

Karp, Larry S.

**Working Paper**

## Provision of a public good with altruistic overlapping generations and many tribes

CESifo Working Paper, No. 3895

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Karp, Larry S. (2012) : Provision of a public good with altruistic overlapping generations and many tribes, CESifo Working Paper, No. 3895, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/61034>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Provision of a Public Good with Altruistic Overlapping Generations and Many Tribes

Larry Karp

CESIFO WORKING PAPER NO. 3895  
CATEGORY 10: ENERGY AND CLIMATE ECONOMICS  
JULY 2012

*An electronic version of the paper may be downloaded*

- *from the SSRN website:* [www.SSRN.com](http://www.SSRN.com)
- *from the RePEc website:* [www.RePEc.org](http://www.RePEc.org)
- *from the CESifo website:* [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

# Provision of a Public Good with Altruistic Overlapping Generations and Many Tribes

## Abstract

Intergenerational altruism and contemporaneous cooperation are both important to the provision of long-lived public goods. Equilibrium climate protection may depend more sensitively on either of these considerations, depending on the type of policy rule one examines. This conclusion is based on a model with  $n$  tribes, each with a sequence of overlapping generations. Tribal members discount their and their descendants' utility at different rates. Agents in the resulting game are indexed by tribal affiliation and the time at which they act. The Markov Perfect equilibrium is found by solving a control problem with a constant discount rate and an endogenous annuity.

JEL-Code: C730, D620, D630, D640, H410, Q540.

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

*Larry Karp*  
*Department of Agricultural and Resource Economics*  
*University of California, Berkeley*  
*California / Berkeley / USA*  
*karp@berkeley.edu*

July 19, 2012

Earlier versions of this paper benefitted from comments by Geir Asheim, Ivar Ekeland, Antii Iho, Jacek Krawczyk, Derek Lemoine, Armon Rezai, Christian Traeger, and seminar participants at Berkeley, Santa Barbara, Paris, Vienna, and SURED 2012. The usual disclaimer holds.

# 1 Introduction

The provision of a long-lived public good depends on the ability of agents who make current decisions to cooperate with each other, and on the extent to which they care about the welfare of the not-yet born. The determination of climate policy illustrates these two considerations, but similar issues arise in many other settings. I develop a model that helps to assess the relative importance, for the provision of a public good, of contemporaneous cooperation and intergenerational altruism. The protection of the climate, like the provision of other long-lived goods, is generically a coordination problem, with many equilibria. The relative importance of contemporaneous cooperation and intergenerational altruism may depend on which equilibrium one examines. I study a model that imbeds overlapping generations in a differential game, and then specialize to the climate setting. The overlapping generations component is critical for disentangling impatience with regard to one's own future consumption, from altruism with regard to unborn generations. The differential game component is critical for representing different degrees of contemporaneous cooperation.

Two decades of negotiation confirm the importance and the difficulty of getting policymakers to cooperate on climate policy. A recent literature emphasizes the role of discounting in selecting climate policy (Stern, 2006; Nordhaus, 2007, Weitzman, 2007). The UK and France currently use low social discount rates to evaluate long-lived public projects, and the US Environmental Protection Agency (2010) is considering a similar proposal. These reforms promote intergenerational equity. Even if such reforms were widely adopted, to what extent would they affect the equilibrium protection of the global commons, at given levels of international cooperation? How would increased international cooperation affect climate policy, for a given level of intergenerational altruism?

The climate problem is intergenerational. If agents currently alive discount their own future utility and the utility of the not-yet born at constant but different rates, then the aggregation of their preferences implies non-constant discounting, and time-inconsistent preferences. The existence of different tribes (countries, in the climate policy context) complicates the policy choice. I consider the case where there are  $n \geq 1$  tribes, each of which has an equal share of the world's fixed population. The current population of each tribe consists of many generations. Members of a tribe care about their own utility stream and – to the extent that they have intergenerational

altruism – about the utility streams of future tribal members. Within each tribe and at each point in time, a social planner aggregates the preferences of tribal members currently alive.

When there is a single tribe in which agents discount their own and the unborn generations' utility at different rates, the result is the familiar problem of non-constant discounting (Strotz, 1956; Laibson, 1997). When there are multiple tribes, each with a constant discount rate, the result is a differential game (Long, 2010; Haurie, Krawczyk and Zaccour, 2012). With  $n$  tribes and non-constant discounting, the result is a game in which each social planner is distinguished by the tribal index and the time that she acts. The different tribes' social planners alive in the same period act simultaneously, and the social planner of any tribe acts before the future social planners. I consider a symmetric Markov Perfect Equilibrium (MPE) to a symmetric game: the optimal action of the social planner in a particular tribe at a particular point in time is a function of only the payoff-relevant state variable.

Most overlapping generations (OLG) models assume either that the individual lifetime is a random variable (Yaari, 1965; Blanchard, 1985; Calvo and Obstfeld, 1988) or a finite known parameter, as in Diamond (1965)'s two period model or in Schneider, Traeger, and Winkler (2010)'s continuous time setting. I show that the discount rates in the two models, where agents' lifetimes are either exponentially distributed or a known constant, are always similar in the short run; the two rates might differ greatly at times close to or beyond the finite known lifetime. If current actions are insensitive to changes in distant discount rates, and if the deterministic lifetime is long, the two models are likely to be observationally similar. Thus, if the agents' lifetime corresponds to their biological span, the two models are likely to yield similar equilibria. If instead, the agents' lifetime corresponds to the time they are in political office, the two models may produce quite different equilibria.

A “paternalistically altruistic” agent cares about the utility flows enjoyed by future generations, but does not take into account that each of those agents also cares about the utility flows of subsequent agents. The “purely altruistic” agent does take into account the fact that her successors also value their own successors' utility flows (Ray, 1987; Andreoni, 1989; Saez-Marti and Weibull, 2005). If lifetime is exponentially distributed, I find that the models of paternalistically and purely altruistic agents are observationally equivalent. The two types of agents have the same preferences if the discount rates that they use to evaluate their successors' payoffs differ by the mortality rate.

After presenting the non-constant discount rates that arise in this OLG setting, I describe the game amongst the tribes and across the generations, present the equilibrium conditions, and discuss the multiplicity of MPE. When agents are “very altruistic”, these MPE can be Pareto ranked, and it is possible to support a “good” outcome.

Because the focus of this paper is to assess the relative importance of altruism and contemporaneous cooperation, in determining the equilibrium climate policy, I need to solve, rather than merely describe the equilibrium. For this reason, I use a linear-quadratic model, which is simple and involves only a few parameters. This parsimony and transparency is particularly valuable in light of the uncertainty about the true costs and benefits of reducing carbon emissions. The linear-quadratic model has been widely used to study differential games (with a constant discount rate) in both industrial organization (Fershtman and Kamien, 1987; Reynolds, 1987) and natural resource economics (van der Ploeg, 1992; Dockner and Long, 1993; Wirl, 1994); it has also been used to study quasi-hyperbolic discounting in the one-agent setting (Karp 2005).

I find that the relative importance of intergenerational altruism and contemporaneous cooperation is sensitive to the type of equilibria that one considers. With a linear equilibrium, the degree of altruism is relatively unimportant at relevant levels of contemporaneous cooperation. In contrast, with non-linear equilibria, the degree of altruism is especially important at relevant levels of contemporaneous cooperation.

## 2 Discounting

The objective of this paper is to evaluate the relative influence, on investment in a public good, of agents’ attitudes toward future generations and on the ability of different groups to cooperate at a point in time. I emphasize the case where lifetime is exponentially distributed and agents’ altruism is paternalistic. To determine whether results are sensitive to model details, I consider two alternatives. The first replaces the assumption of paternalistic with pure altruism, under exponentially distributed lifetime. The second alternative replaces the assumption of exponentially distributed lifetime with a known finite lifetime, under paternalistic altruism. In order to make this paper self-contained, but without distracting from its main message, I collect derivations of the discount functions in an appendix. Ekeland and Lazrak

(2010) provide the formula for the case of exponentially distributed lifetime and paternalistic altruism, but the two alternatives are new.<sup>1</sup>

In every case, the population is constant, so the birth rate equals the death rate. The task of this section is to construct and analyze the discount factor used to aggregate utility streams. For an arbitrary sequence of bounded utility flows,  $\{u_\tau\}_{\tau=t}^\infty$ , I find the discount factor  $D(\tau)$  for which welfare of a social planner at  $t$  is  $\int_t^\infty D(\tau) u_\tau d\tau$ . The discount factor  $D(\tau)$  depends on the both the type of altruism and on whether lifetime is exponentially distributed or deterministic. There is no tribal index here, because I consider a representative tribe.

## 2.1 Exponential lifetime, paternalistic altruism

Consider a public project, e.g. protection of the climate system. An agent’s utility flow at a point in time depends on the current stock of the public good (e.g., the stock of greenhouse gasses) and on her tribe’s current investment in that good (e.g., abatement). This investment cost is shared equally by all tribal members then alive, so *at time  $t$  they all have the same utility flow,  $u_t$* . Agents’ welfare consists of a selfish and an altruistic component. The selfish component equals the present discounted value of the expected flow of the agent’s utility, using a constant pure rate of time preference,  $r$ . The altruistic component consists of the agent’s evaluation of her successors’ stream of utility, which she discounts at rate  $\lambda$ . In this paper, “altruism” refers only to benevolence toward one’s descendants, not toward current or future members of other tribes. At the cost of introducing another parameter, one could distinguish between intergenerational altruism within and across tribes. The agent’s mortality rate is  $\theta$ , so  $\gamma \equiv r + \theta$  is the agent’s risk-adjusted pure rate of time preference.

The memoryless feature of the exponential distribution means that all agents alive at a point in time have the same distribution function for their

---

<sup>1</sup>Several papers use the sum of exponentials to represent non-constant discounting (Li and Lofgren 2000), (Gollier and Weitzman 2010), (Zuber 2010), and (Jackson and Yariv 2011). The convex combination of exponentials may arise because different agents want to use different discount factors to evaluate future flows, and the decisionmaker takes a weighted sum of their preferences; or it may arise because the decisionmaker is uncertain about the correct discount rate, and therefore takes the expectation of the associated discount factors. Ekeland and Lazrak (2010)’s motivation is quite different: the OLG structure gives rise to the sum (but not necessarily convex combination) of exponentials.

remaining lifetime. Because there is no private accumulation in this model, all agents alive at a point in time are identical. In this case, there is a representative agent in the usual sense. The paternalistic agent cares about her successors' utility stream, but does not take into account that each successor also cares about their own successors' utility stream. The purely altruistic agent (studied below) does take into account the fact that her successors care about their own successors' utility streams.

The discount factor for the paternalistic agent with exponentially distributed lifetime is

$$D(t) = \left( \frac{\lambda - r}{\lambda - \gamma} \right) e^{-\gamma t} - \frac{\theta}{\lambda - \gamma} e^{-\lambda t}. \quad (1)$$

## 2.2 Robustness

Both because of its intrinsic interest, and also to gauge the sensitivity of the results to assumptions about agents' lifetime and their type of altruism, I consider two alternatives to the model with paternalistic altruism and exponentially distributed lifetime.

### 2.2.1 Pure altruism, exponentially distributed lifetime

Deriving the discount factor for the representative agent under pure altruism is more complicated than under paternalistic altruism. In the latter case, one can simply write down the discount factor from its definition, and then simplify by changing the order of an integration to obtain equation (1). With pure altruism, in contrast, it is necessary to solve a recursion. I achieve this in two stages. First, I begin with a discrete time model, in which each period lasts for  $\varepsilon$  units of time. I solve the resulting discrete time recursion, to obtain the discrete time discount function for the representative agent. Taking the limit as  $\varepsilon \rightarrow 0$  gives the continuous time discount function under pure altruism. Comparing that function with the expression for  $D(t)$  in equation (1) establishes an isomorphism between paternalistic and pure altruism:

**Proposition 1** *Suppose that the agent with pure altruism discounts future agents' welfare at rate  $\lambda'$ , and that the agent with paternalistic altruism discounts future agents' utility at rate  $\lambda$ . Both have exponentially distributed lifetimes with mortality rate  $\theta$  and the pure rate of time preference  $r$ . The two agents have the same preferences if and only if  $\lambda' = \lambda + \theta$ .*

The following corollary is a consequence of Proposition 1 and the fact that for  $D(t)$  given by equation (1),  $\frac{dD}{d\lambda} < 0$  for  $t > 0$ :

**Corollary 1** *Given  $\theta$  and the same preference parameters  $(r, \lambda)$ , the agent with paternalistic altruism discounts the future flow of utility more heavily than the agent with pure altruism.*

This comparison is not surprising: the agent with pure altruism cares about future utility flows both because they affect the future generations that directly experience those flows, and because they affect the welfare of earlier generations that care about those future generations. In contrast, the agent with paternalistic altruism cares only about the direct affect of future utility flows on the agents who experience them.

### 2.2.2 Paternalistic altruism, known finite lifetime

Each agent lives for  $T = \frac{1}{\theta}$  years. This equality means that the known finite lifetime in this setting equals the expected lifetime in the exponentially distributed setting. Thus, the two models are directly comparable.

With deterministic lifetimes, agents alive at a point in time are different: the older ones will die sooner than the younger ones. In this setting, I assume that the representative agent (social planner) at a point in time is utilitarian; she puts equal weight on the preferences of all tribal members currently alive. Bergstrom (2006) points out that when agents feel benevolence toward others who share both the costs and the benefits of a public good, it is necessary to count both the “sympathetic costs” as well as the “sympathetic benefits” (those arising from the feeling of benevolence). If agents are identical, as in the exponentially distributed case, the sympathetic costs offset the sympathetic benefits, so benevolent feelings toward other tribal members currently living would not affect the cost benefit calculation. In the case of deterministic lifetime, where agents are not identical, I ignore any benevolent feelings that an agent has for others currently alive. Those agents can speak for themselves, and the social planner gives each of their preferences equal weight. The unborn (future) tribal members enjoy the benefits of the public good but do not share the current investment costs, and they cannot directly influence the tribe’s current social planner. Therefore, benevolent feelings toward these agents affect the cost benefit calculation.

With a finite lifetime,  $T$ , the structure of the discount function differs for  $t < T$  and  $t > T$ . For  $t < T$ , some of the agents currently alive (at

time 0) will still be alive. Those agents continue to benefit from the utility flow, and that utility flow contributes to the selfish component of the welfare of the current (time 0) representative agent. For  $t > T$ , all of the agents alive at time 0 will have died, so the representative agent at time 0 places a positive weight on the utility flow at  $t > T$  only to the extent that agents are altruistic ( $\lambda < \infty$ ).

For finitely lived agents, the discount factor is<sup>2</sup>

$$D(t) = \begin{cases} e^{-rt} \left( \frac{1-e^{-(\lambda-r)t}}{T(\lambda-r)} + \frac{T-t}{T} \right) & \text{for } t \leq T \\ e^{-\lambda t} (e^{(\lambda-r)T} - 1) \frac{1}{T(\lambda-r)} & \text{for } t \geq T \end{cases} \quad (2)$$

### 2.3 Discount rates: a comparison

Here I compare the discount rates under the baseline model (exponentially distributed lifetime, paternalistic altruism) and under the finite lifetime alternative. In view of the isomorphism between the types of altruism (under the exponential distribution), a separate treatment of the agent with pure altruism is unnecessary.

The discount rate,  $\eta$ , associated with the discount factor in equation (1) is

$$-\frac{dD}{dt} \frac{1}{D} \equiv \eta(t) = \frac{-\gamma\lambda + \gamma r + \theta\lambda e^{-(\lambda-\gamma)t}}{-\lambda + r + \theta e^{-(\lambda-\gamma)t}}. \quad (3)$$

The definition of the discount rate  $\eta(t)$  implies

$$\text{for } r < \lambda < \infty: \quad \frac{d\eta(t)}{dt} > 0; \quad \eta(0) = r; \quad \lim_{t \rightarrow \infty} \eta(t) = \gamma; \quad (4)$$

$$\eta|_{\lambda=\infty}(t) = \gamma \quad \eta|_{\lambda=r}(t) = r \quad \eta|_{\lambda=0}(t) = \gamma \frac{r}{r + \theta e^{\gamma t}}; \quad (5)$$

$$\text{for } 0 < \lambda < r: \quad \frac{d\eta(t)}{dt} < 0; \quad \eta(0) = r; \quad \lim_{t \rightarrow \infty} \eta(t) = \lambda. \quad (6)$$

Equation (4) states that if the current generation cares less about future generations' welfare than about its own, but has some concern for the future ( $r < \lambda < \infty$ ), the discount rate increases over time from the pure rate of time preference,  $r$ , to the risk adjusted rate,  $\gamma = r + \theta$ . These limits are

---

<sup>2</sup>This model can be viewed as a generalization of the  $\beta, \delta$  of quasi-hyperbolic discounting (Laibson 1997). That relation is easiest to see in a discrete time setting; see Appendix A.1.3

independent of  $\lambda$ . That is, for finite  $t$ , a finite value of  $\lambda > r$  shifts down the entire trajectory of the discount rate, relative to the risk adjusted selfish rate. For  $\lambda < r$ , the discount rate falls, as with hyperbolic discounting (equation 6), and in this case the discount factor is a convex combination of two exponentials.<sup>3</sup>

Equation (5) provides the discount rate for three limiting values of  $\lambda$ :  $\infty$  (using the compactified real line),  $r$ , and 0. The fact that  $\eta|_{\lambda=\infty}(t) = \gamma \forall t \geq 0$  whereas  $\eta(0) = r$  for  $\lambda < \infty$  means that there is a discontinuity in  $\eta(t)$  at  $\lambda = \infty$ .<sup>4</sup> The discount rate is constant over time for  $\lambda = \infty$  and for the borderline case where the agent cares as much about future generations' utility as about her own ( $\lambda = r$ ).

The discount rate under a finite lifetime is:

$$\eta(t) = \begin{cases} \frac{\lambda e^{-(\lambda-r)t} - r(\lambda-r)(T-t) - \lambda}{e^{-(\lambda-r)t} - 1 - (\lambda-r)(T-t)} & \text{for } t \leq T \\ \lambda & \text{for } t \geq T. \end{cases} \quad (7)$$

The shape of the trajectory of the discount rate, for  $t \in (0, T)$ , depends on whether  $\lambda$  is greater or less than  $r$ . For  $\lambda \in (r, \infty)$ , the discount rate increases from  $r$  to  $\lambda$  as  $t$  increases. For  $\lambda \in [0, r)$  the discount rate falls from  $r$  to  $\lambda$ , and for  $\lambda = r$  the discount rate is constant. For  $\lambda = 0$ , the discount rate is  $r^2 \frac{T-\tau}{e^{r\tau} + (T-\tau)r - 1}$  for  $\tau \leq T$  and 0 for  $\tau > T$ . In the other limiting case, where  $\lambda = \infty$ , the discount rate is  $\eta(t) = r + \frac{1}{T-t}$  for  $0 \leq t < T$  and infinite for  $T \geq 0$ . As with exponentially distributed lifetimes, there is a discontinuity at  $\lambda = \infty$ .

The following corollary summarizes the comparison of the discount rates under exponentially distributed and deterministic lifetimes for paternalistic agents.

---

<sup>3</sup>Comparison of the last parts of equations (4) and (6) show that the asymptotic (as  $t \rightarrow \infty$ ) discount rate depends on whether  $\lambda \leq r$ . For  $\lambda > r$ , concern for far distant generations dies out more quickly than the concern for one's own future utility flow, so the asymptotic discount rate equals the risk adjusted rate. For  $\lambda < r$ , self-interest dies out more quickly than altruism, so the asymptotic discount rate equals  $\lambda$ .

<sup>4</sup>A calculation shows that for  $\lambda > r$  the  $L1$  norm  $\mu$  is continuous in  $\lambda$ :

$$\mu(\lambda) \equiv \int_0^\infty (\gamma - \eta(\tau)) d\tau = \ln \left( \frac{\lambda - (r + \theta)}{\lambda - r} \right) \implies \lim_{\lambda \rightarrow \infty} \mu(\lambda) = 0 \quad \text{and} \quad \mu(\infty) = 0.$$

**Corollary 2** (i) *The discount rates under paternalistic altruism with exponentially distributed or deterministic lifetimes are constant and equal to each other iff: (a)  $\lambda = r$ , where  $\eta = r$ ; **or** (b)  $\theta = \infty$  ( $T = 0$ ), where  $\eta = \lambda$ ; **or** (c)  $\theta = 0$  ( $T = \infty$ ), where  $\eta = r$ . (ii) For  $\lambda = \infty$  the discount rate under exponentially distributed lifetime is constant at  $\gamma = r + \theta$ , and the discount rate under a deterministic lifetime begins at  $r + \frac{1}{T} = r + \theta$  and approaches  $\infty$  as  $t \rightarrow T$ . (iii) For finite  $\lambda$  and for positive  $t$  close to 0 and for  $t \geq T$ , the discount rate for agents with exponentially distributed lifetime is greater than the discount rate for agents with deterministic lifetime if and only if  $\lambda > r$ . (iv) For  $\lambda > r$  the difference between the discount rates approaches  $\lambda - r - \theta$  as  $t \rightarrow \infty$  and for  $\lambda < r$  the difference approaches 0.*

**Proof.** Parts (i.a), (ii), (iv) and the claim in part (iii) regarding  $t \geq T$  merely summarize the previous discussion. Parts (i.b, i.c) are by inspection. Verifying the claim in part (iii) regarding positive  $t$  close to 0 uses a third order expansion of the discount factors evaluated at  $t = 0$ : the discount factor with exponentially distributed lifetimes minus the discount factor with deterministic lifetimes equals  $\frac{1}{6}\theta^2(\lambda - r)t^3 + o(t^4)$ . ■

This corollary suggests that in the cases where  $\lambda \approx r$ , or  $\lambda < r$  or  $\lambda \approx r + \theta$  the trajectories of the discount rates are “quite similar”, so that the preferences of the two types of agents (with deterministic or exponentially distributed lifetimes) are also quite similar. However, if  $\lambda \gg r + \theta$  the discount rates, and thus preferences, are quite dissimilar in the two cases for  $t \geq T$ ; for small  $t$  the discount rates are similar. If  $T$  is large, then the fact that the discount rates are dissimilar for  $t \geq T$  is likely to be rather unimportant for current decisions. In summary, unless  $\lambda$  is large and  $T$  is small, preferences under the two models are “quite similar”.

Parts (i.b, i.c) show the importance of the overlapping generations structure. If  $\theta = \infty$  ( $T = 0$ ) a tribe consists of a succession of agents, each of whom lives for a single instant, and has a constant discount rate  $\lambda$ . At the other extreme,  $\theta = 0$  ( $T = \infty$ ) a tribe consists of an infinitely lived agent, with a constant discount rate  $r$ . For these two limiting cases, there is no time consistency problem. For small  $\lambda$  and  $0 < \theta < \infty$  agents would be willing to incur costs to transfer utility to the distant future, but their incentive to do so is limited by the fact that they know that their future selves would consume some of the intended transfer.

Figure 1 shows the discount rates in the two models, for  $r = 0.02 = \theta$  (so  $T = 50$ ). The solid curves (labelled  $E$ ) correspond to exponentially

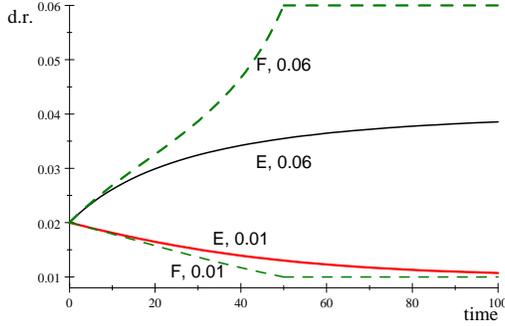


Figure 1: Discount rates (d.r.) for  $\theta=0.02=r=\frac{1}{T}$ . Solid curves (labelled E) correspond to exponentially distributed lifetime and dashed curves (labelled F) correspond to fixed lifetime. Numerical values in label show value of  $\lambda$ .

distributed lifetime, and the dashed curves (labelled  $F$ ) correspond to fixed lifetime. The increasing curves correspond to  $\lambda = 0.06$  and the decreasing curves correspond to  $\lambda = 0.01$ . This figure illustrates Corollary 2 and the comments that follow it.

The qualitative differences illustrated in Figure 1 follow naturally from the underlying assumptions of the two models. With exponentially distributed lifetime, the probability that an agent is alive at a point in time in the future decreases at a constant rate. Therefore, even if she cares nothing about the unborn generations, her discount rate for future utility flows never rises above  $r + \theta$ . In contrast, in the deterministic lifetime case, all agents currently alive will be dead within  $T$  years, so the utilitarian agent who aggregates the preferences of tribal members currently alive puts weight on utility flows after  $T$  years only to the extent that agents currently alive care about future generations. Completely selfish agents put no weight on the welfare of the unborn ( $\lambda = \infty$ ). A smaller value of  $\lambda$  implies a higher level of altruism.

The utility flow  $t$  periods in the future may be shared by some agents currently alive and also by agents born  $t$  periods in the future. The agent currently alive puts the same weight on the utility flow received by her future self and the not-yet born if and only if  $\lambda = r$ . However, for any  $\lambda > 0$ , the agent alive today puts lower value on the lifetime welfare of an agent born in the future, the further in the future that agent is born. A plausible requirement of ethical behavior is to treat the present discounted value of

the stream of utility of all agents symmetrically, regardless of when they are born (Ramsey 1928). That requirement implies  $\lambda = 0$  (*not*  $\lambda = r$ ).

### 3 The game

Both the game and the equilibrium are easiest to describe in a discrete time setting, where the equilibrium conditions can be obtained using elementary methods. The discrete time analog provides at least as good a representation of the world as does the continuous time version. For these reasons, I study the model by taking formal limits of a discrete time model.

There are  $n$  symmetric tribes, each of which consists of a sequence of overlapping generations of the type considered above. A larger value of  $n$  corresponds to greater fragmentation, or less cooperation, amongst agents at a point in time. At time  $t$  a social planner for tribe  $i$  chooses the action  $x_{it}$ . This action affects  $i$ 's current utility flow and the evolution of a state variable,  $S_t$ , common to all tribes. For example,  $S_t$  is a climate-related variable, such as the stock of greenhouse gasses (GHG) or average temperature, and  $x_{it}$  is tribe  $i$ 's GHG emissions or abatement at time  $t$ . In the interest of simplicity, I assume that tribe  $i$ 's flow of utility at  $t$ ,  $u(S_t, x_{it}; n)$ , depends only on the state variable and  $i$ 's action. (Including  $j$ 's current action in  $i$ 's utility flow makes it possible to include leakage in a climate-related model or to consider the case of cross-tribal altruism.) In this symmetric environment, all tribes have the same utility function and discount factor. The state variable evolves according to

$$S_{t+\varepsilon} - S_t = f(S_t, \mathbf{x}_t; n) \varepsilon, \text{ with } \mathbf{x}_t = (x_{1t}, x_{2t} \dots x_{nt}).$$

The parameter  $n$  in the growth function,  $f$ , and the utility function,  $u$ , makes it possible to consider a fragmentation of the economy (larger  $n$ ) that leaves unchanged the set of feasible utility. A change in  $n$  has no intrinsic effect on aggregate (or per capita) utility and stock flows, but it does alter the equilibrium decisions, thereby altering the equilibrium aggregate utility and stock flows:  $n$  has a strategic but not an intrinsic effect on outcomes.

The payoff to the social planner in tribe  $i$  at time  $t$  is the present discounted value of the stream of utility:

$$\sum_{\tau=0}^{\infty} D(\tau; \varepsilon) u(S_{t+\tau}, x_{i,t+\tau}; n) \varepsilon, \tag{8}$$

where the discount factor depends on whether lifetime is exponentially distributed or deterministic. In this discrete time setting, agent  $it$  is the social planner in tribe  $i$  at period  $t$ . The strategic interactions in this game occur amongst different tribes and across periods. Agents  $it$  and  $jt$  move simultaneously (choosing  $x_{it}$  and  $x_{jt}$ , respectively); for all  $\tau > t$ , agent  $it$  moves before agent  $j\tau$ .

I consider a symmetric stationary pure strategy Markov Perfect Equilibrium (hereafter, “MPE”), a function  $\chi^\varepsilon(S)$  that satisfies the Nash condition: If the stock at  $t$  is  $S_t$  and agent  $it$  believes that agents  $jt$ ,  $\forall j \neq i$ , use the decision rule  $x_{jt} = \chi^\varepsilon(S_t)$ , and that agents  $j\tau$  for  $\forall j$  and  $\tau > t$  will use the decision rule  $x_{j\tau} = \chi^\varepsilon(S_\tau)$ , then  $x_{it} = \chi^\varepsilon(S_t)$  is the optimal action for agent  $it$ . The superscript  $\varepsilon$  reminds the reader that a function corresponds to a particular length of a period in the discrete time setting,  $\varepsilon$ .

The symmetry of this problem makes it possible to consider a single tribe’s problem, taking as given the behavior of all other tribes. Let the current period index be  $t$ , and suppose that  $\forall \tau \geq t$  and  $j \neq i$  agents  $i\tau$  believe that agents  $j\tau$  will use the policy  $x_{j\tau} = \chi^\varepsilon(S_\tau)$ . That is, each of the succession of social planners in tribe  $i$  expects each of the succession of social planners in other tribes to use this policy rule. With these beliefs, agent  $it$  expects the state to evolve according to

$$S_{\tau+\varepsilon} - S_\tau = F^\varepsilon(S_\tau, x_{i\tau}) \varepsilon \text{ with } F(S_\tau, x_{i\tau}) \equiv f(S_\tau, \iota_{n-1} \chi^\varepsilon(S_\tau), x_{i\tau}; n), \quad (9)$$

$\forall \tau \geq t$ , where  $\iota_{n-1}$  is an  $n-1$  dimensional vector consisting of 1’s. If agents  $i\tau$   $\forall i$  and  $\tau \geq t$  use the policy  $\chi^\varepsilon(S_\tau)$ , then the value of the program for the agent who faces the current state  $S$  is

$$J^\varepsilon(S) = \sum_{\tau=0}^{\infty} D(\tau; \varepsilon) u(S_{t+\tau}^*, \chi^\varepsilon(S_{t+\tau}^*); n) \varepsilon \quad (10)$$

where  $S_\tau^*$  is the solution to equation (9) when all agents use the policy rule  $\chi^\varepsilon$  and the initial condition is  $S$ .

The payoff in (8) with the equation of motion (9) are the ingredients of the game amongst the sequence of social planners in tribe  $i$ . These planners play a game rather than solving a standard control problem, because the function  $D(\tau; \varepsilon)$  is not exponential: as a consequence, the sequence of actions that is optimal for an agent at a point in time is not optimal for her successors. The social planner at a point in time can choose the action at that time, but cannot commit her successors to particular actions. She therefore chooses

the action that maximizes the discounted flow of current and future utility, understanding that future actions are chosen according to  $\chi^\varepsilon$ .

The continuous time payoff and equation of motion for tribe  $i$  are

$$\int_\tau^\infty D(t - \tau)u(S_t, x_{it}; n)dt \text{ and } \frac{dS_t}{dt} \equiv \dot{S}_t = F(S, x_i) \quad (11)$$

with  $F(S, x) \equiv f(S, \iota_{n-1}\chi(S), x_i; n)$ ,

where  $\chi(S)$  is a MPE policy function in the continuous time setting, and the discount factor,  $D$ , is given by equation (1) in the case of agents with exponentially distributed lifetime, and by equation (2) in the case of agents with deterministic lifetime.

A larger value of  $n$  means that the time  $t$  social planner in a particular tribe internalizes a smaller fraction of the effect of her actions. A larger value of  $\lambda$  means that this social planner internalizes a smaller fraction of the future effect of her actions. The point of this paper is to determine the relative sensitivity of equilibrium investments in a public good, to changes in  $n$  and  $\lambda$ .

### 3.1 Equilibrium conditions

Karp (2007), building on Harris and Laibson (2001), finds the formal limit, as  $\varepsilon \rightarrow 0$ , of the equilibrium conditions to the sequential game defined by the payoff in (10) with the equation of motion (9) for the one-tribe ( $n = 1$ ) case. These conditions require that  $J(S) \equiv \lim_{\varepsilon \rightarrow 0} J^\varepsilon(S)$  and its first derivative exist – an assumption that can be checked given a particular equilibrium. These conditions can be applied directly to the case of  $n > 1$ : when the succession of planners in tribe  $i$  take as given other tribes' policy rule, tribe  $i$  has a standard problem with non-constant discounting. The resulting simplicity relies on the assumption of a symmetric (across both tribe and time indices) equilibrium.

The equilibrium conditions differ in the two cases corresponding to  $\lambda < r$  and  $\lambda > r$  with exponentially distributed lifetime (because  $\lim_{t \rightarrow \infty} \eta(t)$  differs in the two cases), and in the case where agents have deterministic lifetimes (because the function  $K$ , introduced below, has a different form here). For  $\lambda = r$  with both deterministic and exponentially distributed lifetimes, and for  $\lambda = \infty$  with exponentially distributed lifetime, the discount rate is constant. In these cases, the tribes play a standard differential game, i.e. one without the strategic interactions within a tribe, across periods. I provide details

for the model with exponentially distributed lifetime and  $0 < \lambda \leq r$  (where  $\lim_{t \rightarrow \infty} \eta(t) = \lambda$ ), relegating the other two cases to Appendix B.1.

Dropping the agent index  $i$  (because of symmetry) and the superscript  $\varepsilon$  (because I now consider the continuous time limit) Proposition 1 and Remark 1 of Karp (2007) imply that  $\chi(S)$  satisfies the necessary condition to the following “fictitious” optimal control problem with constant discount rate  $\lambda$

$$J(S) = \max \int_0^\infty e^{-\lambda t} (u(S_t, x_t; n) - K(S_t)) d\tau \quad \text{subject to } \dot{S} = F(S, x), \quad (12)$$

with the side condition (definition):

$$K(S_t) = (r - \lambda) \int_0^\infty e^{-\gamma t} u(S_\tau, \chi(S_\tau); n) d\tau. \quad (13)$$

The tribe’s utility flow on the equilibrium path is  $u(S_\tau, \chi(S_\tau); n)$ , and  $S_\tau^*$  is the solution to the differential equation in (12) when all agents use the decision rule  $\chi(S)$ . The function  $K$  can be interpreted as an annuity, which if received in perpetuity and discounted at the rate  $r - \lambda$ , equals the present value of the stream of future utility, discounted at the rate  $\gamma$ .

This model includes familiar special cases. For  $n > 1$ , the endogenously determined function  $F(S, x) = f(S, \iota_{n-1}\chi(S), x; n)$  depends on the policies of the other  $n - 1$  agents; those agents do not exist if  $n = 1$ , in which case  $F(S, x) = f(S, x; 1)$ , an exogenously given function. For  $\lambda = r$ ,  $K \equiv 0$  and the model collapses to a standard (constant discounting) differential game for  $n > 1$  or a control problem for  $n = 1$ .

## 3.2 Nonuniqueness

In general, the equilibrium to this game is not unique. Tsutsui and Mino (1990) note the existence of a continuum of stable steady states (an open interval) in the differential game with constant discounting when decision rules are differentiable. For each point in this interval there is an equilibrium policy function, defined at least in the neighborhood of that point. The economic explanation for this multiplicity in the differential game 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 Markov perfect equilibrium conditions do not pin down these beliefs. In a standard optimal control

problem, the envelope theorem eliminates that kind of consideration, because the first order welfare effect of a deviation from the steady state is 0. This theorem is not applicable in the differential game, because rivals' actions do not maximize an agent's welfare. In this sense, the transversality condition typically applied to study autonomous control problem is "incomplete".

The same kind of consideration applies in the one-tribe model with non-constant discounting. Karp (2007) and Ekeland and Lazrak (2010) characterize the set of stable steady states in this setting. When  $n > 1$  and the discount rate is non-constant, there are two sources of multiplicity of steady states, so we would expect the equilibrium to be unique (within the class that induce differentiable value functions) only under special circumstances. The multiplicity of equilibria means that there is a coordination problem both across tribes and across generations. Some MPE may Pareto dominate others.

Most applications use specific functional forms, because of the difficulty of obtaining general results outside the steady state. Models with logarithmic utility and Cobb Douglas growth functions have been used to study both differential games (Levhari and Mirman 1980) and the non-constant discounting problem with  $n = 1$  (Barro 1999). The Introduction notes the widespread use of the linear-quadratic model; some papers consider only the linear equilibrium, and several discuss the multiplicity of equilibria. (Krusell and Smith (2003) and Vieille and Weibull (2009) discuss multiplicity in different settings.)

Ekeland, Karp, and Sumaila (2012) study the  $n = 1$  case using a model that is linear in the control variable, but otherwise allows general functional forms. That paper uses the exponentially distributed lifetime model, but all of its results carry over to the model with deterministic lifetime. Within the class of equilibria that gives rise to a differentiable value function, the equilibrium is unique: the stock follows a most rapid approach path to a steady state, which for  $\lambda < \infty$  is *independent of*  $\lambda$ .

This linear-in-control model also has a simple equilibrium for  $n > 1$ ; in this equilibrium, a bang-bang policy drives the state variable (e.g. the stock of fish) to a level at which the flow of utility is 0. If all other agents use a bang-bang control to drive the state to particular steady state  $S_\infty$  at which the flow of utility is strictly positive, then agent  $i$  has an incentive to undercut them (e.g. to harvest more than the level that sustains that stock). Thus, the only steady state is a level associated with a zero utility flow. In this example, the unique equilibrium is completely insensitive to (finite)  $\lambda$  and

extremely sensitive to a change from  $n = 1$  to  $n = 2$ .<sup>5</sup>

### 3.3 Interpretation of this game

The purpose of this model is to provide intuition about how limited cooperation amongst contemporaneous agents, and different levels and types of intergenerational altruism, interact to affect equilibrium decisions. Game theoretic models like this one are not normative. However, a limiting case of this model is “close” to being normative. Here I consider the agent with exponentially distributed lifetime and paternalistic altruism.

For  $n = 1$ , there is no conflict amongst contemporaneous agents and for  $\lambda = r$  or  $\lambda = \infty$  the time inconsistency problem also vanishes. These parameter values produce a standard optimization problem, rather than a game, but neither provides a reasonable normative model. The ethical objection to  $\lambda = \infty$  is obvious, and the ethical objection to  $\lambda = r$  is that the evaluation of an agent’s utility stream should not depend on her date of birth.

The choice  $\lambda = 0$  has a stronger ethical basis, and therefore (together with  $n = 1$ ) brings the model “closer” to being normative. Nevertheless,  $\lambda = 0$  and  $n = 1$  implies non-constant discounting and therefore produces a game across generations. An equilibrium outcome to that game is in general not normative, for the same reason that the non-cooperative Nash equilibrium to virtually any game is not normative.

For  $\lambda = 0$ , the discount rate converges to 0, so unless the flow payoff converges to 0 sufficiently rapidly, the payoff in the first line of equation (11) is unbounded. Nevertheless, the equilibrium policy function may be well defined even as  $\lambda \rightarrow 0$ , as in the model in Section 4. For small positive  $\lambda$  and bounded  $u_{it}$ , the payoff is well defined and is approximately proportional to the steady state utility flow, divided by  $\lambda$  (see Appendix B.3):

**Lemma 1** *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 expo-*

---

<sup>5</sup>In a MPE to the linear-in-control model, an agent knows that if she were to drive the state away from a locally stable steady state, the equilibrium response of other agents is to drive it back to this steady state *as rapidly as possible*. Consequently, the indeterminacy in beliefs that gives rise to the multiplicity of stable steady states in the general setting, is absent here. These remarks rest on the requirement that the value function is differentiable; with more general policies, many other MPE can exist (Dutta and Sundaram 1993).

nentially distributed (equation (1)) and with arbitrarily small  $\varepsilon > 0$

$$(1 - \varepsilon) \frac{\theta}{\gamma} < \lim_{\lambda \rightarrow 0} \left( \frac{\int_0^\infty D(t)u(t)dt}{\frac{u_\infty}{\lambda}} \right) < (1 + \varepsilon) \frac{\theta}{\gamma}.$$

For small  $\lambda$ , the payoff in the steady state dominates the evaluation of welfare.

Denote the steady state that maximizes the steady state utility flow as the “optimal static steady state” (the solution to  $\max_S u(S, x; 1)$  subject to  $f(S, x; 1) = 0$ ). Suppose that this static optimization problem is concave, so that levels of the state variable closer to the optimal static steady state have higher utility levels. Consider the case where  $\lambda$  is small and where the MPE state trajectory does not approach the optimal static steady state. In this situation, Lemma 1 together with the concavity assumption imply that a deviation from the equilibrium that causes the state to move closer to the optimal static steady state, is a Pareto improvement over the MPE. Each generation prefers this deviation.

The following proposition states that for small  $\lambda$ , there is a MPE that supports a steady state arbitrarily close to the optimal static steady state.

**Proposition 2** *Consider the class of differentiable MPE policy rules. For  $n = 1$  and for arbitrarily small positive  $\varepsilon$ , it is possible to support a MPE steady state that leads to a utility flow within  $\varepsilon$  of the utility level at the optimal static steady state, provided that  $\lambda$  is sufficiently small (but positive).*

This proposition, together with Lemma 1 and the generic multiplicity of equilibria (when policy functions are differentiable), implies that when  $\lambda$  is small but positive, a MPE that takes the state close to the optimal static steady state Pareto dominates a MPE that takes the state further from the optimal static steady state. The proof of the proposition uses the assumption that the policy function is differentiable. Section 3.2 notes that for this class of policy function, the set of MPE stable steady states is an open interval. Proposition 2 states that a boundary of that interval moves close to the optimal static steady state as  $\lambda$  becomes small. We have no information about the measure of the domain of a policy function that drives the state close to this utility-maximizing level. If that measure is small, it is possible to support such a steady state only for initial conditions close to it. In that case, the existence of this steady state may have little practical importance.

## 4 An application to climate policy

The state variable in the climate model,  $S$ , is the atmospheric stock of carbon, in parts per million (ppm), and the control variable,  $A$ , is abatement, defined as the percent reduction in BAU emissions. The equation of motion is linear in  $S$ ,  $A$  and the aggregate flow payoff is quadratic:  $-\frac{1}{2}(bA^2 + h(S - 280)^2)$ ; 280 is the pre-industrial stock of carbon. The flow cost is the sum of a term proportional to the square of abatement and to the square of the increase in the stock relative to preindustrial levels, thus allowing damages to be convex. Appendix C provides details of the calibration of the model, explains how  $n$  enters the model, and explains how to compute both linear and non-linear equilibria.

The model requires four calibration assumptions:

- (i) The steady state absent anthropogenic emissions is 280 ppm.
- (ii) The half-life of the atmospheric stock is 83 years.
- (iii) The year 2100 BAU stock is 700 ppm, given 2010 stock of 380.
- (iv) A numerical value for  $\Omega \equiv \frac{h(560-280)^2}{b50^2}$ .

The 83-year half life is slightly less than the level used in DICE (Nordhaus 2008). The IPCC's 2007 projections of year 2100 stocks range from 535 to 983 ppm. The assumptions imply that under BAU, the stock reaches 870 ppm after 200 years. That increase might be unrealistically high, but any exaggeration may be offset by the fact that this model ignores the possibility of catastrophic damages.

The parameter  $\Omega$  equals the flow cost of doubling the stock of GHG, relative to preindustrial levels, as a ratio of the flow cost of a 50% reduction in BAU emissions. In view of the amount of uncertainty about both abatement costs and climate damages, I calibrate the model using the ratio of these, instead of the coefficients  $h$  and  $b$ .<sup>6</sup> This approach makes it easy to consider a range of beliefs about relative damages and costs, without the need to consider independently changes in  $h$  and  $b$ . Karp and Zhang (2006) estimate an abatement cost parameter that matches (with  $R^2 = 0.97$  in a

---

<sup>6</sup>The tax  $bA$  supports an abatement level  $A$ . Note also that I use BAU as shorthand for “doing nothing to reduce climate damages”. The model produces equilibrium (not optimal) actions, and these could also be construed as “business as usual” – but that is not how I use the term.

psuedo-regression) the cost assumptions in Nordhaus (1994).<sup>7</sup> That estimate implies that a 50% abatement level reduces Gross World Product (GWP) by about 1.1%. A low-to-moderate estimate of damages resulting from doubling stocks, relative to pre-industrial levels, is 1.33% of GWP. These estimates suggest that  $\Omega = \frac{1.33}{1.12} \approx 1.2$  is consistent with (at least some) previous modeling efforts. I report results for  $\Omega = 1$  (low damages) and  $\Omega = 3$  (moderate damages).

As a plausibility check for the range  $\Omega \in [1, 3]$ , I use the fact that given the first three calibration assumptions, the value of  $\Omega$  determines the value of the optimal static steady state. (Recall that this steady state, defined above, equals the steady state stock that maximizes the steady state flow payoff.) For  $\Omega = 1$ , the optimal static steady state is 553 ppm and for  $\Omega = 3$  this level is 402 ppm.

The three parameters of particular interest are: (i)  $\lambda$ , an inverse measure of the extent to which agents currently alive care about unborn generations in their tribe; (ii)  $n$ , a measure of the degree to which tribal fragmentation impedes agents currently alive from cooperating on current policy; and (iii)  $\Omega$ , a measure of the damages from increased carbon stock relative to the costs of abatement.

## 4.1 Discussion of this model

There is a long list of reasons why the linear-quadratic model cannot “accurately” reflect the complex problem of climate change. A one-state linear model can provide only a rough approximation of the hugely complex carbon cycle (Archer and Brovkin 2008). The quadratic utility function is obviously restrictive. The exclusion of private investment ignores an important channel for the current generation to benefit its successors. The stationarity of the model eliminates exogenous technical change that would result, amongst other things, in changing levels of BAU emissions. By way of compensation, the model requires only four parameters. Given the amount of uncertainty about climate change, such a parsimonious model is particularly useful for the kind of fundamental question that this paper addresses.

---

<sup>7</sup>There are more recent, and perhaps more reliable, cost estimates. I use the earlier estimate partly because it has been fitted to the quadratic function. Significantly, I use this estimate only to obtain an idea of a reasonable magnitude of  $\Omega$ . Despite refinements in cost estimates, I do not think that they have changed by orders of magnitude.

The historical record suggests that future generations are likely to be wealthier than the current generation, a possibility excluded by this stationary model. Is that restriction a disadvantage? In the standard model where welfare equals the sum of the expectation of discounted future utility flows, expected growth can significantly increase the discount rate, rendering changes to distant utility flows unimportant in today's welfare calculation. Recognition that future growth is stochastic has only a second order effect on this calculation. However, this conclusion follows from the fact that in the standard model, the curvature of the utility function captures both risk aversion and the elasticity of intertemporal substitution. Using a model that disentangles these distinct characteristics of preferences, Traeger (2012) shows that under reasonable assumptions, stochastic growth might have little or no effect on the discount rate. The hypothesis that future generations are *likely* to be richer than us, therefore does not necessarily lead to a much higher discount rate. The widely held view to the contrary is a habit of thought, born of our often unquestioning adoption of a model that fails to distinguish (in a risky world) between risk aversion and intertemporal substitutability.

It is worth comparing the linear-quadratic model to a prominent alternative. Golosov, Hassler, Krusell, and Tsyvinski (2011) construct an Integrated Assessment Model that involves both produced capital and the environment. Their specification of logarithmic utility and Cobb Douglas production yields the familiar rule that savings is a constant fraction of output. Importantly, this constant is independent of environmental damages, permitting a decoupling of the savings and abatement decisions. They also assume that damages are log-linear in the environmental variable, resulting in an optimal carbon tax that is a constant fraction of output. Gerlagh and Liski (2012) adapt this model to include, amongst other things, quasi-hyperbolic discounting. Restricting their attention to linear equilibria, they again find that the savings and tax are constant fractions of output.

These functional assumptions (logarithmic utility, Cobb Douglas production and log-linear damages) could be imbedded in my model, to obtain decision rules that depend on  $n$ ,  $r$ ,  $\lambda$  and  $\theta$ . This exercise would be worthwhile, because the decoupling of the investment and abatement decisions would make it possible to include distinct capital stocks for the different tribes. The examination of non-linear equilibria would also be interesting.

The assumptions leading to a constant tax (possibly as a share of output) greatly simplifies the analysis. In general, a model with constant marginal

damages has a constant tax policy. A variation of my linear quadratic model replaces  $h(S - 280)^2$  with  $h(S - 280)$ , making marginal damages constant rather than increasing.

**Remark 1** *For the linear-quadratic model with constant rather than linear marginal damages, there are two globally defined MPE decision rules; one is linear in the stock and the other is a constant. There are also non-linear equilibria.*

Define  $\Omega' = \frac{h(560-280)}{b50^2}$ ,  $d$  as the decay rate for the stock and  $C$  as the abatement coefficient in the equation of motion, and assume that  $\lambda < r$  and lifetime is exponentially distributed lifetime. Abatement in the constant equilibrium equals  $A^* \equiv \Omega' \frac{C(d-\theta-\lambda)}{\sqrt{2}(r+\theta-d)n(\lambda-d)} > 0$ , which is decreasing in both  $\lambda$  and  $n$  and linear in  $\Omega'$ .<sup>8</sup> The (absolute value of the) elasticity of  $A^*$  with respect to  $n$  is 1, and the elasticity of  $y$  with respect to  $\lambda$  is  $\theta \frac{\lambda}{(\lambda-d)(-d+\theta+\lambda)}$ . For the baseline value of  $d$  and  $\theta = 0.02$ , the elasticity of  $A^*$  with respect to  $\lambda$  ranges between 0 and 0.3. Thus, for this special case, the equilibrium decision is much more sensitive to  $n$  than to  $\lambda$ .

The next two sections show that with convex damages, this comparison also holds in the linear equilibrium but may not hold in non-linear equilibria. Hereafter, I consider only convex damages.

## 4.2 The Linear Equilibrium

The linear equilibrium exists for all initial conditions. I report results for the year 2100 stock, a level often used in policy discussions. I set  $r = 0.02 = \theta = \frac{1}{T}$  (a risk-adjusted discount rate of 4%), using  $\lambda \in \{0, 0.02, 0.1, \infty\}$  to represent a range of altruism,  $n \in \{1, 5, 10\}$  to represent a range of fragmentation amongst tribes, and  $\Omega \in \{1, 3\}$  to represent a range of beliefs about the cost of climate change relative to the cost of abatement.

The most important two conclusions from this section are (under the linear equilibrium): (i) The degree of fragmentation ( $n$ ) has a much larger effect on the equilibrium outcome, compared to the degree of altruism ( $\lambda$ ); and (ii) for moderately high levels of fragmentation, the degree of altruism is unimportant.

---

<sup>8</sup>Under the first three calibration assumptions given above,  $d = -8.351 \times 10^{-3}$  and  $C = -2.104 \times 10^{-2}$ .

Table 1 shows the equilibrium reduction in year 2100 stock, as a percent of the BAU stock in that year (hereafter, “stock reduction”). The left panel corresponds to the agent with exponentially distributed lifetime, and the right panel corresponds to the agent with finite lifetime, both with paternalistic altruism. The rows corresponding to  $\lambda = 0.02 = r$  in the two panels are identical, because for  $\lambda = r$  the two problems are equivalent; both have the constant discount rate  $r + \theta = r + \frac{1}{T}$ . Comparisons between other rows illustrate the points made in Section 2.3; the rest of this section discusses only the left panel, for the agent with exponentially distributed lifetime.

The first element in an entry corresponds to low damages,  $\Omega = 1$ , and the second element corresponds to moderate damages,  $\Omega = 3$ . Keeping the year 2100 stock to 400, 500 and 600 ppm requires, respectively, stock reductions of 43%, 29% and 14% of the BAU levels of 700 ppm. For the parameter values here, the equilibrium year 2100 stock ranges from 470 ppm (with  $n = 1$  and  $\lambda = 0$ ) to just under 700 ppm for large values of  $n$  and  $\lambda$ .

$\lambda \backslash n$	1	5	10
0.0	(19, 33)	(6, 12)	(3, 8)
I. 0.02	(11, 25)	(2, 7)	(1, 4)
0.1	(8, 19)	(2, 5)	(1, 2)
$\infty$	(7, 17)	(2, 4)	(1, 2)

$\lambda \backslash n$	1	5	10
0.0	(24, 37)	(7, 15)	(4, 10)
II. 0.02	(11, 25)	(2, 7)	(1, 4)
0.1	(5, 14)	(1, 3)	(0.5, 2)
$\infty$	(3, 9)	(0.7, 2)	(0.3, 1)

Table 1: Reduction in the year 2100 stock as a % of BAU level (“stock reduction”). First element corresponds to  $\Omega = 1$  and second corresponds to  $\Omega = 3$ . Table 1.I corresponds to the agent with exponentially distributed lifetime and Table 1.II corresponds to the agent with finite lifetime.

A decrease in altruism (larger  $\lambda$ ) or contemporaneous cooperation (larger  $n$ ) of course are associated with higher stocks (lower stock reduction). The facts that  $n$  and  $\lambda$  have different units, and that the elasticities that we care about are not constant, complicate a comparison of the equilibrium effects of these parameters. However,  $\lambda = \infty$  is an upper bound, and the equilibria under  $\lambda = 0.1$  and  $\lambda = \infty$  are similar; thus,  $\lambda = 0.1$  corresponds to a very low level of altruism. At  $n = \infty$  the stock reduction equals 0. Stock reductions for  $n = 10$  are small, suggesting that  $n = 10$  corresponds to a high degree of fragmentation. I therefore consider the intermediate values  $\lambda = 0.02$  and  $n = 5$  to represent “moderate” levels of altruism and fragmentation. I emphasize the equilibrium effects of changing only one or

both of the parameters to these levels, from their lower bounds  $\lambda = 0$  and  $n = 1$ .

An increase of  $n$  from 1 to 5 causes a much lower stock reduction, relative to an increase in  $\lambda$  from 0 to 0.02. Changing both parameters from their lower bound to the moderate levels causes, for  $\Omega = 3$ , a nearly 80% fall in equilibrium stock reduction; an increase in only  $\lambda$  causes a 24% fall in stock reduction and an increase in only  $n$  causes a 64% fall in stock reduction. By these measures, a decrease in contemporaneous cooperation has a much larger effect on the equilibrium than does a decrease in altruism. (The comment following Remark 1 notes that this comparison also holds when marginal damages are constant.)

The equilibrium effect of lower altruism is smaller than the literature cited in the Introduction, which emphasizes the social discount rate, might suggest. That literature uses an infinitely lived agent, where the discount rate (or the pure rate of time preference) measures the value that one places on both one's own and on descendants' future consumption (or utility). At least a portion of the disagreement amongst the authors cited arose from ethical questions about the appropriate treatment of future generations. The parameters  $r$  and  $\lambda$  disentangle selfish from altruistic considerations.

Increasing  $\Omega$  from 1 to 3, holding  $n$  and  $\lambda$  fixed, increases the percent stock reduction in the year 2100 stock by a factor of 2 – 3 in most cases. An increase in  $n$  from 1 to 5 leads to a substantial drop in this stock reduction, much larger in both percentage and absolute terms than the difference caused by a change of  $n$  from 5 to 10. This type of result, where cooperation breaks down rapidly even at small levels of fragmentation, is familiar from industrial organization models. For example, the Cournot equilibrium price may be close to the competitive price when there are only a few firms. A larger value of  $n$  decreases the effect of changes in altruism. If the level of contemporaneous cooperation is low, the degree of altruism is rather unimportant. For example, at  $n = 10$  equilibrium stock is close to BAU levels for all values of  $\lambda$ , so changes in  $\lambda$  do not have much effect. In contrast, even at high values of  $\lambda$ , a change in  $n$  has a large effect on the equilibrium.

For  $\Omega = 3$ ,  $n = 1$  and  $\lambda = 0$ , the equilibrium steady state is 496 ppm, an almost 50% reduction in the BAU steady state. The same steady state results from a standard optimization problem with a constant discount rate of 0.9%. As noted above, the stock that maximizes the steady state utility flow (the “optimal static steady state”) is 402 ppm; the same steady state results from the social planner's problem in the limit as the constant discount

rate approaches 0.

### 4.3 Nonlinear Equilibria

I consider non-linear equilibria only for the agent with exponentially distributed lifetime. The most important conclusions from this section (under non-linear equilibria) are: (i) The importance of the degree of altruism is comparable to that of the degree of fragmentation; and (ii) the degree of altruism may be especially important in determining the outcome when the degree of fragmentation is large. Thus, consideration of non-linear equilibria can reverse the conclusions obtained under the assumption of a linear equilibrium.<sup>9</sup>

Define  $\beta$  as the infimum of steady state stocks that can be supported in a MPE, as a ratio of the optimal static steady stock. Proposition 2 implies that for  $n = 1$ ,  $\beta \rightarrow 1$  as  $\lambda \rightarrow 0$ . In general,  $\beta$  is a function of  $n, \lambda$ , and the other model parameters. The solid graphs in Figure 2 show  $\beta$  for  $0 \leq \lambda \leq r = 0.02$  for  $n = 1$  and  $n = 5$ , given  $\Omega = 3$  and the other parameter values discussed above. The dashed graphs show, for different values of  $\lambda$  and  $n$ , the steady state in the linear equilibrium as a ratio of the optimal static steady state.

For  $\lambda = r$  and  $n = 1$  the game is a standard optimal control problem, so the linear equilibrium is the unique MPE; consequently, for  $n = 1$  the solid and dashed graphs converge at  $\lambda = r$ . For  $\lambda < r$  or  $n > 1$  there are multiple equilibria, as explained in Section 3.2. In this circumstance, the gap between the smallest stable MPE steady state and the steady state under the linear equilibrium (reflected in the distance between a solid and a dashed curve) increases with both  $n$  and  $r - \lambda$ . The greater is  $n$  or  $r - \lambda$ , the greater is the coordination problem arising from the multiplicity of equilibria.

The graph of  $\beta$  for  $n = 5$  lies strictly above the graph for  $n = 1$ , although there is only approximately a 1% difference at  $\lambda = 0$ . The difference between the graphs increases with  $\lambda$ . Thus, for small values of  $\lambda$  it is possible to support, as a MPE steady state, a stock of carbon close to the optimal static steady state regardless of whether  $n = 1$  or  $n = 5$ . However, for larger values of  $\lambda$ , the smallest possible steady state in a MPE is much larger for  $n = 5$  compared to  $n = 1$ . Thus, if one considers non-linear equilibria, the altruism

---

<sup>9</sup>It may be important to bear this possible reversal in mind in evaluating results that focus on a single (usually linear) equilibrium in models with non-constant discounting.

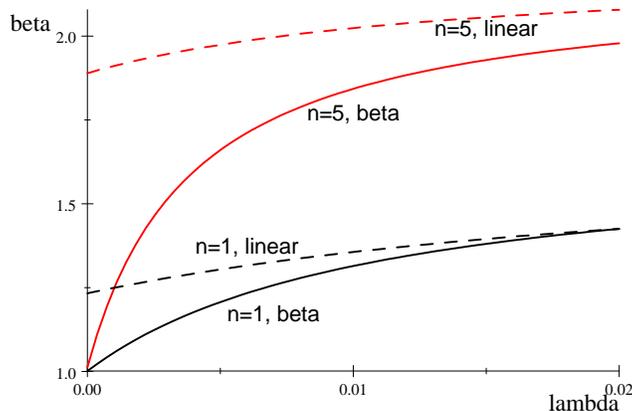


Figure 2: Solid curves show  $\beta$ , the ratio of the infimum of stable MPE steady state to the static optimum steady state. Dashed curves show ratio of steady state in linear equilibrium to the static optimum steady state.  $\Omega = 3$ .

parameter becomes more important for larger  $n$ . The previous section notes that if one considers only the linear equilibria, the altruism parameter is less important for large  $n$ .

Non-linear equilibria – unlike the linear equilibrium – may be defined over only a strict subset of state space. Tsutsui and Mino’s (1990) graphical analysis of the domain of existence of non-linear equilibria is not possible in the three-dimensional system  $(S, A, K)$  here. Nonlinear equilibria can be obtained by solving the system of differential equations for the control variable and  $K$ , as functions of the state variable, as an initial value problem; the initial value is a steady state (sic).

Figure 3 shows three equilibrium decision rules, together with the set of steady states (the dashed line) for  $n = 5$ ,  $\Omega = 3$ , and  $\lambda = 0.00001$  and the other parameters discussed above. For combinations of  $S, A$  strictly below the dashed line, the stock is increasing over time. The label on each solid curve shows the steady state induced by that policy function. The positively sloped line is the linear equilibrium, corresponding to a steady state of 760 ppm. In this equilibrium, the level of abatement increases with the stock. The non-monotonic policy function drives the state to 745 ppm. In this equilibrium, abatement starts out high and falls until the stock gets close to its steady state; in the neighborhood of the steady state, abatement increases.

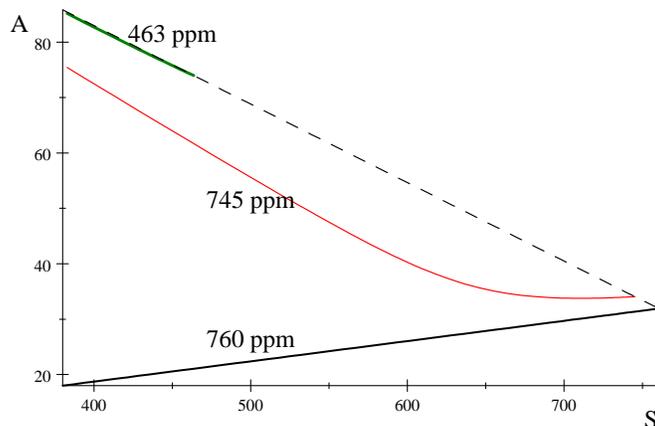


Figure 3: Abatement,  $A$  (as a percent of BAU emissions) as a function of the stock,  $S$  (ppm), for  $n = 5$  and  $\lambda = 0.00001$ . The three MPE policy functions drive the state from the initial value 380 ppm to different steady states, shown by the labels on the curves. The dashed line shows the combinations of abatement and stock that maintain a steady state.

Even though the two steady states, 760 and 745, are similar, the speed of increase is much higher in the linear equilibrium. As a consequence of the scale of the figure, the curve labeled “463 ppm” appears to be coincident with the dashed line. The actual curve lies below the dashed line except at the steady state, so this policy rule drives any stock  $380 \leq S_0 \leq 463$  to the steady state 463. In this equilibrium, abatement falls over time, allowing the stock to grow slowly until it reaches the steady state. The stock 463 is about 10% larger than the theoretically determined minimum steady state in a MPE with  $n = 5$ ,  $\Omega = 3$ , and  $\lambda = 0.00001$ .

Figure 3 provides two important pieces of information. First, the domain of existence of non-linear equilibria that maintain low stock levels may include the current stock level. Thus, the multiplicity of equilibria is not a mere curiosity, as would arguably be the case if non-linear equilibria were defined only in the neighborhood of their steady state. Second, equilibrium decision rules may be strictly increasing, decreasing, or non-monotonic in the state variable. Where the equilibrium decision rule is increasing in the stock, actions can be considered strategic substitutes in the following sense: a higher level of abatement by an agent today leads to a lower future stock,

with lower future equilibrium abatement by all agents. Similarly, when the decision rule is decreasing in the stock, actions can be considered strategic complements. Thus, actions might be strategic complements, substitutes, or changing from the former to the latter over time. Decision rules that produce low stock trajectories imply strategic complementarity, and they require especially large abatement efforts early in the trajectory, while the stock is still low.

The possibility that equilibrium abatement (or the equilibrium carbon tax) falls over time, ignores the likely reality of convex adjustment costs in emissions levels. It also ignores that abatement costs may fall over time, an assumption that is important to the “policy ramp” in DICE. These simplifications, and many others, mean that this model is not suitable for detailed policy advice; as noted above, the model does not have that normative objective. Nevertheless, it is interesting to note that Golosov et. al’s (2011) normative model recommends a decreasing carbon tax. That result is due to the Hotelling aspect of their model, and the related “Green paradox”.

## 5 Discussion

The provision of a long-lived public good depends on the ability of contemporaneous agents to cooperate, and on their degree of altruism towards future generations. In the climate policy context, the academic economic literature has emphasized the role of the social discount rate. The popular press and other disciplines emphasize the difficulty of getting different nations to form an agreement that internalizes the effect of national emissions. Although limited altruism and lack of cooperation are both obstacles to effective climate policy, their relative importance is not obvious. This relation depends on the specifics of the problem, in particular on parameters that determine the costs and benefits of abatement and on the dynamics of the climate system. It also depends on the type of equilibria that one considers.

The manner in which these factors affect the relative importance of the two obstacles to effective policy has, I believe, not previously been studied. Embedding an OLG framework in a differential game provides a means of addressing this issue formally. I use this model to study climate policy.

If one considers only the linear equilibrium in the climate application, two clear conclusions emerge: (i) the degree of contemporaneous cooperation has significantly greater effect on equilibrium policy than does the degree of

intergenerational altruism, and (ii) once the degree of cooperation is even moderately low, the outcome is insensitive to the degree of altruism. However, this class of model admits many MPE, even under the restriction that these induce a differentiable value function. Non-linear equilibria can support a broad range of outcomes. Some of these outcomes result in high levels of abatement and low pollution stocks. Taking into consideration these nonlinear equilibria, the outcome may be more sensitive to the degree of altruism than to the level of cooperation, and the degree of altruism may be important especially when the degree of cooperation is low. That is, broadening the scope of enquiry to include nonlinear equilibria, may reverse conclusions concerning the relative importance of cooperation and altruism.

It is tempting, but I believe unproductive, to try to assess the relative plausibility of linear versus non-linear equilibria. The existence of non-linear equilibria requires an infinite time horizon (or a random finite horizon, with unbounded support). However, the infinite (or random, unbounded) horizon is a realistic feature of the problem. Linear equilibria have the virtue of being defined for the entire real line, whereas non-linear equilibria are defined over only a subset of the real line. However, the examples in Section 4.3 show that the domain of non-linear equilibria is large enough to be relevant. Linear equilibria are easier to compute; but given the degree of abstraction of this model, I doubt that ease of computability is a convincing basis for choosing a particular type of equilibrium. A persuasive argument for equilibrium selection in this setting may be found, but I have not yet seen it.

It is important to recognize that the provision of a long-lived public good by countries that care about only their own current and future citizens is *generically* a coordination problem. The logic of Nash's noncooperative equilibrium does not doom us to bad outcomes, even if we dismiss "trigger" or other punishment strategies as implausible. International negotiations on climate policy are important, even if they do not result in enforceable agreements. Negotiations make coordination easier to achieve.

## References

- ANDREONI, J. (1989): "Giving with impure altruism: applications to charity and Ricardian equivalence," *Journal of Political Economy*, 97, 1447–1458.
- ARCHER, D., AND V. BROVKIN (2008): "The millennial atmospheric lifetime of antropogenic CO<sub>2</sub>," *Climatic Change*, 90, 283–297.
- BARRO, R. (1999): "Ramsey meets Laibson in the neoclassical growth model," *Quarterly Journal of Economics*, 114, 1125–52.
- BERGSTROM, T. (2006): "Benefit-Cost in a Benevolent Society," *American Economic Review*, 96(1), 339–351.
- BLANCHARD, O. J. (1985): "Debts, deficits and finite horizons," *Journal of Political Economy*, 93, 223–247.
- CALVO, G., AND M. OBSTFELD (1988): "Optimal time-consistent fiscal policy with finite lifetimes," *Econometrica*, 56, 411–32.
- DIAMOND, P. A. (1965): "National Debt in a Neoclassical Growth Model," *The American Economic Review*, 55(5), 1126–1150.
- DOCKNER, E., AND N. VAN LONG (1993): "International pollution control: cooperative versus non-cooperative strategies," *Journal of Environmental Economics and Management*, 24, 13–29.
- 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 (2012): "Equilibrium management of fisheries with altruistic overlapping generations," University of British Columbia Working Paper.
- EKELAND, I., AND A. LAZRAK (2010): "The golden rule when preferences are time inconsistent," *Mathematical and Financial Economics*, 4(1).
- ENVIRONMENTAL PROTECTION AGENCY (2010): "Guidlines for preparing economic analyses," EPA 240-R- 0-001.
- FERSHTMAN, C., AND M. KAMIEN (1987): "Dynamic Duopolistic Competition with Sticky Prices," *Econometrica*, 55, 1151–1164.

- GERLAGH, R., AND M. LISKI (2012): “Carbon prices for the next thousand years,” Unpublished Working Paper.
- 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 (2011): “Optimal taxes on fossil fuels in general equilibrium,” NBER Working Paper 1738.
- HARRIS, C., AND D. LAIBSON (2001): “Dynamic Choices of Hyperbolic Consumers,” *Econometrica*, 69(5), 935–957.
- HAURIE, A., J. KRAWCZYK, AND G. ZACCOUR (2012): *Games and Dynamic Games*. Imperial College Press; World Scientific, Hackensack, New Jersey.
- JACKSON, M., AND L. YARIV (2011): “Collective dynamic choice: the necessity of time inconsistency,” Department of Economics, Stanford, Working Paper.
- 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 J. ZHANG (2006): “Regulation with Anticipated Learning about Environmental Damage,” *Journal of Environmental Economics and Management*, 51, 259–280.
- KRUSELL, P., AND A. SMITH (2003): “Consumption-saving decisions with quasi-geometric discounting,” *Econometrica*, 71(1), 365–75.
- LAIBSON, D. (1997): “Golden eggs and hyperbolic discounting,” *Quarterly Journal of Economics*, 62, 443–78.
- LEVHARI, D., AND L. MIRMAN (1980): “The Great Fish War: An Example Using a Dynamic Cournot-Nash Solution,” *Bell Journal of Economics*, 11, 322–334.

- 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. (1994): *Managing the Global Commons: The economics of the greenhouse effect*. Cambridge, MA: MIT Press.
- NORDHAUS, W. D. (2007): “A Review of the Stern Review on the Economics of Climate Change,” *Journal of Economic Literature*, (3), 686 – 702.
- NORDHAUS, W. D. (2008): *A Question of Balance*. Yale University Press.
- RAMSEY, F. (1928): “A Mathematical Theory of Savings,” *Economic Journal*, 38, 543–59.
- RAY, D. (1987): “Nonpaternalistic intergenerational altruism,” *Journal of Economic Theory*, (41), 112–132.
- REYNOLDS, S. (1987): “Capacity Investment, Preemption and Commitment,” *International Economic Review*, 28, 69–88.
- 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 (2010): “Trading off generations: infinitely lived agent versus OLG,” DARE Working Paper.
- 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.
- SUMAILA, U., AND C. WALTERS (2005): “Intergenerational discounting: a new intuitive approach,” *Ecological Economics*, 52, 135–142.
- TRAEGER, C. (2012): “Why uncertainty matters - Discounting under Intertemporal risk aversion and ambiguity,” DARE Working Paper 1092R2.

- TSUTSUI, S., AND K. MINO (1990): “Nonlinear Strategies in Dynamic Duopolistic 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.
- VIELLE, 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.
- YAARI, M. E. (1965): “Uncertain lifetime, life insurance and the Theory of the Consumer,” *Review of Economic Studies*, 32, 137–150.
- ZUBER, S. (2010): “The aggregation of preferences: can we ignore the past?,” *Theory and Decision*, 70, 367 – 384.

## A Appendix Discounting

I begin with a discrete time model, in which each period lasts for  $\varepsilon$  units of time (e.g., years), and obtain the discount factors in the continuous time setting by passing to the limit, as  $\varepsilon \rightarrow 0$ . This approach is useful for explaining the meaning of the different models, and particularly for deriving the discount functions under pure altruism and for a finite lifetime. Ekeland and Lazrak (2010) provide the continuous time discount factor under exponentially distributed lifetime and paternalistic altruism, without the discrete time detour; I include that case so that this appendix is self-contained.

### A.1 The discrete time model

Table 2 introduces notation used to obtain concise expressions of the discrete time discount factors. The pure rate of time preference that an agent uses to evaluate her selfish component of welfare is  $r$ , and  $\lambda$  is the rate she uses to evaluate the utility or welfare of future generations;  $\rho = e^{-r\varepsilon}$  and  $\delta = e^{-\lambda\varepsilon}$  are the corresponding discount factors. For the case of exponentially distributed lifetime,  $\theta$  is the mortality = birth = hazard rate, so  $\gamma = r + \theta$  is the risk-adjusted discount rate and  $\alpha = e^{-\gamma\varepsilon}$  is the corresponding risk-adjusted discount factor. With a constant population normalized to 1,  $b = 1 - e^{-\theta\varepsilon}$  is the mass of agents born at the end of a period of length  $\varepsilon$ . In order to make the models with exponentially distributed and deterministic lifetimes comparable, I assume throughout that the lifetime in the latter case equals  $T = \frac{1}{\theta}$ , the expected lifetime under the exponential distribution.

	mortality	selfish time preference	risk adjusted discounting	altruism weight
continuous time rates	$\theta = \frac{1}{T}$	$r$	$\gamma = r + \theta$	$\lambda$
discrete time factors	$b = 1 - e^{-\theta\varepsilon}$	$\rho = e^{-r\varepsilon}$	$\alpha = e^{-\gamma\varepsilon}$	$\delta = e^{-\lambda\varepsilon}$

Table 2: Parameters that appear in the discount functions

#### A.1.1 Exponentially distributed lifetime, paternalistic altruism

Agents use the risk-adjusted discount factor  $e^{-\gamma\varepsilon}$  to evaluate their own future utility. Agents alive in a period are identical, so any can be chosen as the

social planner who makes the decision about investment in the public good in that period. The social planner alive in period 0 gives weight  $e^{-\gamma\epsilon t}$  to the period  $t$  utility of agents currently alive, and the weight  $(1 - e^{-\theta\epsilon}) e^{-\lambda\epsilon\tau} e^{-\gamma\epsilon(t-\tau)}$  to the period  $t$  utility of agents born in period  $\tau \leq t$ . There are  $1 - e^{-\theta\epsilon}$  of these agents, each of whom discounts her period  $t$  utility at  $e^{-\gamma\epsilon(t-\tau)}$ , and the current social planner values that utility at  $e^{-\lambda\epsilon\tau}$ .

The total weight that the current social planner puts on period  $t$  utility flow (the discount factor) is the sum of the selfish and altruistic components for those currently alive:

$$D(t; \epsilon) = e^{-\gamma\epsilon t} + (1 - e^{-\theta\epsilon}) \sum_{\tau=1}^t e^{-\lambda\epsilon\tau} e^{-\gamma\epsilon(t-\tau)}. \quad (14)$$

### A.1.2 Exponentially distributed lifetime, pure altruism

The expected present value of the flow of utility of an agent alive in period  $\tau$  (the selfish component of her welfare) is

$$\sum_{t=0}^{\infty} e^{-\gamma\epsilon t} u_{\tau+t}.$$

An agent's total welfare equals the sum of her own selfish utility and the utility that she receives from the welfare of agents who are born in the future. Denote the total welfare of the agent born in period  $\tau + t$  as  $V_{\tau+t}$ . In each period,  $1 - e^{-\theta\epsilon}$  agents are born. The agent currently alive attaches the weight  $(1 - e^{-\theta\epsilon}) e^{-\lambda\epsilon t}$  to the welfare of the generation born  $t$  periods in the future. Thus, the welfare of the agent alive at period  $\tau$  is<sup>10</sup>

$$V_{\tau} = \sum_{t=0}^{\infty} e^{-\gamma\epsilon t} u_{\tau+t} + (1 - e^{-\theta\epsilon}) \sum_{t=1}^{\infty} e^{-\lambda\epsilon t} V_{\tau+t}. \quad (15)$$

The following proposition gives the formula for the discount factor  $D(t; \epsilon)$  such that welfare  $V_{\tau}$  equals the present discounted value of future utility flows  $u_s$ , i.e.:

$$V_{\tau} = \sum_{s=0}^{\infty} D(s; \epsilon) u_{\tau+s} \quad \text{with} \quad D(0; \epsilon) = 1. \quad (16)$$

---

<sup>10</sup>Replacing the first term on the right side of equation (15) by  $\tilde{u}_{\tau} \equiv \sum_{t=0}^{\infty} \alpha^t u_{\tau+t}$  and defining  $a(t) = (1 - e^{-\theta\epsilon}) e^{-\lambda\epsilon t}$ , results in equation (2) of Saez-Marti and Weibull (2005).

**Proposition 3** Assume  $e^{-\lambda\varepsilon} (2 - e^{-\theta\varepsilon}) - e^{-\gamma\varepsilon} \neq 0$  and  $\lambda > \theta$ . (i) The additively separable function  $V_t$  defined in equation (16) equals the solution to the recursion in equation (15) if and only if the discount factor equals

$$D(t; \varepsilon) = \frac{e^{-\gamma\varepsilon t} (e^{-\lambda\varepsilon} - e^{-\gamma\varepsilon}) + e^{-\lambda\varepsilon} (e^{-\lambda\varepsilon} (2 - e^{-\theta\varepsilon}))^t \theta\varepsilon}{e^{-\lambda\varepsilon} (2 - e^{-\theta\varepsilon}) - e^{-\gamma\varepsilon}}. \quad (17)$$

(ii) (a)  $D(t; \varepsilon)$  is positive, bounded and approaches 0 as  $t \rightarrow \infty$ . (b)  $\sum_{s=0}^{\infty} D(s; \varepsilon)$  is bounded, so  $V$  is bounded given that  $u$  is bounded.

The term  $e^{-\lambda\varepsilon} (2 + e^{-\theta\varepsilon}) = \delta (1 + b)$  appears in a formula used to prove Proposition 3. The following lemma establishes a bound on this term.

**Lemma 2** The necessary and sufficient condition for  $\delta b + \delta < 1$  for all  $\varepsilon \geq 0$  is  $\lambda > \theta$ .

**Proof.** Necessity: Using the definitions of  $\delta$  and  $b$ ,  $\delta b + \delta = e^{-\lambda\varepsilon} (2 - e^{-\theta\varepsilon})$ . The first order approximation of this expression, evaluated at  $\varepsilon = 0$ , is  $1 + (\theta - \lambda)\varepsilon + o(\varepsilon)$ . This expression is less than 1 if and only if  $\lambda > \theta$ . Sufficiency: Use  $\frac{d(e^{-\lambda\varepsilon}(2 - e^{-\theta\varepsilon}))}{d\varepsilon} = e^{-\lambda\varepsilon} ((\theta + \lambda)e^{-\theta\varepsilon} - 2\lambda)$ . If  $\lambda > \theta$  then  $(\theta + \lambda)e^{-\theta\varepsilon} - 2\lambda < 2\lambda(e^{-\theta\varepsilon} - 1) < 0$  so  $e^{-\lambda\varepsilon} (2 - e^{-\theta\varepsilon})$  is decreasing in  $\varepsilon$ . Therefore, it is negative for all  $\varepsilon \geq 0$  if and only if it is negative for  $\varepsilon = 0$ . ■

**Proof.** Proposition 3. Substituting equation (16) into (15) gives

$$V_\tau = \sum_{s=0}^{\infty} \alpha^s u_{\tau+s} + b \sum_{s=1}^{\infty} \delta^s \left( \sum_{k=0}^{\infty} D_k u_{\tau+s+k} \right)$$

Making a change of variables,  $t = s + k$  and then reversing the order of summation and simplifying yields

$$V_\tau = u_\tau + \sum_{t=1}^{\infty} \left( \alpha^t + b \sum_{s=1}^t \delta^s D_{t-s} \right) u_{\tau+t}. \quad (18)$$

Equating coefficients in equations (16) and (18) implies

$$D_t = \alpha^t + b \sum_{s=1}^t \delta^s D_{t-s} \quad (19)$$

with initial condition  $D_0 = 1$ . The manipulations above are valid because  $\sum_{s=0}^{\infty} D(s; \varepsilon)$  is bounded, as established in part (ii) below.

An inductive proof establishes part (i). Setting  $t = 0$ , the trial solution in equation (17) satisfies the initial condition  $D_0 = 1$ . Suppose that for  $t \geq 0$ , the trial solution solves the recursion (19) for  $\tau \leq t$ . I need to show that this hypothesis implies that the trial solution solves the recursion for  $t + 1$ . The hypothesis implies

$$D_{t+1} = \alpha^{t+1} + b \sum_{s=1}^{t+1} \delta^s D_{t+1-s} = \alpha^{t+1} + b \sum_{s=1}^{t+1} \delta^s \frac{\alpha^{t+1-s} (\delta - \alpha) + \delta b (\delta b + \delta)^{t+1-s}}{-\alpha + \delta b + \delta}.$$

Simplifying the last expression gives

$$D_{t+1} = \frac{\alpha^{t+1} (\delta - \alpha) + \delta b (\delta b + \delta)^{t+1}}{-\alpha + \delta b + \delta},$$

as was to be shown.

(ia) The facts that  $\alpha < 1$  and  $(\delta b + \delta) < 1$  (from Lemma 2) imply that  $D_t$  is bounded and approaches 0 as  $t \rightarrow \infty$ . To show that  $D_t > 0$ , consider three cases, where  $\delta > \alpha$ , where  $\alpha > \delta > \frac{\alpha}{b+1}$ , and where  $\delta < \frac{\alpha}{b+1}$ . In the first case, the numerator and denominator of  $D_t$  are positive by inspection. In the second case, the denominator is positive. The numerator is positive iff

$$L(t) \equiv \left( \frac{\delta b + \delta}{\alpha} \right)^t > \frac{\alpha - \delta}{\delta b}. \quad (20)$$

The function  $L(t)$  is increasing in  $t$ ,  $L(0) = 1$ , and  $\frac{\alpha - \delta}{\delta b} < 1$  because  $\delta > \frac{\alpha}{b+1}$ ; therefore, inequality (20) is satisfied. Consequently, the numerator of  $D_t$  is positive, so  $D_t$  is positive. In the third case, the denominator of  $D_t$  is negative and the numerator is negative iff

$$L(t) \equiv \left( \frac{\delta b + \delta}{\alpha} \right)^t < \frac{\alpha - \delta}{\delta b}. \quad (21)$$

Here,  $L(0) = 1$  and is decreasing in  $t$  and the right side of inequality (21) is greater than 1, so the inequality is satisfied.

(iib) The fact that  $D$  is the sum of two geometrically decreasing terms, means that it's infinite sum is bounded. The sum equals  $\frac{1 - \delta}{(1 - \delta b - \delta)(1 - \alpha)}$ . ■

### A.1.3 Deterministic lifetime, paternalistic altruism

Here, in contrast to the case of exponentially distributed lifetimes, agents alive in a period are different: the older ones will die sooner than the younger ones. Agents therefore have different views about the future benefits of the public good. The social planner in a period is utilitarian: she maximizes the sum of the discounted utility of those currently alive, plus the value that those agents give to the utility of the not-yet born.

Agents live for  $T \geq 1$  periods.<sup>11</sup> For  $T = 2$ , this model produces Laibson's (1997)  $\beta, \delta$  model of quasi-hyperbolic discounting. Next period, the fraction  $(1 - \frac{1}{T})$  of current agents will still be alive. The agents alive at  $t = 0$  discount their future utility using the selfish discount factor  $e^{-r\varepsilon}$ . The utilitarian social planner representing the agents alive at  $t = 0$  discounts their  $t = 1$  utility at  $e^{-r\varepsilon} (1 - \frac{1}{T})$ , which takes into account both the agents' impatience and the death of some of the agents alive at  $t = 0$ . For  $t \leq T - 1$  the discount factor that this social planner uses to evaluate the future utility of agents currently alive is  $e^{-r\varepsilon t} (1 - \frac{t}{T})$  and for  $t \geq T$  the discount factor is 0, because all of the original agents will have died.

With a constant population,  $\frac{1}{T}$  new agents are born in each period. The agents alive in period 0 discount future generations' utility at  $e^{-\lambda\varepsilon}$ ; for example, the weight that the agents alive in period 0 give to those born in year 1 is  $\frac{e^{-\lambda\varepsilon}}{T}$ , a factor that accounts for both the current agents' altruism and the fact that  $\frac{1}{T}$  new agents arrive in each period. The agents who arrive in year 1 discount their own next period utility at  $e^{-r\varepsilon}$ , and because the agents alive in period 0 discount those agents' utility at  $e^{-\lambda\varepsilon}$ , the weight that the period 0 agents place on the  $t = 2$  utility of the agents who arrived at  $t = 1$  is  $\frac{e^{-\lambda\varepsilon} e^{-r\varepsilon}}{T}$ . Those alive at time 0 place the weight  $\frac{e^{-\lambda\varepsilon i} e^{-r\varepsilon(t-i)}}{T}$  on the period  $t > 0$  utility of agents who arrive in period  $i$ , for  $t - T < i \leq t$ .

The weight that the  $t = 0$  social planner places on the utility flow at period  $t \geq 0$ ,  $D(t; \varepsilon)$ , equals the sum of the weights that the agents whom she represents place on the selfish and the altruistic components of their preferences. As noted above, the selfish component of this discount factor has a different structure for  $t < T$  and for  $t \geq T$ . Therefore, the function  $D(t; \varepsilon)$  also has a different structure for  $t < T$  and for  $t \geq T$ .

In the discrete time setting, the weight that the social planner places on

---

<sup>11</sup>Sumaila and Walters (2005) propose a similar model, but their formulae account for birth but not for death. With constant population, the discount factor must account for both death and birth.

utility flow at a future time equals the sum of the weight attributed to agents currently alive and to those have not yet been born. This sum has a different structure, depending on whether  $t \leq T - 1$  or  $t \geq T$ . For  $\lambda \neq r$  the discount factor is

$$D(t; \varepsilon) = \begin{cases} e^{-r\varepsilon t} \left(1 - \frac{t}{T}\right) + \frac{1}{T} \sum_{\tau=1}^t e^{-\lambda\varepsilon\tau} e^{-r\varepsilon(t-\tau)} = \\ \frac{1}{T} \left( \frac{e^{-\lambda\varepsilon} (e^{-\lambda\varepsilon t} - e^{-r\varepsilon t})}{(e^{-\lambda\varepsilon} - e^{-r\varepsilon})} + (T - t) e^{-r\varepsilon t} \right) \text{ for } t \leq T - 1, \\ \frac{1}{T} \sum_{\tau=t-T+1}^t e^{-\lambda\varepsilon\tau} e^{-r\varepsilon(t-\tau)} = \\ \frac{1}{T} e^{-\lambda\varepsilon(t+1-T)} \frac{e^{-\lambda\varepsilon T} - e^{-r\varepsilon T}}{e^{-\lambda\varepsilon} - e^{-r\varepsilon}} \text{ for } t \geq T. \end{cases} \quad (22)$$

For  $T = 2$ , the discount factor at  $t = 1$  is  $\frac{e^{-r\varepsilon} + e^{-\lambda\varepsilon}}{2}$  and the discount factor at  $t > 1$  is  $\frac{1}{2} e^{-\lambda\varepsilon(t-1)} (e^{-\lambda\varepsilon} + e^{-r\varepsilon})$ . Defining  $\beta = \frac{e^{-r\varepsilon} + e^{-\lambda\varepsilon}}{2}$ ,  $\delta = e^{-\lambda\varepsilon}$  produces the  $\beta, \delta$  model.

## A.2 The continuous time limit

In the discrete time setting I do not need to distinguish between the number of units of time (e.g. years) and the number of periods. This distinction is important in using the discrete time model to obtain (by passing to a limit) the continuous time discount factors. That is,  $t$  and  $\tau$  refer to period indices in the discussion of discrete time models, and they refer to units of calendar time in the discussion of continuous time models; similarly,  $T$  refers to the number of periods that the agent lives in the discrete time setting, and the number of units of time that she lives in the continuous time setting. For fixed  $\varepsilon$ , this notation does not present an issue, but here I want to consider the limiting case as  $\varepsilon \rightarrow 0$ . For this purpose, it is important to maintain the distinctions between the period index and units of time.

For this appendix (only) I use  $\tau$  and  $\Gamma$  to refer exclusively to units of time and  $t$  and  $T$  to refer exclusively to number of periods. If a period lasts  $\varepsilon$  units of time,  $\tau = t\varepsilon$  and  $\Gamma = T\varepsilon$ . Using these definitions, I set  $t = \frac{\tau}{\varepsilon}$  and  $T = \frac{\Gamma}{\varepsilon}$  and use the definitions in the last row of Table 2 to write the discrete time discount factors as a function of units of time (rather than number of periods) and the length of each period,  $\varepsilon$ . It is then a simple matter to take limits as  $\varepsilon \rightarrow 0$ .

### A.2.1 Exponentially distributed lifetime, paternalistic altruism

Equation (14) can be simplified to

$$D(t; \varepsilon) = \frac{\left( (e^{-\lambda\varepsilon} - e^{-\gamma\varepsilon}) e^{-\gamma\varepsilon t} + (1 - e^{-\theta\varepsilon}) e^{-\lambda\varepsilon} (e^{-\lambda\varepsilon t} - e^{-\gamma\varepsilon t}) \right)}{(e^{-\lambda\varepsilon} - e^{-\gamma\varepsilon})}.$$

This simplification assumes that  $\gamma \neq \lambda$ ; the case  $\lambda = \gamma$  follows from L'Hospital's Rule.

Define  $\mu(\varepsilon) = \frac{b\delta}{\alpha - \delta} = \frac{(1 - e^{-\theta\varepsilon})e^{-\lambda\varepsilon}}{e^{-(r+\theta)\varepsilon} - e^{-\lambda\varepsilon}}$  and use L'Hospital's Rule to obtain  $\lim_{\varepsilon \rightarrow 0} \mu(\varepsilon) = \frac{\theta}{\lambda - \theta - r} = \frac{\theta}{\lambda - \gamma}$ . Using this definition of  $\mu$  and the last row of Table 2, equation (14), the discount function expressed as a function of time, rather than number of periods, is

$$(1 + \mu(\varepsilon)) e^{-(r+\theta)\tau} - \mu(\varepsilon) e^{-\lambda\tau}.$$

Letting  $\varepsilon \rightarrow 0$  gives the continuous time discount factor for calendar time  $\tau$ , equation (1). Ekeland and Lazrak (2010) obtain this formula directly (without the discrete time detour).

### A.2.2 Exponentially distributed lifetime, pure altruism

This section provides the proof of Proposition 1.

**Proof.** Proposition 1 In order to establish the claim, denote the discount rate that the agent with pure altruism applies to future generations as  $\lambda'$  (instead of  $\lambda$ , the rate under paternalistic altruism). Define the function

$$\xi(\varepsilon) = \frac{-\alpha + \delta(b+1)}{\varepsilon} = \frac{-e^{-\gamma\varepsilon} + e^{-\lambda'\varepsilon} (2 - e^{-\theta\varepsilon})}{\varepsilon}$$

and use  $\lim_{\varepsilon \rightarrow 0} \xi(\varepsilon) = r + 2\theta - \lambda'$ . Also define  $\phi(\varepsilon) = \frac{\ln(2 - e^{-\theta\varepsilon})}{\varepsilon}$  and use  $\lim_{\varepsilon \rightarrow 0} \phi(\varepsilon) = \theta$ . Finally, note that

$$(b+1)^t = (2 - e^{-\theta\varepsilon})^{\frac{t}{\varepsilon}} = \exp\left(\tau \frac{\ln(2 - e^{-\theta\varepsilon})}{\varepsilon}\right) = e^{\tau\phi(\varepsilon)},$$

so  $\lim_{\varepsilon \rightarrow 0} (2 - e^{-\theta\varepsilon})^{\frac{t}{\varepsilon}} = e^{\tau\theta}$ . With these definitions and the last row of Table 2, the discrete time discount factor in equation (17) can be written

$$\frac{e^{-\gamma\tau} \left( \frac{e^{-\lambda'\varepsilon} - e^{-\gamma\varepsilon}}{\varepsilon} \right) + e^{-\lambda'\varepsilon} \left( \frac{1 - e^{-\theta\varepsilon}}{\varepsilon} \right) e^{-\lambda'\tau} e^{\tau\phi(\varepsilon)}}{\xi}.$$

Using the limiting expressions given above, the limit of this function as  $\varepsilon \rightarrow 0$  gives the continuous time discount factor for the agent with pure altruism:

$$D(t) = \frac{e^{-\gamma t} (\gamma - \lambda') + \theta e^{-(\lambda' - \theta)t}}{\gamma + \theta - \lambda'}. \quad (23)$$

The right side of equations (1) and (23) are equivalent if and only if  $\lambda' = \lambda + \theta$ .

■

### A.2.3 Finite lifetime, paternalistic altruism

Define  $\nu(\varepsilon) = \frac{\delta - \rho}{\varepsilon} = \frac{(e^{-\lambda\varepsilon} - e^{-r\varepsilon})}{\varepsilon}$  and use  $\lim_{\varepsilon \rightarrow 0} \nu(\varepsilon) = r - \lambda$ . With this definition of  $\nu$  and the last row of Table 2, the discrete time discount factor for  $t < \Gamma$  can be written

$$\begin{aligned} \frac{\varepsilon}{\Gamma(\delta - \rho)} \left( \delta (\delta^t - \rho^t) + (\Gamma - \tau) \rho^t \frac{(\delta - \rho)}{\varepsilon} \right) \\ = \frac{e^{-\lambda\varepsilon} (e^{-\lambda t} - e^{-r t})}{\Gamma \nu} + \frac{(\Gamma - \tau)}{\Gamma} e^{-r t}. \end{aligned}$$

Taking the limit as  $\varepsilon \rightarrow 0$  and rearranging the resulting expression produces the first line of equation (2).

For  $t \geq \Gamma$  the discount factor is

$$\frac{\varepsilon}{\Gamma(\delta - \rho)} e^{\lambda(\Gamma - t - \varepsilon)} (e^{-\lambda\Gamma} - e^{-r\Gamma}).$$

Taking the limit as  $\varepsilon \rightarrow 0$  produces the second line of equation (2).

## B Appendix: Proofs and Remarks

This appendix collects a number of calculations and remarks.

### B.1 Equilibrium conditions for other cases

There are two cases under exponentially distributed lifetime and a single case with a deterministic lifetime, because in the former but not in the latter,  $\lim_{t \rightarrow \infty} \eta(t)$  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 (12) and ), a necessary condition for the MPE is that

$$x_t = \chi(S_t) \equiv \arg \max (u(S_t, x_t) - K(S_t) + J_S(S) F(S, x)), \quad (24)$$

and that  $J(S)$  satisfy the dynamic programming equation<sup>12</sup>

$$\lambda J(S) = (u(S, \chi(S)) - K(S) + J_S(S) F(S, \chi(S))). \quad (25)$$

For the exponential case with  $\lambda > r$ , where  $\lim_{t \rightarrow \infty} \eta(t) = \gamma$ , the fictitious control problem is

$$J(S) = \max \int_0^\infty e^{-\gamma t} (u(S_t, x_t) - K(S_t)) d\tau \quad \text{subject to } \dot{S} = F(S, x), \quad (26)$$

with the side condition (definition):

$$K(S_t) \equiv \int_0^\infty D(\tau) (\eta(\tau) - \gamma) u(S_\tau^*, \chi(S_\tau^*)) d\tau. \quad (27)$$

Equation (1) and the first line of equation (3) imply  $D(t) (\eta(t) - \gamma) = -\theta e^{-\lambda t}$  so equation (27) simplifies to

$$K(S_t) = -\theta \int_0^\infty e^{-\lambda \tau} u(S_{t+\tau}^*, \chi(S_{t+\tau}^*)) d\tau. \quad (28)$$

The integral in equation (28) is the present discounted value of the equilibrium future flow of payoff, computed using the discount rate  $\lambda$ . Thus,  $-K(S_t)$  is an annuity, which if received in perpetuity and discounted at  $\theta$  (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 payoff in the original model, plus this annuity. A necessary condition for the MPE is that

$$x_t = \chi(S_t) \equiv \arg \max (u(S_t, x_t) - K(S_t) + J_S(S) F(S, x)), \quad (29)$$

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

$$\gamma J(S) = (u(S, \chi(S)) - K(S) + J_S(S) F(S, \chi(S))). \quad (30)$$

---

<sup>12</sup>Appendix B.2 explains why these necessary conditions, together with the definition in equation (13), are also sufficient for a MPE.

For agents with deterministic lifetime  $T$ , the function  $K$  has a slightly different form than above. Using equation (4) of Karp (2007) and equations (7) and (2),  $-K$  equals<sup>13</sup>

$$-K(S_t) \equiv (\lambda - r) \int_0^T \frac{(T - \tau)}{T} e^{-r\tau} U(S_{t+\tau}^*, \chi(S_{t+\tau}^*)) d\tau. \quad (31)$$

The integral on the right side of definition (31) is the present value, discounted at the selfish rate  $r$ , of the payoff that those alive at  $t$  receive from the flow  $U(S_{t+\tau}^*, \chi(S_{t+\tau}^*))$  over  $[t, T + t]$ . Over that interval, the number of agents remaining from the time  $t$  population decreases linearly. Again, we can interpret  $-K$  as an annuity, which if received in perpetuity and discounted at the rate  $(\lambda - r)$ , equals the present value to those currently alive of the program  $U(S_{t+\tau}^*, \chi(S_{t+\tau}^*))$ .

The dynamic programming equation in this case is

$$\lambda J(S) = \max_x (U(S_t, x_t) - K(S_t) + J_S(S) f(S, x)) \quad (32)$$

with the annuity  $K$  given by equation (31).

## B.2 Sufficiency

The discussion of sufficiency in Karp (2007) is misleading, and I take this opportunity to clarify it. The endogeneity of the function  $K(S)$ , and the resulting difficulty in determining its curvature, renders inapplicable the standard sufficiency conditions for the fictitious control problem, defined by equations (26) and (27). However, the sufficiency regarding the fictitious control problem is a red herring, because that problem is merely a device for describing the equilibrium to the sequential game induced by non-constant discounting; for that purpose, we use only the necessary conditions to the fictitious problem. The maximization problem in equation (21) of Karp (2007) is merely a statement of the problem for the planner in a particular period in a discrete time setting, under the assumption of Markov perfection. Equation (5) of that paper (equivalently, equation (30) above) is merely the limiting form of the discrete time condition, as the length of a period of commitment goes to

---

<sup>13</sup>Karp (2007) sets up the problem using a discount rate  $r(t)$  for  $t \leq T$  and  $r(t) = \bar{r}$  for  $t \geq T$ . The results for the case of exponentially distributed lifetime use the limiting case as  $T \rightarrow \infty$ . In the OLG model with finitely lived agents,  $T$  is finite and  $\bar{r} = \lambda$ .

0. Therefore, provided that we are willing to restrict attention to the limiting game (as the length of a period goes to zero in the discrete time game), and provided that the value function is differentiable, a sufficient condition for the MPE is that the control rule satisfy equation (29) above, and that the value function satisfy the DPE (30).

The primitive functions of some interesting optimal control problems do not have the curvature need to satisfy familiar sufficient conditions. Sufficiency in optimal control problems is therefore sometimes a difficult issue, and the analysis sometimes proceeds without reference to sufficiency. The difficulty arises because sufficiency is a global property in optimal control problems. In contrast, sufficiency is a much simpler issue in the type of sequential game induced by non-constant discounting and the requirement of Markov perfection. In this game, each of the succession of social planers chooses a single action; given her beliefs about successors' policy function, each policy maker thus solves a static optimization problem. Because each of the policymakers treats the functions  $J(S)$  and  $K(S)$  as predetermined (although they are endogenous to the game), sufficiency requires (in the limit as  $\varepsilon \rightarrow 0$ ) only that  $x = \chi(S)$  maximize  $(U(S_t, x_t) + J_S(S) f(S, x))$ .

### B.3 Proof of Lemma 1

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

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

Note that  $\tau < \infty$ . Use

$$\begin{aligned} & \lim_{\lambda \rightarrow 0} \left( \frac{\int_0^\infty D(t)U(t)dt}{\frac{U_\infty}{\lambda}} \right) \\ &= \lim_{\lambda \rightarrow 0} \frac{\lambda \left[ \left( \int_0^\tau D(t)U(t)dt \right) + \left( U_\infty \int_\tau^\infty D(t)dt \right) + \left( \int_\tau^\infty U(t)D(t)dt - U_\infty \int_\tau^\infty D(t)dt \right) \right]}{U_\infty}. \end{aligned} \tag{33}$$

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

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

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  $\lambda \rightarrow 0$  of this expression, implies that the second term on the second line of equation (33) equals  $\frac{\theta}{\gamma}$ . By definition of  $\tau$ ,

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

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

## B.4 Proof of Proposition 2

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 optimal static steady state for  $\lambda$  close to 0. Because I am interested in the case where  $\lambda$  is small, I assume throughout that  $\lambda < r$ .

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

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

where the subscript  $\infty$  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

$$S_{t+\tau} = e^{z\tau} S_t + S_{\infty} (1 - e^{z\tau}) + o(S_t - S_{\infty}) \implies \frac{dS_{t+\tau}}{dS_t} \approx e^{z\tau} \quad (34)$$

for  $\tau \geq 0$ . Equation (13) implies

$$K'(S_t) = (r - \lambda) \int_0^{\infty} e^{-\gamma\tau} (u_S(S_{t+\tau}, \chi(S_{t+\tau})) + u_x(S_{t+\tau}, \chi(S_{t+\tau})) \chi'(S_{t+\tau})) \frac{dS_{t+\tau}}{dS_t} d\tau. \quad (35)$$

Using equation (34) and evaluating equation (35) at  $S_t = S_\infty$  gives

$$\begin{aligned} K'(S_\infty) &= (r - \lambda) (u_S + u_x \chi')|_\infty \int_0^\infty e^{-\gamma t} e^{zt} dt \\ &= \frac{(r - \lambda)(u_S + u_x \chi')|_\infty}{\gamma - z} = \frac{(r - \lambda)(u_S + u_x \frac{\Psi'}{n})|_\infty}{\gamma - z}. \end{aligned} \quad (36)$$

The Hamiltonian corresponding to the fictitious optimal control problem in equation (12) is

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

where  $\mu$  is the current value costate variable. The necessary conditions for optimality are

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

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

$$\begin{aligned} &\left[ -u_S + K' + \frac{u_x}{f_x} (f_S + \frac{n-1}{n} f_x \Psi' - \lambda) \right]_{|\infty} = \\ &\left[ -u_S + \frac{(r - \lambda)(u_S + u_x \frac{\Psi'}{n})}{\gamma - z} + \frac{u_x}{f_x} (f_S + \frac{n-1}{n} f_x \Psi' - \lambda) \right]_{|\infty} = 0, \end{aligned} \quad (37)$$

where the first equality uses equation (36). Using the definition of  $z$  and rearranging the second line of equation (37) implies that  $\Psi' = \Psi'(S_\infty)$  is a solution to the quadratic equation

$$Q \times (\Psi')^2 + L \times \Psi' + C = 0 \quad (38)$$

with

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

Hereafter I set  $n = 1$ , so

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

$$z = f_S + f_x \frac{(\lambda + \theta - f_S) \frac{u_S}{u_x} - \frac{1}{f_x} (f_S - \lambda)(\gamma - f_S)}{r - f_S + \frac{u_S}{u_x} f_x} = \frac{\theta \left( \frac{u_S}{u_x} - \frac{f_S}{f_x} \right) f_x + \lambda \left( \frac{u_S}{u_x} - \frac{f_S}{f_x} + \frac{\gamma}{f_x} \right) f_x}{r + \left( \frac{u_S}{u_x} - \frac{f_S}{f_x} \right) f_x}. \quad (39)$$

The optimal static steady state 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 optimal static steady state), a stock level slightly greater than the optimal static steady state satisfies

$$\left( \frac{u_S}{u_x} - \frac{f_S}{f_x} \right) u_x = \varepsilon < 0 \quad \text{or} \quad \left( \frac{u_S}{u_x} - \frac{f_S}{f_x} \right) = \frac{\varepsilon}{u_x} > 0, \quad (40)$$

for  $\varepsilon$  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 (40) in (39) 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  $\varepsilon$ . For  $\varepsilon$  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  $\lambda$  is sufficiently small, as was to be shown.

## C The linear-quadratic model with $n$ tribes

This appendix is not intended for publication. In the interest of generality, I provide the formulae for the model that includes an interaction between the control and state variables in the utility function. The climate model does not require this interaction. I first make a linear transformation of the state and control variables in order to reduce the dimension of parameter space in the linear-quadratic one-tribe model. I then convert the  $n = 1$  model to the multi-tribe setting. The next subsections provide formulae for the linear equilibria for general  $n$ , under both exponentially distributed lifetimes and deterministic lifetimes. I also explain how to obtain nonlinear equilibria for the model with exponentially distributed lifetime. I then discuss the calibration.

The text notes that a specialization of the linear-quadratic problem has the flow payoff quadratic in the control and linear in the state. It is easy to see why there are two globally defined MPE decision rules for this variation, a constant and a linear decision rule. The constant equilibrium decision can be obtained in closed form; to obtain the linear equilibrium it is necessary to solve a quadratic rather than a cubic equation. Nonlinear equilibria must be solved numerically. If agent  $it$  believes that other agents set their decision to a constant level, agent  $it$  faces a problem that is linear in the state; the equilibrium decision to that problem is a constant. If agent  $it$  believes that other agents use a linear control rule, the function  $K$  is quadratic in the state; agent  $it$  therefore has a fictitious control problem that is quadratic in the state; the solution to that problem is a linear control rule.

### C.1 Reduction in parameter space

Here I make a linear transformation of the state and control variables that reduces the dimension of parameter space from 8 to 4. Begin with a one-tribe model in which the state variable is  $\sigma$  and the control variable is  $\varphi$ . Given the 8 parameters  $w, W, v, V, M, G, d, C$ , the flow payoff in the one-tribe setting is

$$-(w\varphi + W\varphi^2 + v\sigma + V\sigma^2 + M\varphi\sigma) \quad (41)$$

and the equation of motion is

$$\dot{\sigma} = G + d\sigma + C\varphi. \quad (42)$$

Define a new state and control variable:

$$\begin{aligned} S &= \sqrt{2V} \left( \frac{2vW - Mw}{4VW - M^2} + \sigma \right) \\ X &= \sqrt{2W} \left( \frac{2wV - Mv}{4VW - M^2} + \varphi \right) \end{aligned} \quad (43)$$

With these definitions, the flow payoff and the equation of motion are, respectively,

$$-\frac{1}{2} (X^2 + S^2 + mXS) + \text{constant}, \text{ and } \dot{S} = g + dS + cX \quad (44)$$

with  $m \equiv \frac{M}{\sqrt{VW}}$  and

$$g \equiv \sqrt{2V} \left( G - D \frac{2vW - Mw}{4VW - M^2} + C \frac{Mv - 2wV}{4VW - M^2} \right) \text{ and } c \equiv \frac{C\sqrt{V}}{\sqrt{W}}. \quad (45)$$

The constant in the flow payoff does not affect behavior, so I ignore it henceforth.

## C.2 From one tribe to $n$ tribes

I want to define a flow payoff and an equation of motion (the technology) such that the aggregate feasible payoff does not depend directly on  $n$ . The equilibrium aggregate payoff and decision rule depends on  $n$  only insofar as  $n$  alters the equilibrium decision rules of the individual tribes. That is,  $n$  has a strategic but not an intrinsic effect on agents' payoffs.

Define the flow payoff and the constraint facing the  $i$ 'th tribe as

$$-\frac{1}{2} \left( \frac{1}{n} S^2 + n x_i^2 + m x_i S \right) \quad \dot{S} = g + dS + c \left( x_i + \sum_{j \neq i} x_j \right), \quad (46)$$

where  $x_i$  is the control variable of the  $i$ 'th tribe at an arbitrary point in time. If all tribes use the same decision,  $x$ , the aggregate action is  $X = \sum_i x_i = nx$ . The aggregate payoff when all tribes use the same action is  $n$  times the first expression in system (46), which equals the first expression in system (44). When all tribes use the same action, the equations of motion in the two systems are obviously identical. The equilibrium flow payoff (where  $x = \frac{X}{n}$ ) simplifies to  $-\frac{1}{2n} (S^2 + X^2 + mXS)$

### C.3 Calculating the equilibria

This section provides details for the linear and non-linear equilibria in the model where agents have exponentially distributed lifetime, and for the linear equilibrium where agents have known finite lifetime.

#### C.3.1 Exponentially distributed lifetime

For the purpose of obtaining the linear equilibrium under exponentially distributed lifetimes for the two cases,  $\lambda < r$  and  $\lambda > r$ , and for the limiting case  $\lambda = \infty$ , I introduce constants  $\eta, \epsilon, \nu$  that take the values given in Table 3

	$\eta =$	$\epsilon =$	$\nu =$
if $\lambda > r$	$\lambda$	$\theta$	$\gamma$
if $\lambda < r$	$\gamma$	$\lambda - r$	$\lambda$
if $\lambda = \infty$	n/a	$\epsilon = 0$	$\gamma$

Table 3: values of  $\eta, \epsilon, \nu$  for different cases

Here,  $\eta$  is the discount rate used in the definition of  $K$ ,  $\epsilon$  is the factor that multiplies the integral in the definition of  $-K$ , and  $\nu$  is the discount rate used in the fictitious control problem.

The Hamiltonian,  $H$ , for the fictitious optimal control problem, the necessary conditions for that problem, and the definition of  $K$ , are

$$\begin{aligned}
 H &= -\frac{1}{2} \left( \frac{1}{n} S^2 + n x^2 \right) - K(S) + \mu \left( g + dS + c \left( x + \frac{n-1}{n} \Psi(S) \right) \right) \\
 \frac{\partial H}{\partial x} &= -n x + \mu c = 0 \implies \mu = \frac{n x}{c} = \frac{X}{c} \\
 \dot{\mu} &= \nu \mu - \frac{\partial H}{\partial S} = \nu \mu - \left( -\frac{S}{n} - K' + \mu \left( d + \frac{n-1}{n} \Psi'(c) \right) \right) \\
 K &= -\epsilon \int_0^\infty e^{-\eta t} \left( -\frac{1}{2n} \right) (S^2 + X^2) dt.
 \end{aligned} \tag{47}$$

In a symmetric equilibrium,  $n x = \Psi \equiv X$ .

Differentiating the second equation in system (47) with respect to time and using the third equation gives

$$\frac{\dot{X}}{c} = \nu \frac{X}{c} - \left( -\frac{S}{n} - K' + \frac{X}{c} \left( d + \frac{n-1}{n} c X' \right) \right).$$

Dividing this equation by  $\dot{S}$ , using  $\frac{\dot{X}}{\dot{S}} = \frac{dX}{dS} = X'$ , gives

$$\begin{aligned} \frac{\dot{X}}{c\dot{S}} = \frac{X'}{c} &= \frac{\nu\frac{X}{c} - \left(-\frac{S}{n} - K' + \frac{X}{c}\left(d + \frac{n-1}{n}cX'\right)\right)}{g+dS+cX} \implies \\ X' &= \frac{(Xn(\nu-d)+Sc+cXK')}{n(g+dS)-Xc(1-2n)}. \end{aligned} \quad (48)$$

The fourth equation in system (47) implies

$$\begin{aligned} \eta K &= \frac{\epsilon}{2n}(S^2 + X^2) + K'(g + dS + cX) \implies \\ K' &= \frac{\eta K - \frac{\epsilon}{2n}(S^2 + X^2)}{(g+dS+cX)}. \end{aligned} \quad (49)$$

A MPE must solve the ODEs in the second lines of equations (48) and (49).

**Finding the linear equilibrium** In the linear equilibrium,  $X$  is a linear and  $K$  is a quadratic function of  $S$ :

$$\begin{aligned} X &= a + \Delta S \\ K &= \kappa_0 + \kappa_1 S + \frac{1}{2}\kappa_2 S^2. \end{aligned}$$

The objective is to find the equilibrium values of  $a$ ,  $\Delta$ . Substituting these two functions in the ODEs in equations (48) and (49) and equating coefficients in  $S$  and 1 for the first equation, and coefficients of  $S^2$  and  $S$  in the second equation, implies

$$\begin{aligned} (c - 2cn)\Delta^2 + n(\nu - 2d)\Delta + c + cn\kappa_2 &= 0 \\ (2dn - n\eta + 2cn\Delta)\kappa_2 + (\epsilon\Delta^2 + \epsilon) &= 0 \\ (c\Delta + n\nu - dn - 2cn\Delta)a + cn\kappa_1 - gn\Delta &= 0 \\ (2dn - 2n\eta + 2cn\Delta)\kappa_1 + (2gn\kappa_2 + 2a\Delta\epsilon + 2acn\kappa_2) &= 0 \end{aligned} \quad (50)$$

Solving the second equation for  $\kappa_2$  and substituting the result into the first equation implies

$$\kappa_2 = -\frac{\epsilon\Delta^2 + \epsilon}{2dn - n\eta + 2cn\Delta}$$

$$\begin{aligned} \Sigma &\equiv 2c^2(2n-1)\Delta^3 + (((8d-2\eta-2\nu)n + (\epsilon + \eta - 2d))c)\Delta^2 \\ &+ ((\nu\eta - 2d\eta - 2d\nu + 4d^2)n - 2c^2)\Delta + c(\epsilon + \eta - 2d) = 0. \end{aligned}$$

The correct root of the cubic  $\Sigma = 0$  must yield a stable equilibrium, i.e. the inequality  $d + c\Delta < 0$  must hold. Use this root (or roots) to obtain the equilibrium value(s) of  $\kappa_2$ . Using these values, the last two equations in the system (50) are a pair of linear equations in  $a$ ,  $\kappa_1$ .

**Finding nonlinear equilibria** I begin by using the argument presented in Appendix B.4, generalized to allow arbitrary values of  $\lambda$ , but specialized to the linear-quadratic setting. Denote  $\Phi$  as the value of  $X'$  at a steady state, i.e. where  $0 = g + dS + cX$ . With this definition,  $z = d + c\Phi$ . A derivation that parallels that which establishes equation (36) gives

$$K'(S_\infty) = \frac{\epsilon}{n} (S + X\Phi)|_\infty \int_0^\infty e^{-\eta t} e^{zt} dt = \frac{\frac{\epsilon}{n}(S+X\Phi)|_\infty}{\eta-z} = \frac{\frac{\epsilon}{n}(S-\frac{1}{c}(g+Sd)\Phi)}{\eta-d-c\Phi}. \quad (51)$$

Substitute this steady state value of  $K'$  into the steady state value of  $X'$ , using equation (48), and evaluate the result at a steady state ( $X = -\frac{1}{c}(g + Sd)$ ) to obtain the quadratic in  $\Phi$ :

$$\begin{aligned} Q\Phi^2 + L\Phi + C &= 0 \quad \text{with} \\ Q &= (1-n)c(g+Sd) \\ L &= (d-\eta-\epsilon+n\nu+2n\eta-3dn)(g+Sd) + gn(d-\eta) + S(d^2n-dn\eta-c^2) \\ C &= Sc(\epsilon+\eta-d) + \frac{n}{c}(g+Sd)(d\eta-\nu\eta+d\nu-d^2). \end{aligned} \quad (52)$$

The set of  $S$  that can be supported as steady states in a MPE is

$$\Lambda = \{S \mid \exists \Phi \text{ for which } d + c\Phi < 0 \wedge Q\Phi^2 + L\Phi + C = 0\}.$$

To construct a non-linear MPE for this model, pick a value of  $S \in \Lambda$  with  $X = -\frac{1}{c}(g + Sd)$  and solve the pair of ODEs in equations (48) and (49) with these initial conditions. In general, these ODEs must be solved numerically. A possible difficulty arises from the fact that we do not know the domain of such a solution. For example, if we were to solve these ODEs using a function approximation over a domain that is strictly larger than the domain of existence of the ODEs, then the solver might return a poor approximation of the true solution.

MuPad can solve the initial value problem described above; MuPAD's default method is DOPRI78, an embedded Runge-Kutta pair of orders 7 and 8. This procedure runs into difficulty because the expression for  $K'$  (and also for  $X'$  in the case where  $n = 1$ ) is an indeterminate form at a steady state. In order to avoid this difficulty, choose a small  $\epsilon$  and replace the initial condition  $(S_\infty, X_\infty, K_\infty)$  with a nearby point  $(S_\infty + \epsilon, X_\infty + \Phi\epsilon, K_\infty + K'(S_\infty)\epsilon)$ . This point lies close to the correct boundary, and it lies on the tangent to

the trajectory that takes the system to this boundary. The solution to this initial value problem,  $X(S)$ , can be plotted, to confirm that it does not cross the line  $n(g + dS) - Xc(1 - 2n)$ , where  $X'$  is undefined. (As a consistency check, I confirmed that this approach returns the linear equilibrium when given an initial condition near the steady state to the linear equilibrium.) The results reported in the text use this method of numerically obtaining non-linear equilibria.

In order to calculate the function  $\beta$  graphed in Figure 2 I set  $d + c\Phi = \tau$ , where  $\tau$  is a non-positive number close to 0. For  $\tau < 0$ , the value of  $\Phi$  that solves this equation is associated with a stable steady state;  $\tau = 0$  gives the infimum of values of  $\Phi$  associated with a stable steady state. Substituting  $\Phi = \frac{-d}{c}$  into  $Q\Phi^2 + L\Phi + C = 0$  results in an equation that is linear in  $S$ . The value of  $S$  that solves this equation equals the infimum of  $S$  that can be supported as a steady state in a stable MPE. Replacing  $S$  with  $-\beta \frac{dg}{d^2+c^2}$  expresses the infimum of stable steady states as a multiple  $\beta$  of the optimal static steady state,  $-\frac{dg}{d^2+c^2}$ . The resulting equation is linear in  $\beta$  and depends on all other parameter values. Setting those values equal to the levels given in Section C.4 results in a  $\beta$  as a function of  $n, \lambda$  and  $\Omega$ , shown in Figure 2.

### C.3.2 Deterministic lifetime

The simplicity of the model with exponentially distributed lifetime results from the fact that for that case it is trivial to obtain the ODE for the function  $K(S)$ . It is not clear how one would obtain the ODE for  $K$  in the case where agents have known finite lifetime. I therefore use a slightly different argument to obtain the linear equilibrium in this case, and I leave the non-linear case for future work.

The linear equilibrium  $X = a + \Delta S$  must solve the ODE in the second line of equation (48), a fact that leads to the first and the third line of system (50). Under the linear policy  $X = a + \Delta S$ , and given a stock  $S(0) = s$ ,

$$S(\tau) = \frac{-(g + ca)}{d + c\Delta} (1 - e^{(d+c\Delta)\tau}) + e^{(d+c\Delta)\tau} s. \quad (53)$$

Substituting this formula into the flow payoff  $U = -\frac{1}{2n} (S^2 + (a + \Delta S)^2)$

and collecting terms,

$$\begin{aligned}
U &= \sigma_2 s^2 + \sigma_1 s + \sigma_0 \quad \text{with} \\
\sigma_2 &= -\frac{1}{2n} \left( e^{2((d+c\Delta)\tau)} + \Delta^2 e^{2((d+c\Delta)\tau)} \right) \\
\sigma_{11} &= -\frac{1}{2n} \left( -2\frac{c}{d+c\Delta} \left( 1 - e^{(d+c\Delta)\tau} \right) e^{(d+c\Delta)\tau} + 2 \left( 1 - c\frac{\Delta}{d+c\Delta} \left( 1 - e^{(d+c\Delta)\tau} \right) \right) \Delta e^{(d+c\Delta)\tau} \right) \\
\sigma_{12} &= -\frac{1}{2n} \left( -2\frac{g}{d+c\Delta} \left( 1 - e^{(d+c\Delta)\tau} \right) e^{(d+c\Delta)\tau} - 2\Delta^2 \frac{g}{d+c\Delta} \left( 1 - e^{(d+c\Delta)\tau} \right) e^{(d+c\Delta)\tau} \right) \\
\sigma_1 &= a\sigma_{11} + \sigma_{12} \\
\sigma_0 &= -\frac{1}{2n} \left( \frac{(-g-ca)^2}{(d+c\Delta)^2} \left( 1 - e^{(d+c\Delta)\tau} \right)^2 + \left( a + \Delta \frac{-g-ca}{d+c\Delta} \left( 1 - e^{(d+c\Delta)\tau} \right) \right)^2 \right)
\end{aligned} \tag{54}$$

Under the linear policy,  $K(s)$  is a quadratic function:  $K(s) = \kappa_2 s^2 + \kappa_1 s + \kappa_0$ . This fact and the definition of  $K$  imply

$$\kappa_2 s^2 + \kappa_1 s + \kappa_0 = (r - \lambda) \int_0^T \frac{(T - \tau)}{T} \left( e^{-r\tau} \sigma_2 s^2 + \sigma_1 s + \sigma_0 \right) d\tau.$$

Equating coefficients gives

$$\begin{aligned}
\kappa_2 &= (r - \lambda) \int_0^T \frac{(T - \tau)}{T} e^{-r\tau} \sigma_2 d\tau \\
\kappa_1 &= (r - \lambda) \int_0^T \frac{(T - \tau)}{T} e^{-r\tau} \sigma_1 d\tau
\end{aligned} \tag{55}$$

The first line of system (50), together with the definitions of  $\kappa_2$  and  $\sigma_2$  is a non-linear equation in  $\Delta$ . A solution that satisfies  $d + c\Delta < 0$  is a stable root. Given a numerical value for  $\Delta$ , the third line of system (50), together with the definitions of  $\kappa_1$  and  $\sigma_1$  imply

$$a = -n \frac{c\sigma_{14} - g\Delta}{cn\sigma_{13} + c\Delta + n\lambda - dn - 2cn\Delta}.$$

## C.4 Calibration of the climate model

I first express the model in “natural units” and then rewrite the model so that the control and state variables are percentages that have a convenient interpretation. I then use the transformation in Appendix C.1 to reduce the dimension of parameter space.

Let  $Y_t$  be the stock variable at time  $t$ , ppm of carbon and  $x$  be the flow, measured in ppm per year. (One ppm by volume equals 2.13 GtC.) The equation of motion for the stock of greenhouse gases is

$$\dot{Y} = \bar{g} + \delta Y + x. \tag{56}$$

The parameters  $\bar{g}$ ,  $\delta$ , and 1 on the right side correspond to  $G$ ,  $d$ ,  $C$  in equation (42);  $Y$  corresponds to  $\sigma$  and  $x$  corresponds to  $\varphi$ . To calibrate the model, I set the half life of carbon to 83 years, the steady state in the absence of anthropogenic emissions equal to the pre-industrial level to 280 ppm, and assume that under BAU the stock increases from the current level, 380, to 700 in 90 years. These assumptions imply

$$\delta = -8.3511709 \times 10^{-3}, \quad \bar{g} = 2.3383278, \quad x^{\text{bau}} = 5.8926860. \quad (57)$$

With these parameters, the steady state stock under BAU is 986 ppm, and after 200 years of BAU the stock reaches 872 ppm.

For ease of interpretation, it is convenient to express the control variable,  $A$ , as abatement as a percent of BAU emissions, and the stock,  $s$ , as the percent increase over preindustrial levels:

$$A = \frac{x^{\text{bau}} - x}{x^{\text{bau}}} 100 \quad \text{and} \quad s = \frac{Y - 280}{280} 100$$

With these definitions and the parameter values in equation (57), the equation of motion is

$$\dot{s} = \delta s - \frac{1}{280} x^{\text{BAU}} A + \left( \frac{5}{14} x^{\text{BAU}} + 100\delta + \frac{5}{14} \bar{g} \right) = ds + CA + G \quad (58)$$

with

$$G = 2.1045307 \quad d = \delta = -8.3511709 \times 10^{-3} \quad C = -2.1045307 \times 10^{-2}. \quad (59)$$

In this stationary setting, denote the constant  $Z$  as gross world product (GWP) exclusive of climate related damage and abatement costs. In this linear-quadratic model, the flow cost of abatement, as a percent of GWP, and the flow cost of the stock, as a percent of GWP are, respectively,

$$\frac{b}{2} \frac{A^2}{Z} 100 \quad \text{and} \quad \frac{h}{2} \frac{s^2}{Z} 100.$$

For calibration, suppose that the flow cost of a 50% reduction in emissions relative to BAU ( $A = 50$ ), is  $Q$  percent of GWP; and the flow cost of doubling of ppm relative to pre-industrial level ( $s = 100$ ), is  $P$  percent of GWP. These assumption imply

$$b = \frac{1}{125000} ZQ, \quad h = \frac{1}{500000} PZ \quad \implies \quad \frac{h}{b} = \frac{P}{4Q}.$$

Define  $\Omega = \frac{P}{Q}$ , so that the total flow costs, as a percent of GWP, corresponding to actual abatement  $A$  and actual stock  $s$  equal

$$\left(\frac{b100}{2Z}\right) \left((A)^2 + \frac{\Omega}{4}(s)^2\right).$$

Hereafter I drop the positive factor  $\frac{b100}{2Z}$  to write flow benefits (negative costs) as

$$-\left(A^2 + \frac{\Omega}{4}s^2\right).$$

Table 4 shows the correspondence between the variables and parameters in the general model in Section C.1 and in the climate model here.

	control	state								
general	$\varphi$	$\sigma$	$w$	$W$	$v$	$V$	$M$	$G$	$d$	$C$
climate	$A$	$s$	0	1	0	$\frac{\Omega}{4}$	0	$G$	$d$	$C$

Table 4: Correspondence between parameter values in general model from Section C.1 and the climate model in this section

Equation (59) gives the numerical values of  $G, d, C$  that appear in the last row of Table 4, for this calibration. Use Table 4, the numerical values in equation (59) and the formulae in equation (45), to obtain the values of the model parameters that are used in finding the equilibria in the climate model:

$$g = 1.4881279\sqrt{\Omega} \quad \text{and} \quad c = -1.0522654 \times 10^{-2}\sqrt{\Omega} \quad (60)$$

In interpreting the equilibrium results, it is important to keep in mind that the model is solved in terms of the transformed state and control  $S$  and  $X$ . The percent increase in the stock relative to preindustrial level,  $s$ , and aggregate abatement as a percent of BAU emissions,  $A$ , are related to  $S, X$  using system (43) and the correspondences in Table 4:

$$s = \sqrt{\frac{2}{\Omega}}S, \quad \text{and} \quad A = \frac{X}{\sqrt{2}}.$$

The initial condition for the problem is  $s = \frac{380-280}{280}100 = 35.714286$  or  $S = 35.714286\sqrt{\frac{\Omega}{2}}$ . The text describes the results in terms of the stock of atmospheric carbon expressed in ppm.