

models yield the same equilibrium tax. For the baseline parameters in the two-dimensional model, with $n = 1$, $R(\lambda; r)$ *increases by a factor of almost 10, from* 0.003 *(at* $\lambda = 0$*) to* 0.02 *(at* $\lambda = r = 0.02$*).*

Observational equivalence is not general, but is due to functional assumptions.

The tax and elasticity formulae are even simpler if $\mathbf{S}$ is a scalar, s; here, $\dot{s} = Bs + X$, B is a scalar, $\Lambda = B < 0$ and $\tilde{q} = \frac{1}{\gamma - B}$. Specializing Corollary 1 implies:

Corollary 3 *For the case where* $\mathbf{S}$ *is a scalar,*

$$\tau = \frac{\kappa(\theta - B + \lambda)}{n(\lambda - B)(\gamma - B)}, \text{ and } \varepsilon = \frac{\lambda\theta}{(\lambda - B)(\theta + \lambda - B)};$$

the elasticity ε *is independent of* r *and increases with* B*, reaching its upper bound at* $B = 0$*, where* $\varepsilon = \frac{\theta}{\theta + \lambda} < 1$*: as the pollutant becomes more persistent, the tax becomes more sensitive to* λ*. As* B *varies over* $(-\infty, 0)$*, the tax varies monotonically over* $\left(0, \frac{\kappa(\theta+\lambda)}{n(r+\theta)\lambda}\right)$.

The dashed graph in Figure 1 shows the elasticity of the tax with respect to λ in the two-dimensional model. This elasticity is independent of n, and the elasticity of the tax with respect to n is identically 1. The curve is always less than 0.6. In the scalar limit equilibrium, Corollary 3 gives the formula for the elasticity ε. This elasticity is maximized at $\lambda = \sqrt{B^2 - B\theta}$. At a 2%/year mortality, the maximum elasticity ranges from 0.33 to 0.69 as the half-life of the stock ranges from 100 to 1000 years. For the baseline value of a 200 year half-life, the maximum elasticity equals 0.45. Comparison of the two-dimensional and the scalar models shows that the lag between emissions and damages increases the elasticity with respect to altruism (having no effect on the elasticity with respect to cooperation). The magnitude of the elasticity is similar for both the scalar and two-dimensional model. This similarity is important, because I rely on the scalar model for the non-limit equilibria.

4.3 "Non-limit" MPE

The limit equilibrium for the LIS model is dominant and state-independent. Many other, qualitatively different MPE exist in the infinite horizon setting. The procedure for obtaining the necessary conditions for these "non-limit" equilibria in the *scalar* model is straightforward.[13] This section uses exclusively the scalar model, where $\dot{s} = Bs + X$ and $B < 0$. Non-limit equilibria (unlike the limit equilibrium) depend on the flow payoff, $v(\cdot)$. I use the GeaS model, where $C(X, t; 1) = A_t X_t^{\alpha}$, requiring α to be constant, to consider stationary equilibria; A_t can depend on time.

I define $\Phi(s) \equiv \frac{dX(s)}{ds}_{|s=s_\infty}$, the derivative of aggregate equilibrium emissions, evaluated at the steady state; $\Phi(s)$ provides a means of comparing steady states of different MPE equilibria. The analysis here subsumes the limit equilibrium, $\Phi = 0$. Local asymptotic stability ("stability") of any equilibrium requires $B + \Phi < 0$. Thus, $\Phi = -B > 0$ is the supremum of values of Φ that correspond to stable equilibria.

The GRR (a steady state stock) for this model is $\frac{\alpha}{\kappa}$. To eliminate α, κ, I define $\Upsilon(\Phi; n, \lambda, B, r, \theta)$ as the steady state stock induced by Φ, expressed as a ratio of the GGR. Using equation 11 and the GeaS model, $\tau(s) = \frac{\alpha}{X(s)}$; here, the steady state tax equals $\tau_{|\infty} = -\frac{\alpha}{Bs_{|\infty}}$. Thus, Υ also equals the tax that supports the GGR divided

[13] Numerical methods using function iteration can find a differentiable MPE for a non-scalar model. My experience with these methods suggests that they identify only the "limit equilibrum" (Fujii and Karp 2008). Other papers that consider multiplicity of equilibria under non-constant discounting (with $n = 1$) in the scalar case include Karp, 2005 and Karp and Tsur, 2011.

by the MPE steady state tax.

Lemma 1 *(i) The ratio of the steady state stock induced by* Φ*, and the GGR, is*

$$\begin{array}{c}\Upsilon(\Phi;\cdot) = -\frac{\left(n\left(B-\lambda+\frac{1}{n}\Phi(n-1)\right)(B+\Phi-\gamma)-\Phi(r-\lambda)\right)}{B(\lambda-r-\Phi-B+\gamma)} \Rightarrow \\ \Upsilon(0) = \frac{n}{B}(B-\lambda)\frac{B-\gamma}{B+r-\lambda-\gamma} \text{ and } \Upsilon(-B) = 1 - n\lambda\frac{\gamma}{B(\theta+\lambda)} > 1.\end{array} \quad (16)$$

(ii) The ratio $\Upsilon(\Phi)$ *is greater than* 1 *and decreases in* Φ*; therefore, the infimum of stable steady states corresponds to* $\Phi = -B$*.*

Lemma 1.i is a formula, and part ii establishes that larger values of Φ correspond to smaller steady states. Hereafter I consider only $\Phi \geq 0$; negative values induce higher steady state stocks, strengthening the results discussed below.

Proposition 5 *(i) For any* $0 \leq \Phi < -B$*, increased cooperation or altruism (smaller* n *or* λ*) move the MPE steady state closer to the GGR:* $\frac{d\Upsilon}{d\lambda} > 0$*,* $\frac{d\Upsilon}{dn} > 0$*. (ii) Increased cooperation (respectively, increased altruism) makes the steady state less sensitive to altruism (respectively, cooperation):* $\frac{d^2\Upsilon}{dnd\lambda} > 0$*. (iii) For all values of* n*, there exists a MPE steady state arbitrarily close to the GGR for* λ *close to 0. This steady state is supported by a policy function corresponding to* Φ *close to its upper bound (*$\Phi = -B$*). In contrast, even for full cooperation (*$n = 1$*), the steady state is bounded away from the GGR for* λ *bounded away from 0.*

Proposition 5.i confirms that increased cooperation or altruism lead to lower steady state carbon stocks (higher taxes). Part (ii) states, for example, that increased altruism (smaller λ) leads to a larger fall in the steady state stock, the lower is cooperation (the larger is n). Part (iii) implies that for sufficiently high levels of altruism (λ close to 0), there is a MPE steady state (corresponding to Φ close to $-B$) near the GGR, regardless of the degree of international cooperation. In contrast, perfect international cooperation ($n = 1$) does not lead to a steady state near the GGR if altruism is limited (λ bounded away from 0). Proposition 5.iii complements Proposition 3; the latter holds for $n = 1$ and general functional forms, whereas the former holds for general n and the LIS model.

Remark 2 *Because the equilibrium steady state decreases in* Φ*, the move from* $\Phi = 0$ *to* $\Phi = -B$ *maximizes the percent change in the steady state (for* $\Phi \geq 0$*). This maximum percent change (*MPC*) provides a measure of the importance of the multiplicity of equilibria in the neighborhood of the steady state.* $MPC = 0$ *if* $\lambda = r$ *and* $n = 1$*; because of continuity,* MPC *is small at neighboring values, where there is little loss in generality in focusing exclusively on the limiting equilibrium. For* $\theta = r = 0.02$ *and the 200 year half-life,* MPC *is negligible for* $n = 1$ *and* $\frac{r}{2} < \lambda < r$*, but is large (50 – 80%) for other combinations of* n, λ*.*

Dynamic Strategic complements Because $\tau(s) = \frac{\alpha}{X(s)}$, in the steady state $\frac{d\tau}{ds} = -\frac{\tau}{X}\Phi$. Thus, for $\Phi > 0$, the equilibrium tax decreases in the state in the neighborhood of the steady state. Here, taxes are dynamic strategic complements (Jun and Vives 2004): If a coalition reduces its current tax below the equilibrium level, the future stock is above the equilibrium level, causing future taxes to be lower than the non-deviation level; future taxes respond in the same direction as a current deviation. Equilibria with lower steady state stocks correspond to dynamic strategic complementarity. When taxes are strategic complements, coalitions have an incentive to use high taxes, in order to keep the stock low, thereby encouraging future decision-makers to use high taxes and also maintain a low stock. This incentive is absent in the limit equilibrium with state-independent taxes, where the steady state stock is consequently higher.

Strategic complementarity arises from two features that likely hold in more general models: (i) A higher tax reduces emissions everywhere, not just at the steady state. (ii) Whenever there is a natural decay rate, i.e. when $\dot{s}$ can be written as $f(s,X)$, with $f_s < 0 < f_X$, there are locally asymptotically stable steady states with $X'(s) > 0$ near the steady state (because stability requires only $f_S + f_X\Phi < 0$). For equilibrium tax rules that support these steady states, the equilibrium tax decreases in the stock, in the neighborhood of the steady state. For the GeaS model (*but not more generally*) the limit equilibrium (where $\Phi = 0$) divides the equilibrium set between dynamic strategic substitutes and strategic complements.

4.4 The relative importance of altruism and cooperation

This section examines the relative importance of either small or large changes in altruism and cooperation. The utility-denominated tax is constant in the limit equilibrium and varies with the state variable in non-limit equilibria. To nest these equilibria, I compare the effect of altruism and cooperation on the steady state utility-denominated tax, or equivalently the steady state stock. I use $I^{coop} \equiv \frac{1}{n} \in [0,1]$ as an index of cooperation and $I^{altruism} \equiv \frac{r-\lambda}{r} \in [0,1)$ as an index of altruism.[14]

Small changes in altruism or cooperation The "local criterion" (LC) equals the semi-elasticity of the steady state with respect to the altruism index, divided by the semi-elasticity with respect to the cooperation index; LC is the ratio of two derivatives.[15] Figure 2 shows two level sets of LC; the solid curves correspond to the limit equilibrium ($\Phi = 0$) and the dashed curves correspond to the non-limit equilibrium with $\Phi = -B$. Moving north-east in the unit square raises the level of cooperation and altruism (lowers n and λ).

At combinations of altruism and cooperation north-east of the solid line labelled $LC = 1$, $LC > 1$ in the limit equilibrium. Over this region, the outcome is more sensitive to altruism than to cooperation (as measured by the ratio of semi-elasticities). The outcome is more sensitive to cooperation south-west of this curve, where cooperation or altruism are low. Each dashed curve lies below the corresponding solid curve. A larger value of Φ leads to a lower equilibrium steady state carbon stock, and also increases the parameter set where a small increase in altruism has a larger equilibrium effect than a small increase in cooperation.

The policy interpretation of this figure depends on which part of parameter space

[14]As noted above Proposition 2, the equilibrium condition differs depending on whether $\lambda < r$ or $\lambda > r$. My analysis is restricted to the case $\lambda \leq r$. My altruism index reflects this restriction, assigning a zero index value to $\lambda = r$. Treating intra- and inter-personal transfers as equivalent ($\lambda = r$) is not literally the same as "zero altruism". An alternative index $\hat{I}^{altruism} \equiv \frac{\theta+r}{\theta+r+\lambda} \in [0,1)$ implies (for $\theta = r$) $\hat{I}^{altruism} = \frac{2}{3}$, not 0, at $\lambda = r$.

[15]These derivatives use the formulae for $\Upsilon(0)$ and $\Upsilon(-B)$ in equation 16, expressed as functions of the cooperation index $\frac{1}{n}$ and the altruism index $\frac{r-\lambda}{r}$. Earlier versions of this paper used a different "Local Criterion", the ratio of elasticities, instead of derivatives. This alternative local criterion is less than 1 for all equilibria with $0 \leq \Phi < -B$.

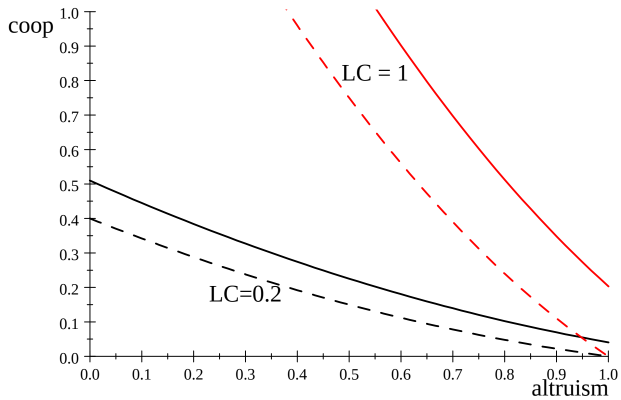

Figure 2: Level sets of $LC = 1$ and $LC = 0.2$ in the limit equlbrium (solid) and for the non-limit equilibrium with $\Phi = -B$. Above $LC = 1$, a small increase in altruism has a larger equilibrium effect than a small increase in cooperation.

most accurately characterizes the current policy environment. Most integrated assessment models are based on the infinitely lived agent model, at which $\lambda = r$. To the extent that policy would be guided by these models, in the event that nations began to cooperate, the altruism index is 0. In addition, governments clearly do not act in unison. A cooperation index well below 0.5 ($n = 2$) seems reasonable. With low levels of altruism and cooperation, the figure suggests that a small increase in cooperation would have a much larger equilibrium effect, compared to a small increase in altruism; $LC = 0.2$ or lower seems plausible.

Large changes in altruism or cooperation Figure 2 shows that the relative importance (measured by semi-elasticities) of small changes in altruism and cooperation depends on the value of those indices. Therefore, a large increase in cooperation might have either a larger or a smaller equilibrium effect than a large increase in altruism. Figure 3 graphs the level sets of Υ (the ratio of a MPE steady state to the GGR), showing that a large increase in altruism has a larger equilibrium effect. The magnitudes $\Upsilon = 6.8$ and $\Upsilon = 27.1$ correspond to the equilibrium Υ in the limit equilibrium for the infinitely lived agents ($\lambda = r$, or $I^{altruism} = 0$) when

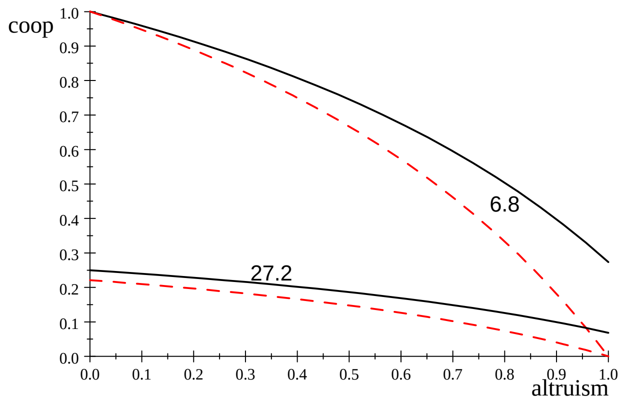

Figure 3: Level sets of $\Upsilon(\Phi)$, the ratio of a MPE steady state to the GGR. Solid curves correspond to $\Phi = 0$ and dashed curves correspond to $\Phi = -B$. $\Upsilon = 6.8$ and $\Upsilon = 27.1$ are the ratio of eqquilibrium-to-GGR steady states in the limit equilibrium for the infinitely lived agent ($\lambda = r$, or $I^{altruism} = 0$) for $n = 1$ and $n = 4$. Half-life $= 200$ years and $\theta = r = 0.02$.

$n = 1$ and $n = 4$, respectively.

If planners act like the infinitely lived agent, then $\lambda \approx r$ is appropriate. If the world is less cooperative than it would be if three blocs, the EU, the BRIC nations, and North America, determine policy, then $n \geq 4$. At $n = 4$ and $\lambda = r$, a move to maximum international cooperation decreases the steady state by a factor of four. A move close to maximum altruism, holding international cooperation at $n = 4$, decreases the steady state by a factor of over 27. For the baseline parameter values, large increases in altruism have a larger equilibrium effect than large increases in cooperation As Proposition 5.iii notes, the system can get close to the GGR steady state if altruism is high, even if cooperation is low; it cannot get close to the GGR if altruism is low, even if cooperation is high.

5 Discussion

The provision of a long-lived public good, such as a stable climate, depends on the ability of contemporaneous agents to cooperate, and on their degree of altruism towards future generations. A differential game/overlapping generations model shows how these two features interact, and provides estimates of their relative importance in determining equilibrium policy. Altruism is especially influential when cooperation is low, and cooperation is especially influential when altruism is low.

At empirically plausible levels of altruism and cooperation, a small increase in cooperation has a much larger equilibrium effect than does a small increase in altruism. This comparison is stronger for the limit equilibrium, but also holds for all other differentiable equilibria resulting in lower steady state carbon stocks. Large changes in altruism and cooperation can reverse this comparison. A move to full cooperation increases the provision of the public good, but is unlikely to get the steady state close to the "Green Golden Rule" level. In contrast, a move close to maximal altruism can get the state close to this level, even at low levels of cooperation.

The multiplicity of equilibria opens the possibility that actions are dynamic strategic complements, rather than strategic substitutes (or dominant, as in the limit equilibrium). The logic of Nash's noncooperative equilibrium does not doom us to bad outcomes, even if we exclude trigger or other punishment strategies. (Consideration of such strategies increases the equilibrium set.) This conclusion, although not specific to this paper, is nevertheless worth stating, because many non-cooperative models of climate policy build in strategic substitutability, implying that agents have an incentive to undertake less public investment, partly to induce their successors to invest more. This built-in free riding causes the models to be quite pessimistic about the chance of a meaningful climate agreement amongst sovereign nations. Recognition of the possibility of strategic complementarity, where agents have an incentive to increase their current investment partly to induce higher future investment, moderates this pessimism. International negotiations on climate policy are important, even if they do not result in enforceable agreements. Negotiations make coordination on a good equilibrium easier to achieve.

References

Aldy, J., and R. Stavins (eds.) (2007): *Architectures for Agreement: Addressing Global Climate change in the Post-Kyoto World*. Cambridge Univerisity Press.

Arrow, K., M. Cropper, B. Groom, G. Heal, R. Newell, W. Norhaus, R. Pindyck, W. Pizer, P. Portney, T. Sterner, R. Tol, and M. Weitzman (2013): "Determining Benefits and Costs for Futur Generations," *Science*, 341, 349–350.

Barrett, S. (2005): "The theory of international environmental agreements," in *The Handbook of Environmental Economics*, ed. by K.-G. Maler, and J. Vincent. North Holland, Amsterdam.

Calvo, G., and M. Obstfeld (1988): "Optimal time-consistent fiscal policy with finite lifetimes," *Econometrica*, 56, 411–32.

Chichilnisky, G., G. Heal, and A. Beltratti (1995): "The Green Golden Rule," *Economics Letters*, 49(2), 175–179.

Drupp, M., M. Freemand, B. Groom, and F. Nesje (2014): "Disounting disentangled: an expert survey on the determinants of teh long term social discount rate," London School of Economics working paper.

Dutta, P. K., and R.-K. Sundaram (1993): "How Different Can Strategic Models Be?," *Journal of Economic Theory*, 60, 42–61.

Ekeland, I., L. Karp, and R. Sumaila (2015): "Equilibrium management of fisheries with altruistic overlapping generations," *Journal of Environmental Economics and Management*, 70, 1–16.

Ekeland, I., and A. Lazrak (2010): "The golden rule when preferences are time inconsistent," *Mathematical and Financial Economics*, 4(1), 29–55.

Forster, P., and many others (2007): "Changes in Atmospheric Constituents and in Radioactive Forcing," in *Climate Change 2007: The Physical Science Basis. Contribution to Working Group 1 of the Fourth Assessment Report*, ed. by S. Solomon, and many others. Cambridge: Cambridge University Press.

Fudenberg, D., and D. Levine (1983): "Subgame perfect equilibria of finite and infinite horizon games," *Journal of Economic Theory*, 31, 251–268.

Fujii, T., and L. Karp (2008): "Numerical analysis of non-constant pure rate of time preference: a model of climate policy," *Journal of Environmental Economics and Management*, 56, 83–101.

Gerlagh, R., and M. Liski (2012): "Carbon prices for the next thousand years," Unpublished Working Paper.

Gerlagh, R., and B. van der Zwaan (2001): "Overlapping generations versus infintely- lived agent: the case of global warming," in *The long-term economics of climate change: beyond a doubling of greenhouse gas concentrations*, ed. by D. Hall, and R. Howarth, pp. 461–516. Emerald Group Publishing.

Gollier, C., and M. Weitzman (2010): "How should the distant future be discounted when discount rates are uncertain?," *Economic Letters*, 107, 350–353.

Golosov, M., J. Hassler, P. Krusell, and A. Tsyvinski (2014): "Optimal taxes on fossil fuels in general equilibrium," *Econometrica*, 82, 41–88.

Guersnerie, R. (2004): "Calcul economique et developpement durable," *Revue economique*, 55(3), 363–382.

Guesnerie, R., and H. Tulkens (eds.) (2008): *The Design of Climate Policy.* MIT Press.

Harstad, B. (2012): "Climate Contracts: a game of emissions, investmests, negotiations, and renegotiations," *Review of Economic Studies*, 79, 1527–1555.

Haurie, A., J. Krawczyk, and G. Zaccour (2012): *Games and Dynamic Games*. Imperial College Press; World Scientific, Hackensack, New Jersey.

Hoel, M., and T. Sterner (2007): "Discounting and Relative Prices," *Climatic Change*, 84, 265– 80.

Howarth, R. B. (1998): "An Overlapping Generations Model of Climate-Economy Interactions," *Scandinavian Journal of Economics*, 100(3), 575–91.

Hwang, I., F. Reynes, and R. Tol (2013): "Climate policy under fat-tailed risk:an application of DICE," *Environmental and Resource Economics*, 56(3), 415–436.

Iverson, T. (2013): "Optimal carbon taxes with non-constant time preference," http://mpra.ub.uni-muenchen.de/43264/.

Jackson, M., and L. Yariv (2015): "Collective dynamic choice: the necessity of time inconsistency," *American Economic Journal: Microenconmics*, 00-00, 00.

John, A., R. Pecchenino, D. Schimmelpfennig, and S. Schreft (1995): "Short-lived agents and the long-lived environment," *Journal of Public Economics*, 58(1), 127–141.

Jun, B., and X. Vives (2004): "Strategic incentives in dynamic duoploy," *Journal of Economic Theory*, 116, 249–81.

Karp, L. (2005): "Global Warming and hyperbolic discounting," *Journal of Public Economics*, 89, 261–282.

——— (2007): "Non-constant discounting in continuous time," *Journal of Economic Theory*, 132, 557 – 568.

Karp, L., and Y. Tsur (2011): "Time perspective and climate change policy," *Journal of Environmental Economics and Management*, 62, 1–14.

KEMP, M. C., AND N. V. LONG (1979): "The Under-Exploitation of Natural Resources: A Model with Overlapping Generations," *The Economic Record*, 55(3), 214–221.

KOSKELA, E., M. OLLIKAINEN, AND M. PUHAKKA (2002): "Renewable Resources in an Overlapping Generations Economy Without Capital," *Journal of Environmental Economics and Management*, 43(3), 497–517.

KRUSELL, P., AND A. SMITH (2003): "Consumption-saving desisions with quasi-geometric discounting," *Econometrica*, 71(1), 365–75.

KUBISZEWSKI, I., R. COSTANZA, C. FRANCO, P. LAWN, J. TALBRETH, T. JACKSON, AND C. AYLMER (2013): "Beyond GDP: Measuring and achieving global genuine progress," *Ecological Economics*, 93, 57–68.

LAIBSON, D. (1997): "Golden eggs and hyperbolic discounting," *Quarterly Journal of Economics*, 62, 443–78.

LAURENT-LUCCHETTI, J., AND A. LEACH (2011): "Generational welfare under a climate-change policy with induced innovation," *Scandinavian Journal of Economics*, 113, 904–936.

LI, C. Z., AND K. G. LOFGREN (2000): "Renewable Resources and Economic Sustainability: A Dynamic Analysis with Heterogeneous Time Preferences," *Journal of Environmental Economics and Management*, 40, 236–250.

LONG, N. V. (2010): *A Survey of Dynamic Games in Economics*. World Scientific, Singapore.

NORDHAUS, W. D. (2007): "A Review of the Stern Review on the Economics of Climate Change," *Journal of Economic Literature*, 45(3), 686 – 702.

PHELPS, E., AND R. POLLACK (1968): "On Second-best National Savings and Game: Equilibrium Growth," *Review of Economic Studies*, 35, 185–199.

Rasmussen, T. (2003): “Modeling the Economics of Greenhouse Gas Abatement: An Overlapping Generations Perspective,” *Review of Economic Dynamics*, 6(1), 99–119.

Rezai, A. (2010): “Recast the DICE and its policy recommendations,” *Macroeconomic Dynamics*, 14, 275–289.

Ricke, K., and K. Caldeirar (2014): “Maximum warming occurs about one decade after a carbon dioxide emission,” *Environmental Research Letters*, 9, http://dx.doi.org/10.1088/1748-9326/9/12/124002.

Roemer, J. (2009): “Equality: its justification, nature and domain,” in *The Oxford Handbook of Economic Inequality*, ed. by W. Salverda, B. Noland, and T. Smeeding, pp. 139–187. Oxford University Press, Oxford.

Roemer, J. E. (2011): “The ethics of intertemporal distribution in a warming planet,” *Environmental and Resource Economics*, 48, 363–390.

Saez-Marti, M., and J. Weibull (2005): “Discounting and altruism to future decision- makers,” *Journal of Economic Theory*, (122), 254–66.

Schneider, M., C. Traeger, and R. Winkler (2012): “Trading off generations: equity, efficiency and climate change,” *European Economic Review*, 56, 1621–1644.

Stern, N. (2007): *The Economics of Climate Change*. Cambridge University Press.

Strotz, R. (1956): “Myopia and Inconsistency in Dynamic Utility Maximization,” *Review of Economic Studies*, 23, 165–180.

Traeger, C. (2011): “Sustainabililty, limited substitutability and non-constant social discount rates,” *Journal of Environmental Economics and Management*, 62, 215–228.

——— (2014): “Why uncertainty matters - Discounting under Intertemporal risk aversion and ambiguity,” *Economic Theory*, 56, 627–664.

Tsutsui, S., and K. Mino (1990): "Nonlinear Strategies in Dynamic Duopolisitc Competition with Sticky Prices," *Journal of Economic Theory*, 52, 136–161.

van der Ploeg, R., and A. de Zeeuw (1992): "International aspects of pollution control," *Environmental and Resource Economics*, 2, 117–39.

Vieille, N., and J. Weibull (2009): "Multiple solutions under quasi-exponential discounting," *Economic Theory*, 39, 513–26.

Weitzman, M. (2007): "A Review of the Stern Review on the Economics of Climate Change," *Journal of Economic Literature*, 45(3), 703–724.

Wirl, F. (1994): "Pigouvian taxation of energy for stock and flow externalities," *Environmental and Resource Economics*, 26, 1–18.

Zuber, S. (2010): "The aggregation of preferences: can we ignore the past?," *Theory and Decision*, 70, 367 – 384.

A Proofs

The supplementary material provides more details of some of these proofs.

Proof. (Proposition 1) The restrictions $\lambda > 0$ and $\tilde{\lambda} > \theta$ and the assumption that u_τ is bounded (so U_t is bounded) imply that the integrals W and V exist. (i) Differentiating the identities in equations 1 and 3 produce, respectively,

$$\frac{dW}{dt} = U' - (\theta + \lambda) U + \lambda W \text{ and}$$
$$\frac{dV}{dt} = U' - \tilde{\lambda} U + \left(\tilde{\lambda} - \theta\right) V$$

For arbitrary utility streams $\{u_\tau\}_{\tau=0}^{\infty}$, the solutions to these two equations (W and V, respectively) are identical if and only if $\tilde{\lambda} = \lambda + \theta$. Part (ii) uses part (i) and the fact that for $D(t)$ given by equation (2), $\frac{dD}{d\lambda} < 0$ for $t > 0$. ■

Proof. (Proposition 2) Treating (provisionally) the function $F(\mathbf{S}, x_i)$ as given, the constraint and objective in equation 7 are (apart from obvious notational differences) identical to the constraint and payoff in equations 1 and 2 of Karp (2007). In addition, the discount function in equation 2 is (for $\lambda \leq r$) a special case of the discount factor used in Remark 2 of Karp (2007); equation 9 above merely reproduces equation 7 of that paper, taking into account the different notation. Equation 8 then follows from Remark 1 of that paper. ■

Remark 3 *Unless $n = 1$, the function $F(\mathbf{S}, x_i)$ involves the endogenous function χ, whereas in Karp (2007) (where $n = 1$) the corresponding function is given. This difference is important in computing an equilibrium χ, but it has no effect on the statement of the "auxiliary" problem used to compute that equilibrium.*

Proof. (Proposition 3) I use a Taylor expansion to evaluate $K'(\mathbf{S}_\infty)$. This information, together with the necessary conditions for the auxiliary control problem, evaluated at the steady state, and the requirement that the steady state is locally asymptotically stable, produces a set of $\mathbf{S}$, each element of which can be supported as a locally stable steady state in a MPE. I then show that for $n = 1$, the infimum of this set approaches the GGR as $\lambda \to 0$. ■

Proof. (Proposition 4) The argument proceeds by showing that if all other coalitions, and if future planners in one's own coalition, use state-independent decision rules, then the solution to the auxiliary control problem is linear in the state, with value function and control rule given in the proposition. ■

Corollaries 1 – 3 involve straightforward calculation.

Proof. (Lemma 1) Part (i) parallels the method that Tsutsui and Mino (1990) use to find non-linear equilibria in a linear quadratic differential game with constant discounting. However, the endogenous function $K(s)$ in the game with non-constant discounting creates an additional dimension, ruling out Tsutsui and Mino's use of two-dimensional phase portrait analysis. Extending their approach to the non-constant discounting setting, one manipulates the necessary conditions to the auxiliary control problem (equation 8) to obtain a two-dimensional system of ordinary differential equations that a stationary differentiable MPE must solve. There is some latitude in choosing which two variables to use for this system, but a natural choice consists of emissions (or the tax) and the annuity function K, both expressed as a function of the state variable. The condition that the steady state of the resulting equilibrium state trajectory be locally asymptotically stable identifies an open interval; the steady state to a stable MPE lies in this interval. Values of state variables in this interval, and the corresponding steady state values of the control variable, are boundary conditions for the system of ODEs. ■

Proof. (Proposition 5) This proposition relies on Lemma 1.i and routine but tedious calculations. ■

B Intended for Online Publication only

This material contains supplementary material. The first part of this appendix contains additional details for the proofs that are only sketched in the main text. The second part elaborates on three issues mentioned in the text.

B.1 Proofs

Proposition 3 requires the following lemma.

Lemma 2 *For any bounded utility flow $u(t)$ that converges to $u_\infty \neq 0$, and given the discount factor under paternalistic altruism where lifetime is exponentially distributed (equation (2)),* $\lim_{\lambda\to 0}\left(\frac{\lambda}{u_\infty}\int_0^\infty D(t)u(t)dt\right) = \frac{\theta}{\gamma}$.

Proof. (Lemma 2) For small $\varepsilon > 0$ define the τ as the smallest time beyond which $\left|\frac{U(t)-U_\infty}{U_\infty}\right| \leq \varepsilon$. That is

$$\tau = \inf_t \left\{ t : \left| \frac{U(\tau) - U_\infty}{U_\infty} \right| \leq \varepsilon \forall \tau \geq t \right\}.$$

Note that $\tau < \infty$. Use

$$\begin{aligned} &\lim_{\lambda\to 0}\left(\frac{\int_0^\infty D(t)U(t)dt}{\frac{U_\infty}{\lambda}}\right) \\ &= \lim_{\lambda\to 0}\frac{\lambda\left[\left(\int_0^\tau D(t)U(t)dt\right)+\left(U_\infty\int_\tau^\infty D(t)dt\right)+\left(\int_\tau^\infty U(t)D(t)dt - U_\infty\int_\tau^\infty D(t)dt\right)\right]}{U_\infty}. \end{aligned} \tag{17}$$

Consider each of the three terms on the right side of this equation. The fact that $\tau < \infty$ implies that

$$\lim_{\lambda\to 0}\lambda\int_0^\tau D(t)U(t)dt = 0$$

A calculation confirms that

$$\int_\tau^\infty \left(\left(\frac{\lambda - r}{\lambda - \gamma}\right)e^{-\gamma t} - \frac{\theta}{\lambda-\gamma}e^{-\lambda t}\right)dt = \frac{-e^{-\gamma\tau}\lambda^2 + e^{-\gamma\tau}\lambda r + \theta e^{-\lambda\tau}\gamma}{(-\lambda+\gamma)\gamma\lambda}.$$

Taking the limit as $\lambda \to 0$ of this expression, implies that the second term on the second line of equation (17) equals $\frac{\theta}{\gamma}$. By definition of τ,

$$\left| \frac{\int_{\tau}^{\infty} U(t)D(t)dt - U_{\infty} \int_{\tau}^{\infty} D(t)dt}{U_{\infty}} \right| < \varepsilon \int_{\tau}^{\infty} D(t)dt.$$

The limit as $\lambda \to 0$ of the last expression is $\varepsilon \frac{\theta}{\gamma}$. ■

Proof. (Proposition 3) I first derive the necessary and sufficient condition, for general n, that must be satisfied at a stable steady state in a differentiable MPE. I then specialize to $n = 1$ and show that the boundary of the open interval of states that satisfies this condition is arbitrarily close to the GGR for λ close to 0. Because I am interested in the case where λ is small, I assume throughout that $\lambda < r$.

Denote agent i's policy function as $\chi(\mathbf{S})$ and the aggregate decision as $\Psi \equiv n\chi$, so $\Psi' = n\chi'$. Define

$$z = \left(f_S + f_{\Psi}\Psi'\right)_{|\infty},$$

where the subscript ∞ denotes that the function is evaluated at a steady state. Stability requires $z < 0$. For $S_t \approx S_{\infty}$, a first order approximation gives

$$S_{t+\tau} = e^{z\tau} S_t + S_{\infty}\left(1 - e^{z\tau}\right) + o\left(S_t - S_{\infty}\right) \Longrightarrow \frac{dS_{t+\tau}}{dS_t} \approx e^{z\tau} \tag{18}$$

for $\tau \geq 0$. Equation (9) implies

$$\begin{gathered} K'\left(S_t\right) = \\ (r-\lambda) \int_0^{\infty} e^{-\gamma\tau} \left(u_S(S_{t+\tau}, \chi\left(S_{t+\tau}\right)) + u_x(S_{t+\tau}, \chi\left(S_{t+\tau}\right))\chi'\left(S_{t+\tau}\right)\right) \frac{dS_{t+\tau}}{dS_t} d\tau. \end{gathered} \tag{19}$$

Using equation (18) and evaluating equation (19) at $S_t = S_{\infty}$ gives

$$\begin{aligned} K'\left(S_{\infty}\right) &= (r-\lambda)\left(u_S + u_x\chi'\right)_{|\infty} \int_0^{\infty} e^{-\gamma t} e^{zt} dt \\ &= \frac{(r-\lambda)\left(u_S + u_x\chi'\right)_{|\infty}}{\gamma - z} = \frac{(r-\lambda)\left(u_S + u_x\frac{\Psi'}{n}\right)_{|\infty}}{\gamma - z}. \end{aligned} \tag{20}$$

The Hamiltonian corresponding to the fictitious optimal control problem in equation (8) is

$$H = u(S,x) - K(S) + \mu f\left(S, x + \frac{n-1}{n}\Psi(S)\right),$$

where μ is the current value costate variable. The necessary conditions for optimality are

$$u_x + \mu f_x = 0 \Longrightarrow \mu = -\frac{u_x}{f_x} \text{ and } \dot{\mu} = \lambda\mu - \left(u_S - K' + \mu\left(f_S + \frac{n-1}{n}f_x\Psi'\right)\right).$$

Using the first necessary condition and evaluating the costate equation at a steady state (setting $\dot{\mu} = 0$) gives the condition

$$\begin{gathered}\left[-u_S + K' + \tfrac{u_x}{f_x}\left(f_S + \tfrac{n-1}{n}f_x\Psi' - \lambda\right)\right]_{|\infty} = \\ \left[-u_S + \frac{(r-\lambda)\left(u_S + u_x\frac{\Psi'}{n}\right)}{\gamma - z} + \tfrac{u_x}{f_x}\left(f_S + \tfrac{n-1}{n}f_x\Psi' - \lambda\right)\right]_{|\infty} = 0,\end{gathered} \tag{21}$$

where the first equality uses equation (20). Using the definition of z and rearranging the second line of equation (21) implies that $\Psi' = \Psi'(S_\infty)$ is a solution to the quadratic equation

$$Q \times (\Psi')^2 + L \times \Psi' + C = 0 \tag{22}$$

with[16]

$$\begin{gathered}Q \equiv u_x\tfrac{n-1}{n}f_x \\ L \equiv \left((r-\lambda)\tfrac{1}{n} + \tfrac{n-1}{n}(\gamma - f_S) - \left((f_S - \lambda) - \tfrac{u_S}{u_x}f_x\right)\right)u_x \\ C \equiv (-\lambda - \theta + f_S)\,u_S + \tfrac{u_x}{f_x}(f_S - \lambda)(\gamma - f_S)\end{gathered}$$

Hereafter I set $n = 1$, so

$$\Psi'_{\lfloor\infty} = \frac{(\lambda + \theta - f_S)\frac{u_S}{u_x} - \frac{1}{f_x}(f_S - \lambda)(\gamma - f_S)}{r - f_S + \frac{u_S}{u_x}f_x} \Longrightarrow$$

[16]The function C defined here is unrelated to consumption, used in the section on climate policy.

$$z = f_S + f_x \frac{(\lambda+\theta-f_S)\frac{u_S}{u_x} - \frac{1}{f_x}(f_S-\lambda)(\gamma-f_S)}{r - f_S + \frac{u_S}{u_x} f_x} = \frac{\theta\left(\frac{u_S}{u_x} - \frac{f_S}{f_x}\right) f_x + \lambda\left(\frac{u_S}{u_x} - \frac{f_S}{f_x} + \frac{\gamma}{f_x}\right) f_x}{r + \left(\frac{u_S}{u_x} - \frac{f_S}{f_x}\right) f_x}. \tag{23}$$

The GGR is a solution to

$$\left(\frac{u_S}{u_x} - \frac{f_S}{f_x}\right) u_x = 0.$$

I define the state and the control variables so that $u_S < 0$ and $u_x < 0$. For example, in the climate model, S is the stock of atmospheric carbon and x is the level of abatement, so the flow of utility is decreasing in both variables. These definitions (the state variable is a "bad" and the action is costly) mean that the model is sensible if and only if $f_x < 0$ (so that incurring a cost reduces the public bad). Given the concavity of the static optimization problem (which determines the GGR), a stock level slightly greater than the GGR satisfies

$$\left(\frac{u_S}{u_x} - \frac{f_S}{f_x}\right) u_x = \varepsilon < 0 \ \text{ or } \left(\frac{u_S}{u_x} - \frac{f_S}{f_x}\right) = \frac{\varepsilon}{u_x} > 0, \tag{24}$$

for ε small in absolute value. Such a stock level yields approximately the maximum steady state level of utility. (Given that the costly action x reduces the stock, it would never be part of an equilibrium to drive the stock below the optimal static level.)

Using equation (24) in (23) gives

$$z = \frac{\theta \frac{\varepsilon}{u_x} f_x + \lambda\left(\frac{\varepsilon}{u_x} + \frac{\gamma}{f_x}\right) f_x}{r + \frac{\varepsilon}{u_x} f_x}.$$

The denominator is positive for small ε. For ε small in absolute value (so that $\frac{\varepsilon}{u_x} f_x + \gamma > 0$), the numerator is negative if and only if

$$\frac{-\theta \frac{\varepsilon}{u_x} f_x}{\left(\frac{\varepsilon}{u_x} f_x + \gamma\right)} > \lambda,$$

i.e. if and only if λ is sufficiently small, as was to be shown. ■

Proof. (Proposition 4) Because P is of full rank, $B = P\Lambda P^{-1}$. In a symmetric MPE, i.e. where all coalitions emit $x(t) = \chi(t)$, $\dot{\mathbf{S}}(t) = BS(t) + bn\chi(t)$. Here, the equilibrium value of the state t periods in the future, given the current value S_0 is: $S_t = Pe^{\Lambda t}P^{-1}S_0 + H(t)$, where $H(t)$ depends on the trajectory of controls from time 0 to t. If χ is a constant, then $H(t) = P\Omega(t)Pn^{-1}(bn\chi)$. Because all eigenvalues are negative, $\Omega(t)$ is a diagonal matrix with element $\frac{\exp(\Lambda_i t-1)}{\Lambda_i}$ in the i'th diagonal position.

Under the assumption that the policy maker in coalition i at time t expects future emissions to be independent of the state, and using the flow payoff $v(x,t;n) - \frac{\kappa s}{n} = \frac{1}{n}[v(nx,t;1) - \kappa i_1' S]$ and equation (9), the annuity function is:

$$K(S_t,t) = \tfrac{(r-\lambda)}{n}\int_0^\infty e^{-\gamma\tau}\left[\left(v\left(n\chi\left(t+\tau\right),t+\tau;1\right) - \kappa i_1' S_{t+\tau}\right)\right]d\tau$$
$$= \tfrac{(r-\lambda)}{n}\int_0^\infty e^{-\gamma\tau}\left[\left(v\left(n\chi\left(t+\tau\right),t+\tau;1\right) - \kappa i_1'\left(Pe^{\Lambda\tau}P^{-1}S_t + H\left(t+\tau\right)\right)\right)\right]d\tau.$$

From this formula, it is apparent that K is linear in S, $K(S_t,t) = q_{0,t} + q' S_t$, with

$$q' = -\kappa\frac{r-\lambda}{n}\tilde{q}',$$

where $\tilde{q}'$ is given by the last equation in (12). If future policies are constant, and v does not depend on time, then $q_{0,t}$ is a constant, q_0.

Using $K = q_{0,t} + q'S$ and the utility flow $v(x,t;n) - \frac{\kappa s}{n}$ in equation (8), produces the dynamic programming equation (DPE)

$$\lambda J(S,t) = \max_x\{v(x,t;n) - \tfrac{\kappa}{n} i_1' S$$
$$-\left(q_{0,t} + q'S\right) + J_S'(S,t)\left[BS + b\left(x + (n-1)\chi_t\right)\right]\}.$$

Because this problem is linear in the state, the obvious trial solution is a linear function, $J(S,t) = g_{0,t} + g_t' S$. Using this trial solution, the DPE becomes

$$\begin{aligned} \lambda\left(g_{0,t} + g_t' S\right) &= \max_x \left\{ v\left(x,t;n\right) + g_t' b x \right\} \\ -\tfrac{\kappa}{n} i_1' S - \left(q_{0,t} + q' S\right) &+ g_t' \left[BS + b\left(n-1\right)\chi \right]. \end{aligned} \tag{25}$$

The first order condition (which is sufficient due to concavity of v) is

$$\frac{\partial v\left(x,t;n\right)}{\partial x} + g_t' b = 0. \tag{26}$$

The solution, x^*, possibly depends on time, but is independent of the state. Substituting the optimal flow payoff into the DPE gives the maximized DPE

$$\begin{aligned} \lambda\left(g_{0,t} + g_t' S\right) &= v\left(x^*(t),t;n\right) + g_t' b x^* \\ -\tfrac{\kappa}{n} i_1' S - \left(q_{0,t} + q' S\right) &+ g_t' \left[BS + b\left(n-1\right)\chi \right]. \end{aligned}.$$

Equating coefficients of S gives

$$\lambda g_t' = -\frac{\kappa}{n} i_1' - q' + g_t' B.$$

Because B and q' are constants, g_t' is also a constant, g', given by the first equation in (12). Equation 25 and the fact that g is a constant implies equation 13. If v is independent of t, then equilibrium emissions are also constant, in which case, q_0, and g_0 are also constants.

In summary, regardless of planner i,t's beliefs about other planners' state-independent policies, planner i,t's optimal policy is given by equation 13. Because g is independent of other planners' policies, the equilibrium policy is dominant both respect to actions by future planners in one's own coalition, and by all current and future planners in other coalitions. ■

Proof. (Corollary 1) The first order condition for the problem in equation (13) is $\frac{dv(x;n)}{dx} = -g\left(n\right)' b$, where I make the dependence of g on n explicit for emphasis, and I drop the argument t in v to simplify notation. By concavity of v, this first order condition is sufficient. Using the definition $v(x;n) = \frac{1}{n}v(nx;1)$, and the chain rule, I have $\frac{dv(x;n)}{dx} = \frac{1}{n}\frac{dv(nx;1)}{d(nx)}\frac{dnx}{dx} = \frac{dv(X(n);1)}{dX}$. Using this relation in the first order condition gives $\frac{dv(X(n);1)}{dX} = -g(n)' b$. Using the definition of $g(n)'$ from equation (12) in this first order condition gives equation (14).

By inspection, the absolute value of the elasticity of this tax with respect to n is 1. In order to obtain equation (15), use

$$\frac{d\tau}{d\lambda} = -\frac{\kappa}{n}\frac{d\left[\left(i_1' - (r-\lambda)\tilde{q}'\right)(B-\lambda I)^{-1}b\right]}{d\lambda} =$$
$$-\frac{\kappa}{n}\left[\tilde{q}'(B-\lambda I)^{-1}b + \left(i_1' - (r-\lambda)\tilde{q}'(B-\lambda I)^{-2}b\right)\right] =$$
$$-\frac{\kappa}{n}\left[\tilde{q}' + \left(i_1' - (r-\lambda)\tilde{q}'(B-\lambda I)^{-1}\right)\right](B-\lambda I)^{-1}b.$$

Multiplying this expression by $-\frac{\lambda}{\tau}$ to convert to an absolute value elasticity, gives equation (15). ■

Proof. (Corollary 2) For the GeaS model, where $v(x;n) = \frac{\alpha}{n}\ln(nx)$, the first order condition for the problem in (13) is $\frac{\alpha}{nx} + g'b = 0 \Longrightarrow x = -\frac{\alpha}{ng'b} \Longrightarrow nx = X = -\frac{\alpha}{g'b} = \frac{-\alpha n}{\kappa\left(i_1' - (r-\lambda)\tilde{q}'\right)(B-\lambda I)^{-1}b}$. Emissions are positive ($g'b < 0$), so aggregate emissions are an increasing linear function of n.

For the quadratic model, $v(x,t;n) = a_t x - \frac{d_t n}{2}x^2$. Suppressing the time subscripts, the first order condition for the problem in (13) is $a - dnx + g'b = 0 \Longrightarrow x = \frac{a+g'b}{dn} \Longrightarrow nx = X = \frac{a+g'b}{d}$. Using $\frac{d(g'b)}{dn} = -\frac{\kappa}{n^2}\left(i_1' - (r-\lambda)\tilde{q}'\right)(B-\lambda I)^{-1}b > 0$, aggregate emissions is an increasing strictly concave function of n. ■

Proof. (Corollary 3) For the scalar case, $\tilde{q} = \frac{1}{\gamma - B}$. Using this result in equation (14) produces the formula for τ^* in the scalar case. Straightforward calculation establishes the other claims in the corollary. ■

Proof. (Lemma 1) Here the state is a scalar, s. I use the necessary conditions for a differentiable MPE to find two ordinary differential equations (ODEs) in s, with dependent variables X and K (emissions and the annuity function). For ease of interpretation, I then translate the ODEs in X and K into equivalent ODEs in τ (the tax in units of utility) and K. I then use the stability condition to find the set of stable steady states in a MPE.

Construct the ODEs in X and K. For a given symmetric decision rule, $\chi(s)$ define aggregate emissions as $\digamma(s) = n\chi(s)$. In the scalar linear model, $\dot{s} = Bs + \sum_i x_i$. With $\sum_{j \neq i} x_j = \frac{n-1}{n}\digamma(s)$, a representative coalition (so I drop the dynastic index) faces the equation of motion

$$\dot{s} = F(s, x) = Bs + \frac{n-1}{n}\digamma(s) + x.$$

The GeaS model allows non-stationarity, entering via the time-dependent term A_t. Thus, the annuity function, K, also has time as an argument. This fact requires a slight change in notation, but because the non-stationarity enters additively, it does not complicate the derivations. In the GeaS model, $v(x; n) = \frac{\ln A_t}{n} + \frac{\alpha}{n}\ln(nx)$. To incorporate the non-stationarity, I replace the definition of the annuity, equation 9, by

$$\begin{aligned} \tilde{K}(s_t, t) &= (r-\lambda)\int_0^\infty e^{-\gamma\tau}\left(v(x_{t+\tau}; n) - \tfrac{\kappa}{n}s_{t+\tau}\right)d\tau \\ &= (r-\lambda)\int_0^\infty e^{-\gamma\tau}\left(\tfrac{\ln A_{t+\tau}}{n} + \tfrac{\alpha}{n}\ln(nx_{t+\tau}) - \tfrac{\kappa}{n}s_{t+\tau}\right)d\tau \\ &= \bar{A}(t) + K(s_t), \end{aligned}$$

which uses the definitions

$$\begin{aligned} \bar{A}(t) &= (r-\lambda)\int_0^\infty e^{-\gamma\tau}\left(\tfrac{\ln A_{t+\tau}}{n}\right)d\tau \\ K(s_t) &= (r-\lambda)\int_0^\infty e^{-\gamma\tau}\left(\tfrac{\alpha}{n}\ln(nx_{t+\tau}) - \tfrac{\kappa}{n}s_{t+\tau}\right)d\tau. \end{aligned}$$

In a symmetric equilibrium, the function K satisfies the ODE

$$\begin{aligned} \gamma K(s) &= (r-\lambda)\left(\tfrac{\alpha}{n}\ln(X) - \tfrac{\kappa}{n}s\right) + K'(s)(Bs + X) \Rightarrow \\ K'(s) &= \frac{\gamma K(s) - (r-\lambda)\left(\frac{\alpha}{n}\ln(X) - \frac{\kappa}{n}s\right)}{Bs + X}. \end{aligned} \tag{27}$$

The second line of equation 27 is the ODE for K. The numerator and denominator are both 0, evaluated at a steady state. I therefore use a Taylor expansion (below) to find the value of K' at a steady state.

The auxiliary control problem for the GeaS model (see equation 8) is

$$\tilde{J}(s_t,t) = \max \int_0^\infty e^{-\lambda\tau}\left(\tfrac{\ln A_{t+\tau}}{n} + \tfrac{\alpha}{n}\ln\left(nx_{t+\tau}\right) - \tfrac{\kappa}{n}s_{t+\tau} - \tilde{K}(s_{t+\tau})\right)d\tau$$
$$\text{subject to } \dot{s} = F(s,x)$$
$$= \int_0^\infty e^{-\lambda\tau}\left(\tfrac{\ln A_{t+\tau}}{n} - \bar{A}_{t+\tau}\right)d\tau + \max\int_0^\infty e^{-\lambda\tau}\left(\tfrac{\alpha}{n}\ln\left(nx_{t+\tau}\right) - \tfrac{\kappa}{n}s_{t+\tau} - K(s_{t+\tau})\right)d\tau$$
$$\text{subject to } \dot{s} = F(s,x).$$

Defining

$$J\left(s_t\right) = \tilde{J}(s_t,t) - \int_0^\infty e^{-\lambda t}\left(\frac{\ln A_t}{n} - \tilde{A}\right)dt,$$

produces the stationary auxiliary control problem

$$J\left(s\right) = \max\int_0^\infty e^{-\lambda t}\left(\frac{\alpha}{n}\ln\left(nx\right) - \frac{\kappa}{n}s - K(s)\right)dt \quad \text{subject to } \dot{s} = F(s,x).$$

Defining μ as the costate variable for s, the Hamiltonian and necessary conditions to this problem (taking the function K and F as given for the time being) are

$$H = \max_x\left[\tfrac{\alpha}{n}\ln\left(nx\right) - \tfrac{\kappa}{n}s - K(s) + \mu\left(Bs + \tfrac{n-1}{n}F\left(s\right) + x\right)\right]$$
$$\begin{aligned} &\tfrac{\alpha}{n}\tfrac{1}{nx}n + \mu = 0 \\ \dot{\mu} &= \lambda\mu + \tfrac{\kappa}{n} + K'\left(s\right) - \mu\left(B + \tfrac{n-1}{n}F'\left(s\right)\right) \\ &= \tfrac{\kappa}{n} + K'\left(s\right) + \mu\left(\lambda - B - \tfrac{n-1}{n}F'\left(s\right)\right). \end{aligned} \tag{28}$$

Evaluating the necessary conditions at a symmetric equilibrium (replacing F' with X') gives

$$\begin{aligned} \mu &= -\tfrac{\alpha}{X} \\ \dot{\mu} &= \tfrac{\kappa}{n} + K'\left(s\right) + \mu\left(\lambda - B - \tfrac{n-1}{n}X'\left(s\right)\right) \\ &= \tfrac{\kappa}{n} + K'\left(s\right) - \tfrac{\alpha}{X}\left(\lambda - B - \tfrac{n-1}{n}X'\left(s\right)\right). \end{aligned} \tag{29}$$

Differentiating the first equation with respect to time, using the second, gives

$$\frac{\alpha}{X^2}\dot{X} = \frac{\kappa}{n} + K'\left(s\right) - \frac{\alpha}{X}\left(\lambda - B - \frac{n-1}{n}X'\left(s\right)\right).$$

Dividing this equation by $\dot{s} = Bs + X$, using $\frac{\dot{X}}{\dot{s}} = \frac{dX}{ds} = X'(s)$ gives

$$\begin{aligned}
\frac{\frac{\alpha}{X^2}\dot{X}}{\dot{s}} &= \frac{\frac{\kappa}{n} + K'(s) - \frac{\alpha}{X}\left(\lambda - B - \frac{n-1}{n}X'(s)\right)}{Bs + X} \\
\frac{\alpha}{X^2}\frac{\dot{X}}{\dot{s}} &= \frac{\alpha}{X^2}X' = \frac{\left[\frac{\kappa}{n} + K'(s) - \frac{\alpha}{X}\left(\lambda - B - \frac{n-1}{n}X'(s)\right)\right]}{Bs + X}.
\end{aligned}$$

Solving for X' gives

$$X' = \frac{\left[\frac{\kappa}{n} + K'(s) - \frac{\alpha}{X}(\lambda - B)\right]}{\alpha\left(Bs + \frac{1}{n}X\right)}X^2. \tag{30}$$

A differentiable MPE solves the ODEs 27 and 30.

Translation from X to τ. In the GeaS model, $\tau = \frac{\alpha}{X}$, or $X = \frac{\alpha}{\tau}$, so $\frac{dX}{ds} = -\frac{\alpha}{\tau^2}\frac{d\tau}{ds}$. Using this fact in equations 27 and 30 produces

$$\frac{d\tau}{ds} = -\frac{\left(\frac{\kappa}{n} + K'(s)\right)\tau - (\lambda - B)\tau^2}{\left(Bs\tau + \frac{\alpha}{n}\right)} \text{ and } \frac{dK}{ds} = \frac{\gamma K(s) - (r - \lambda)\left(\frac{\alpha}{n}\ln\left(\frac{\alpha}{\tau}\right) - \frac{\kappa}{n}s\right)}{Bs + \frac{\alpha}{\tau}}. \tag{31}$$

Find the set of boundary conditions for these ODEs. A steady state, a triple s_∞, X_∞ and K_∞, is a feasible boundary condition for the ODEs if it is locally asymptotically stable. Denote Φ as the value of X' at a steady state, i.e. where $0 = Bs + X$, and define $z = B + \Phi$. Stability requires $z < 0$. Equation 20, repeated here, is

$$K'(s_\infty) = \frac{(r - \lambda)\left(u_s + u_x\frac{F'}{n}\right)_{|\infty}}{\gamma - z}. \tag{32}$$

In the GeaS model, $u = v(x; n) - \frac{\kappa}{n}s = \frac{\alpha}{n}\ln(nx) - \frac{\kappa}{n}s$, so $u_s = -\frac{\kappa}{n}$ and $u_x = \frac{\alpha}{n}\frac{1}{nx}n = \frac{\alpha}{X}$. Using these results and equation 32 gives

$$\begin{aligned}
K'(s_\infty) &= \left(\frac{(r-\lambda)\left(-\frac{\kappa}{n} + \frac{\alpha}{X}\frac{X'}{n}\right)}{\gamma - B - X'}\right)_{|\infty} \\
&= \left(\frac{(r-\lambda)\left(-\frac{\kappa}{n} + \frac{\alpha}{-Bs}\frac{\Phi}{n}\right)}{\gamma - B - \Phi}\right)_{|\infty},
\end{aligned} \tag{33}$$

where the second equality uses the fact that $X = -Bs$ at a steady state, and the notation $\Phi = X'_{|\infty}$.

Using equation 29 and setting $\dot{\mu} = 0$ gives

$$0 = \frac{\kappa}{n} + K'(s) - \frac{\alpha}{X}\left(\lambda - B - \frac{n-1}{n}X'(s)\right).$$

Using this equation and the second line of equation 33, and $X = -Bs$ at a steady state, I obtain

$$\begin{aligned} 0 &= \tfrac{\kappa}{n} + K'(s) - \tfrac{\alpha}{-Bs}\left(\lambda - B - \tfrac{n-1}{n}\Phi\right) \\ &= \tfrac{\kappa}{n} + \left[\tfrac{(r-\lambda)\left(-\frac{\kappa}{n} + \frac{\alpha}{-Bs}\frac{\Phi}{n}\right)}{\gamma - B - \Phi}\right] - \tfrac{\alpha}{-Bs}\left(\lambda - B - \tfrac{n-1}{n}\Phi\right). \end{aligned}$$

Rearranging this equation gives

$$0 = Q\Phi^2 + L\Phi + h,$$

using the definitions

$$\begin{aligned} Q = \tfrac{1}{Bn}\alpha(n-1), \quad L = \left(-\tfrac{1}{n}\kappa\right)s + \alpha\tfrac{(1-n)\lambda + B(2n-1) + (\gamma(1-n) - r)}{Bn} \\ h = -\tfrac{1}{n}\kappa(B - \theta - \lambda)s + \tfrac{1}{B}\alpha(B-\lambda)(B-\gamma). \end{aligned} \tag{34}$$

with boundary condition

$$(s_\infty, \tau_\infty, K_\infty) = \left(s_\infty, -\frac{\alpha}{Bs_\infty}, \frac{(r-\lambda)}{\gamma n}\left[\alpha \ln(-Bs_\infty) - \kappa s_\infty\right]\right) \tag{35}$$

for $s_\infty \in \Delta$, where[17]

$$\Delta = \left\{ s \,\middle|\, \exists \Phi \text{ for which } B + \Phi < 0 \wedge Q\Phi^2 + L\Phi + h = 0 \right\}. \tag{36}$$

The relation between Υ *and* Φ. Solving $Q\Phi^2 + L\Phi + h = 0$ for the steady state s and dividing by the GGR gives the ratio Υ in the first line of equation 16. Evaluating this ratio at $\Phi = 0$ and $\Phi = -B$ establishes the second line of equation 16, thus establishing part (i) of the proposition.

*Part (ii):*The derivative of Υ with respect to Φ is

$$\begin{gathered} \frac{d\Upsilon}{d\Phi} = \frac{G}{B(B+\Phi-\theta-\lambda)^2} \text{ with} \\ G(\Phi) = (n-1)\Phi^2 + 2(n-1)(B-\theta-\lambda)\Phi + \\ \left((B-\theta)^2 + \left(\theta\lambda + \lambda^2 - 2B\lambda\right)\right)(n-1) + \theta(nr-\lambda). \end{gathered} \tag{37}$$

G is convex in Φ, and calculation confirms that

$$\begin{gathered} G(0) > 0,\ G(-B) > 0 \\ \text{and } \frac{dG}{d\Phi} < 0 \text{ for } 0 \leq \Phi \leq -B \end{gathered}$$

These inequalities confirm that G is positive and decreasing for $\Phi \in [0, -B]$. The first line of equation 37 then implies that $\frac{d\Upsilon}{d\Phi} < 0$ for $\Phi \in [0, -B]$. Using this derivative and the expression for Υ at $\Phi = -B$, from part (i), confirms part (ii) of the proposition. ■

Proof. (Proposition 5) Part (i) evaluates the derivatives of $\Upsilon(\Phi; \cdot)$, given in Lemma 1.i, with respect to n and λ. Part (ii) evaluates the cross partial derivative to establish $\frac{d^2\Upsilon}{dnd\lambda} > 0$. Part (iii) notes that $\Upsilon(-B) \to 1$ as $\lambda \to 0$, for any finite n. In contrast, at $n = 1$, $\Upsilon(-B)$ remains bounded away for 1 for λ bounded away from 0.

[17]As consistency checks, note that for $n = 1$ and $\lambda = r$, where the game collapses to a standard control problem, Δ is a singleton, with $s = \frac{\alpha(B-r)}{B\kappa}$. As required, this is the steady state in the control problem ($n = 1$ and constant PRTP, r). The GGR steady state ($s = \arg\max_s \alpha \ln X - \kappa s$, subject to $0 = s + BX$) equals the limit of this value as $r \to 0$; here the GGR steady state is $\frac{\alpha}{\kappa}$.

Negative values of Φ are feasible, but these yield higher, i.e. less cooperative, steady state stocks. I do not consider these, because it seems that agents should be able to coordinate on an equilibrium at least as good as the unique limit equilibrium, corresponding to $\Phi = 0$.

■

B.2 Additional supporting material

I provide the necessary conditions for the case not considered in the text, discuss calibration of the two dimensional GeaS model, and then provide some comments on robustness.

B.2.1 Equilibrium conditions

There are two cases under the exponentially distributed lifetime, because $\lim_{t\to\infty}\eta(t)$ depends on whether $\lambda < r$ or $\lambda > r$. For the exponential case with $0 < \lambda \leq r$, and using the differentiability of $J(S)$ (already assumed in deriving the problem comprised of (8) and (9), a necessary condition for the MPE is that

$$x_t = \chi(S_t) \equiv \arg\max\left(u(S_t, x_t) - K(S_t) + J_S(S)\, F(S, x)\right), \tag{38}$$

and that $J(S)$ satisfy the dynamic programming equation

$$\lambda J(S) = \left(u\left(S, \chi(S)\right) - K(S) + J_S(S)\, F\left(S, \chi(S_t)\right)\right). \tag{39}$$

With $\lambda > r$, where $\lim_{t\to\infty}\eta(t) = \gamma$, the auxiliary control problem is

$$J(S) = \max \int_0^\infty e^{-\gamma t}\left(u(S_t, x_t) - K(S_t)\right) d\tau \quad \text{subject to } \dot{S} = F(S, x), \tag{40}$$

with the side condition (definition):

$$K(S_t) \equiv \int_0^\infty D(\tau)\left(\eta(\tau) - \gamma\right) u(S^*_\tau, \chi(S^*_\tau)) d\tau. \tag{41}$$

Equation (2) and the first line of equation (4) imply $D(t)(\eta(t) - \gamma) = -\theta e^{-\lambda t}$ so equation (41) simplifies to

$$K(S_t) = -\theta \int_0^\infty e^{-\lambda\tau} u\left(S^*_{t+\tau}, \chi\left(S^*_{t+\tau}\right)\right) d\tau. \tag{42}$$

The integral in equation (42) is the present discounted value of the equilibrium future flow of payoff, computed using the discount rate λ. Thus, $-K(S_t)$ is an annuity, which if received in perpetuity and discounted at θ (the constant birth = death rate), equals the value of this future stream of payoff. The flow payoff in the fictitious control problem equals the flow playoff in the original model, plus this annuity. A necessary condition for the MPE is that

$$x_t = \chi\left(S_t\right) \equiv \arg\max\left(u(S_t, x_t) - K(S_t) + J_S\left(S\right) F(S, x)\right), \tag{43}$$

and that $J(S)$ satisfy the dynamic programming equation

$$\gamma J(S) = \left(u(S, \chi\left(S\right)) - K(S) + J_S\left(S\right) F(S, \chi\left(S_t\right))\right). \tag{44}$$

B.2.2 The two-dimensional GeaS model

The two-dimensional model consists of two state variables, the scalars s and $\tilde{S}$. I define $\tilde{S}$ as the difference between actual and pre-industrial atmospheric CO_2 measured in Teratons (Tt CO_2). The emissions flow, X, increases this stock, which decays at a constant rate, υ, but has no immediate effect on s. Because I use only two state variables to represent complicated dynamics, the scalar s is only a proxy for temperature change; unlike the carbon stock ($\tilde{S}$), s does not have a simple physical meaning. The proxy $s \leq 0$ responds to increased $\tilde{S}$ with a lag. With multiplicative exponential damages and the normalization $\kappa = -1$, the fractional consumption loss at t due to climate change is $1 - e^{s(t)}$.

With these assumptions, the parameters in equation 10 are

$$B = \begin{pmatrix} \varsigma & \varrho \\ 0 & \upsilon \end{pmatrix} \text{ and } b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The matrix of eigenvalues and eigenvectors corresponding to B are

$$\Lambda = \begin{pmatrix} \upsilon & 0 \\ 0 & \varsigma \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} -\frac{\varrho}{\varsigma-\upsilon} & 1 \\ 1 & 0 \end{pmatrix}. \tag{45}$$

Solving the pair of differential equations for s and $\tilde{S}$, with initial conditions $s(0)$ and $\tilde{S}(0)$ and aggregate emissions flow $X(y)$, gives

$$\begin{pmatrix} s(t) \\ \tilde{S}(t) \end{pmatrix} = Pe^{\Lambda t}P^{-1} \begin{pmatrix} s(0) \\ \tilde{S}(0) \end{pmatrix} + P \begin{pmatrix} e^{\upsilon t} \int_0^t e^{-\upsilon y} X(y)\, dy \\ \frac{\varrho}{\varsigma-\upsilon} e^{\varsigma t} \int_0^t e^{-\varsigma y} X(y)\, dy \end{pmatrix}.$$

Using this equation and

$$Pe^{\Lambda t}P^{-1} = \begin{pmatrix} e^{t\varsigma} & \frac{\varrho e^{t\varsigma} - \varrho e^{t\upsilon}}{\varsigma-\upsilon} \\ 0 & e^{t\upsilon} \end{pmatrix}$$

implies that a one unit increase increase in $\tilde{S}(0)$ causes an $E(t) \equiv \frac{\varrho e^{t\varsigma} - \varrho e^{t\upsilon}}{\varsigma-\upsilon}$ increase in $s(t)$ ("$E(t)$" for "effect".)

Using the two-dimensional analog of equation 10 and the formula in equation 14, the equilibrium tax is

$$\tau = -\frac{\varrho}{n}\frac{1}{\lambda-\upsilon} M_1' M_2 \quad \text{with}$$

$$M_1 = \begin{pmatrix} (\varsigma-\gamma)(r-\lambda)+1 \\ \frac{r-\lambda}{(\varsigma-\upsilon)(\gamma-\upsilon)}\left((\gamma-\upsilon)(\varsigma-\gamma)+1\right) \end{pmatrix} \quad M_2 = \begin{pmatrix} \frac{-1}{(\varsigma-\lambda)} \\ 1 \end{pmatrix}. \tag{46}$$

I need three assumptions to calibrate the three parameters ς, ϱ, and υ. I obtain υ using an estimate of the half-life of atmospheric carbon. A half-life of h years implies $\upsilon = \frac{\ln(0.5)}{h}$. Thus, $h = 200$ implies $\upsilon = \frac{\ln(0.5)}{200} = -3.47 \times 10^{-3}$, while $h = 400$ implies $\upsilon = \frac{\ln(0.5)}{400} = -1.733 \times 10^{-3}$. I calibrate ς and ϱ using Gerlagh and Liski (2012), who calibrate a higher dimensional linear model, based primarily on DICE. I adopt their assumptions that: (i) doubling atmospheric stocks (relative to preindustrial levels) reduces output (in my setting, consumption) by 2.6%, once s has adjusted, and (ii) following a pulse increase in atmospheric CO_2 at time 0, the loss rises (from 0) during the first 60 years, and then falls slowly. These three assumptions imply the "baseline parameter values" $\upsilon = -3.47 \times 10^{-3}$, $\varsigma = -4.685 \times 10^{-2}$, and $\varrho = -5.66 \times 10^{-4}$. Note that $E(t) < 0$ for $t > 0$ (The baseline also includes $r = 0.02 = \theta$.)

I now explain these calculations in more detail. Doubling atmospheric stocks relative to pre-industrial levels implies an increase of 280 parts per million by volume (ppm) of CO_2. One ppm corresponds to 2.13 Gigatons of carbon, or 2.13 (3.66) Gigatons of CO_2 or $\frac{2.13(3.66)}{1000}$ Teratons of CO_2 (Tt CO_2). Therefore, an increase of 280 ppm represents an increase of $\frac{2.13(3.66)}{1000}280 \approx 2.18$ Tt CO_2. The steady state level of s_∞ equals $-\frac{\varrho}{\varsigma}\tilde{S}_\infty$. I assume that doubling the steady state of $\tilde{S}$ (relative to preindustrial levels), leads to a 2.6% reduction in output (and consumption). This assumption implies $\left(1 - e^{-\frac{\varrho}{\varsigma}\tilde{S}_\infty}\right) = \left(1 - e^{-\frac{\varrho}{\varsigma}2.18}\right) = 0.026$ or $-\frac{\ln 0.974}{2.18} = \frac{\varrho}{\varsigma}$.

An instantaneous one unit increase in the stock at time 0 leads to a change in $s(t)$ of $E(t) = \frac{\varrho e^{t\varsigma} - \varrho e^{t\upsilon}}{\varsigma - \upsilon}$ units. The increase in percent reduction in output after t years, due to this time-zero increase in $\tilde{S}$, is $100\left(\left(1 - e^{s(t)+E(t)}\right) - \left(1 - e^{s(t)}\right)\right) = 100e^{s(t)}\left(1 - e^{E(t)}\right)$, where $s(t)$ is the equilibrium value of s absent the initial increase in stock $\tilde{S}$. The time profile of $100e^{s(t)}\left(1 - e^{E(t)}\right)$ depends on $s(t)$. For purpose of calibration, I consider the case where $s(t)$ is a constant, as in a steady state. In this case, the increase in loss is maximized where

$$\frac{d\left(1 - e^{E(t)}\right)}{dt} = -e^{E(t)}\frac{dE}{dt} = -e^{E(t)}\varrho\frac{\varsigma e^{t\varsigma} - \upsilon e^{t\upsilon}}{\varsigma - \upsilon} = 0.$$

Following Gerlagh and Liski (2012), I assume that the increase in loss is maximized at $t = 60$ years, giving the calibration equation

$$\frac{\varsigma e^{60\varsigma} - \upsilon e^{60\upsilon}}{\varsigma - \upsilon} = 0.$$

Solving the last equation, using the calibration assumption $\upsilon = -3.47 \times 10^{-3}$, corresponding to a half-life of 200 years, gives $\varsigma = -4.685 \times 10^{-2}$. Using this result and $-\frac{\ln 0.974}{2.18} = \frac{\varrho}{\varsigma}$ implies $\varrho = -5.66 \times 10^{-4}$.

With these values, the additional consumption loss, due to a unit increase in $\tilde{S}$, beginning at a steady state, increases for the first 60 years and then slowly falls. The increased loss after t years, as a fraction of the (maximal) increased loss after 60 years, is

$$\frac{\left(1-\exp\left(\frac{\varrho e^{t\varsigma}-\varrho e^{t\upsilon}}{\varsigma-\upsilon}\right)\right)}{\left(1-\exp\left(\frac{\varrho e^{60\varsigma}-\varrho e^{60\upsilon}}{\varsigma-\upsilon}\right)\right)}. \tag{47}$$

Figure 4 shows the graph of this ratio assuming a half-life of 200 (solid curve) or 400 (dashed curve) years. Given the other two calibration assumptions, the graph is insensitive to the half-life for $t < 60$. Beyond that time, a longer half-life causes the additional stock, and thus the additional damage, to decay more slowly, causing the curve to rotate up.

Comparing Figure 4 with Gerlagh and Liski's (2012) Figure 1 shows that for a half-life of 200 years, the damage trajectory in my two-dimensional model has the same shape as their representation of the DICE results. For a 400 year half-life, the trajectory has the same shape as in their model for the first 600 years or so. The profiles in my two-dimensional model and in their four dimensional model differ in the very long run. Their climate system is closed, whereas for $\upsilon < 0$, emissions and thus damages eventually dissipate in my model. Therefore, $\lim_{t\to\infty} E(t) = 0$ in my setting, whereas it approaches a positive constant in theirs. Their closed climate system would present a problem in the GeaS setting, where equilibrium emissions are constant; with constant emissions, damages in the closed climate system become unbounded. In contrast, in my setting, steady state damages for constant X equals $\frac{\varrho X}{\varsigma\upsilon}$.

I use the expressions for Λ and P in equation 45 to calculate

$$\tilde{q}' = \int_0^\infty i_1' P e^{-(\gamma I-\Lambda)\tau} P^{-1} d\tau = i_1' P \int_0^\infty e^{-(\gamma I-\Lambda)\tau} d\tau P^{-1} =$$
$$\begin{pmatrix} \gamma-\varsigma, & -\frac{\varrho}{(\varsigma-\upsilon)(\gamma-\upsilon)}\left(\varsigma\gamma-\gamma^2-\varsigma\upsilon+\gamma\upsilon+1\right) \end{pmatrix}$$

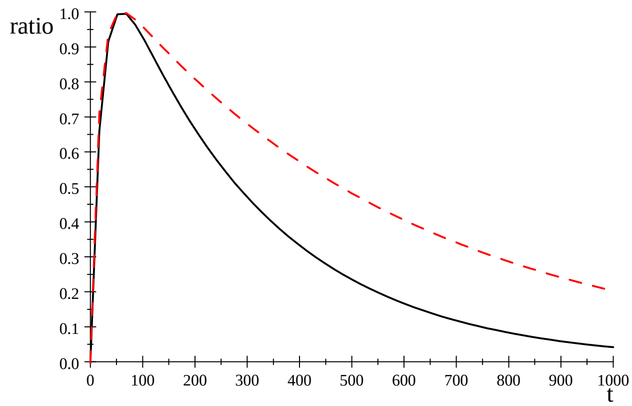

Figure 4: The ratio given in expression 47 for a half life of 200 years (solid curve) and a half life of 400 years (dashed curve). Other two calibration assumption as in the text.

and

$$B-\lambda I=\begin{pmatrix}\varsigma-\lambda & \varrho\\ 0 & \upsilon-\lambda\end{pmatrix}\Rightarrow(B-\lambda I)^{-1}=\begin{pmatrix}\frac{1}{\varsigma-\lambda} & -\frac{\varrho}{\lambda^2-\varsigma\lambda+\varsigma\upsilon-\lambda\upsilon}\\ 0 & -\frac{1}{\lambda-\upsilon}\end{pmatrix}.$$

Using these results in the formula for the tax (with $\kappa=-1$), equation 14, and simplifying the result, gives equation 46.

B.2.3 Robustness

Different OLG structure I first compare the discount rate under the assumption of exponentially distributed lifetime and under a deterministic alternative where agents live for Γ periods. Setting $\Gamma=\frac{1}{\theta}$, the expected lifetime under the exponential distribution, makes the two comparable. When agents have deterministic lifetime, currently living agents have different incentives to invest in a long-lived public good, because some will die sooner than others. Other complications also become important in this setting, such as the possibility of transfers among currently living members of a coalition. For the exponentially distributed lifetime, where all citizens

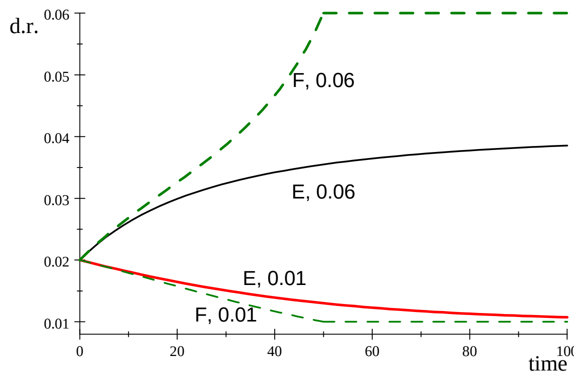

Figure 5: Discount rates (d.r.) for $\theta = 0.02 = r = \frac{1}{\Gamma}$. Solid curves (labelled E) correspond to exponentially distributed lifetime and dashed curves (labelled F) correspond to fixed lifetime. Numerical values in label show value of λ.

are identical, there would be no reason for those transfers.

The discount factor for the deterministic case (with paternalistic altruism) can be calculated under Assumption 1 and the additional assumption that a utilitarian social planner aggregates the preferences of currently living citizens, giving the same weight to each of these. The discrete time analog to this model in which people live for two periods produces quasi-hyperbolic (β, δ) discounting; the continuous time model (and the discrete model where agents live more than two periods) generalizes quasi-hyperbolic discounting. Figure 5 shows the graphs of the discount rates under the exponential distribution and under finitely-lived agents, for $\Gamma = \frac{1}{\theta}$, with $\theta = 0.02 = r$, for $\lambda = 0.01$ (the negatively sloped curves) and for $\lambda = 0.06$ (the positively sloped curves). For $\lambda < r$, the two assumptions about lifetime lead to very similar discount profiles. For $\lambda > r$, the two profiles are similar for the first 15 - 20 years. However, for large λ, the future ceases to matter much after a few decades.[18]

[18] An earlier version of this paper shows numerically that equilibrium policy under the two assumptions about lifetime are similar for $\lambda < r$ and for large λ. To satisfy space constraints, I hereafter consider only the case of exponentially distributed lifetime, emphasizing the situation where $\lambda \leq r$.

Departure from the linear-in-state model An earlier version of this paper considered a model that is non-linear in the state. That model is equivalent to the Quadratic model in Example 2, if one replaces the damage function $e^{-\kappa s}$ by $e^{-\kappa s^2}$, so that consumption is $A_t \exp\left(a_t X - \frac{d_t}{2} X^2 - \kappa s^2\right)$ and utility is $\ln A_t + a_t X - \frac{d_t}{2} X^2 - \kappa s^2$, resulting in the familiar linear-quadratic payoff. The greater complexity of this model requires numerical analysis. The qualitative results obtained above also hold in the numerical results under the stationary linear-quadratic specification.

The tractability of the LIS model is appealing, especially for a research question that seeks general insights. However, that model implies that the climate-related loss is linear in the state when measured in utility units, and is concave when measured in output units. The linear-quadratic model implies that the climate-related loss in utility units is convex in s; the loss in output units is convex at low s and concave for large s. Many environmental economics models assume convex damages. The limit equilibrium to the LIS model implies that strategies are dominant; the limit equilibrium in the linear-quadratic model does not involve dominant strategies. There are strategic interactions in the linear-quadratic model that are absent from the LIS model (for the limit equilibrium). In short, both the LIS and the linear-quadratic models have distinct advantages.