

Discounting Utility and the Evaluation of Climate Policy

Larry Karp

1 Introduction

This essay considers three aspects of the relation between utility discounting and climate policy: the sensitivity of policy recommendations to discounting assumptions, the relation between discounting and catastrophic risk, and the difference between discounting for intra- and inter-personal intertemporal transfers.

The utility discount rate converts future and current utility into the same units, thereby making inter-temporal comparisons sensible.¹ Much of our intuition about the sensitivity of climate policy to discounting comes from cost-benefit examples. In this context, we are willing to spend very little to influence a non-catastrophic event that occurs in the distant future; the amount that we are willing to spend may be sensitive to the discount rate. These conclusions might be reversed if the date of the policy outcome is random instead of deterministic (Sect. 2). Analytic examples and a review of numerical climate models show varying levels of sensitivity of optimal policy to discounting assumptions. The complexity of these models makes an explanation for these differences unattainable. But it may be interesting to note that

¹The utility discount rate (the pure rate of time preference, or PRTP) is related to, but distinct from, the social discount rate, used to compare consumption (as distinct from utility) across two points in time. In the standard deterministic setting, the social discount rate equals the PRTP plus the growth rate multiplied by the inverse of the elasticity of intertemporal substitution. The latter provides a measure of the willingness to substitute consumption between “infinitely close” periods.

I thank an anonymous reviewer for comments on an earlier draft of this paper. The usual disclaimer applies. Sects. 3.2 and 4 of this paper are adapted from Karp (2009), and Sect. 5 is based on Karp (2013a, b).

L. Karp (✉)
Department of Agricultural and Resource Economics,
University of California, 207 Giannini Hall, Berkeley, CA 94720, USA
e-mail: karp@are.berkeley.edu

when the underlying model is “more linear”, the solution appears more sensitive to discounting. The definition of “more linear” is context-specific (Sect. 3).

Time-discounting and catastrophe-avoidance are logically distinct topics, but recent papers claim that the risk of catastrophe swamps any consideration of discounting. Sect. 4 explains why I am skeptical of this claim.

Most policy models use an infinitely lived agent model. That model makes no distinction between intertemporal transfers involving the same agent, e.g. a person when they are young and old, and transfers involving two different people. These two types of transfers are quite different. Even if we use a constant discount rate to evaluate each, there is no reason that we should use the same constant for the two types of transfers. A two-parameter discounting model can distinguish between these two types of transfers. A planner who gives equal weight to the welfare of all people currently living, but distinguishes between intertemporal transfers for the same person and intertemporal transfers between different people, has time inconsistent preferences. Sect. 5 illustrates the discount rate induced by such a planner in an overlapping generations setting.

2 The Cost-Benefit Setting

Although society cannot literally insure itself against climate-related events, current expenditures can reduce the probability of those events (abatement), and the cost associated with them (mitigation). I consider the extreme case where society has a binary choice, either to do nothing and face the risk (or the certainty) of the event, or to take a costly action that eliminates the risk. I refer to the action as buying insurance, and the cost of the action as the cost of the premium. The formal question is to determine how the pure rate of time preference (PRTF) affects society’s maximum willingness to pay for “perfect” insurance, which eliminates the risk of climate change (Karp 2009).

Society’s actual policy choice is not binary, but instead requires choosing among many different types and levels of control; and the actions we take might reduce, but cannot eliminate the risk or the consequences of climate change. However, the simplicity of the model makes it easy to see how key parameters, in particular the PRTF, affect society’s willingness to incur current costs to ameliorate future damages. Abstracting from the complications of more realistic policy-driven models throws into relief certain relations between parameter assumptions and model recommendations.

I compare society’s willingness to pay to eliminate the risk in two extreme cases: where the time of the event, T , is known with certainty, and where the event time, \tilde{T} , is a random variable. I set the expected time, $E(\tilde{T})$, in the stochastic case, equal to the known time in the deterministic case, T , so that the two models are comparable. The two noteworthy qualitative results are the same for both zero and positive consumption growth. First, moving from a certain event time to a random event time increases the maximum premium that society is willing to pay, especially for low

probability events. Second, the premium is less sensitive to the PRTP in the stochastic case, compared to the deterministic case.

The intuition for these results, based on Jensen’s inequality, is quite simple. In moving from the deterministic setting, which concentrates all of the probability mass at the expected event time, to the stochastic case, we transfer some of that concentrated probability mass to earlier times, and some to later times. The higher probability of an earlier event increases the expected present value cost of the event, and the higher probability of a later event decreases that expected present value. However, because of discounting, the effect of the first change is larger than the effect of the second. Therefore, moving from a deterministic to a stochastic event time increases the premium that society is willing to pay.

The explanation for the greater sensitivity (to the PRTP) of the maximum premium in the deterministic case is only slightly more involved. The elasticity of the maximum premium, with respect to the PRTP, in the stochastic case, is a *weighted* expectation of the elasticity under certainty. This formula assigns higher weight to the elasticities corresponding to earlier event times; those elasticities are lower than their counterparts at later times.

Table 1 collects the parameter definitions used in this section.

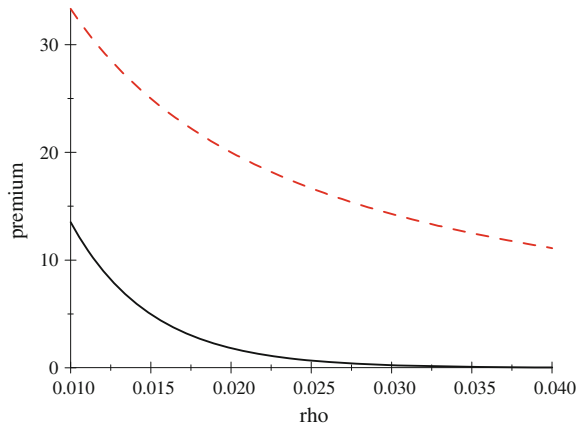
2.1 Zero Growth

The payment of a premium and the loss associated with the event both reduce consumption, and therefore reduce utility. With zero growth, these are the only factors that affect consumption and utility. I express the costs associated with the event and with payment of the premium in units of utility.

Table 1 Parameter definitions

Parameter	Definition
ρ	Pure rate of time preference
$h = \frac{1}{T}$	Hazard (inverse of expected event time)
Δ	Fraction of income lost due to event
g	Consumption growth rate
η	Elasticity of marginal utility: $-\frac{u''(c)}{u'(c)}c$
r	Social discount rate
x	Maximum willingness to pay to avoid event (the premium)
ϕ	Elasticity of premium wrt ρ : $-\frac{dx}{d\rho} \frac{\rho}{x}$

Fig. 1 Maximum premium as a percent of the flow loss when event time is $T = 200$ (solid) and when event time is exponentially distributed with $E(T) = 200$ (dashed)



2.1.1 Deterministic Event Time

For the deterministic case, suppose that by paying a premium that has a utility flow cost of x per unit of time, society insures itself against (and in that sense avoids) a utility flow loss of 100 in each period from time T to ∞ . With a PRTP ρ , the present value of the utility loss that begins at T is $e^{-\rho T} \frac{100}{\rho}$, and the present value of the premium payment, beginning today, is $\frac{x}{\rho}$. Equating these expressions implies that society would be willing to pay a premium of at most $x(T, \rho) = e^{-\rho T} 100$ over $(0, \infty)$. In this deterministic setting, the premium, x , is a convex function of T ; this fact is key to understanding the effect of moving to an uncertain event time. If $T = 200$, the premium, x , changes by a factor of 55, ranging from 13.5 to 0.25, as ρ ranges from 0.01 to 0.03 (1 to 3% per annum); see Fig. 1. The elasticity of x with respect to ρ is $\phi(x) = \rho T$. Thus, x is particularly sensitive to the PRTP when the event occurs in the distant future. This example illustrates the role of discounting in the simplest cost-benefit calculation.

2.1.2 Stochastic Event Time

I begin with a two point distribution to provide intuition, and then move to the exponentially distributed event time, which yields a simpler formula for the premium and its elasticity with respect to ρ . With the two-point distribution, \tilde{T} takes two values, $T - \epsilon$ and $T + \epsilon$, each with probability 0.5, so $E(\tilde{T}) = T$. The maximum premium for the risk neutral planner, x' , is the expectation over \tilde{T} of $x(\tilde{T}, \rho)$:

$$\begin{aligned}
 x' &= E_{\tilde{T}}x = (e^{-\rho(T-\varepsilon)} + e^{-\rho(T+\varepsilon)}) 50 \Rightarrow \\
 \phi(x') &= -\frac{dx'}{d\rho} \frac{\rho}{x'} = \rho \left((T - \varepsilon) \frac{e^{-\rho(T-\varepsilon)}}{e^{-\rho(T-\varepsilon)} + e^{-\rho(T+\varepsilon)}} + (T + \varepsilon) \frac{e^{-\rho(T+\varepsilon)}}{e^{-\rho(T-\varepsilon)} + e^{-\rho(T+\varepsilon)}} \right) \\
 &= \rho \left(T - \varepsilon \left(\frac{1 - e^{-2\rho\varepsilon}}{1 + e^{-2\rho\varepsilon}} \right) \right) < \rho T.
 \end{aligned}$$

The fact, noted above, that $x(T)$ is convex, together with Jensen’s inequality, implies that $E_{\tilde{T}}x(\tilde{T}) > x(E_{\tilde{T}}\tilde{T}) = x(T)$, so the first line implies that moving from a certain to a random event time increases the maximum premium. The second line shows that the elasticity of the premium, with respect to the PRTP, is a *weighted* sum of the probability-weighted elasticities; the first weight exceeds the second. The previous subsection shows that the elasticity is higher at the high event time. Thus, the elasticity in the stochastic case, $\phi(x')$, is lower than the expectation of the deterministic elasticities, ρT .

Where \tilde{T} is an exponentially distributed random variable with hazard rate h , $E(\tilde{T}) = \frac{1}{h}$. The expected present value cost of the uncertain event is $E\left(e^{-\rho T} \frac{100}{\rho}\right) = \frac{100h}{\rho(\rho+h)}$. Equating this expression to the cost of the premium, $\frac{x'}{\rho}$, gives the maximum premium that society would pay for perfect insurance, $x' = \frac{h100}{h+\rho}$. Setting $\frac{1}{h} = T$ makes the stochastic and deterministic models comparable, and yields $x' = \frac{100}{1+\rho T}$.

2.1.3 Comparison of Deterministic and Stochastic Event Time

For $T = 200$ as above, the maximum acceptable premium in the stochastic setting, x' , ranges from 33.3 to 14.3 as ρ ranges from 0.01 to 0.03. A change from the deterministic to the stochastic event time increases the maximum premium by a factor that ranges from 2.5 for $\rho = 0.01$ to 57 for $\rho = 0.03$ (Fig. 1).

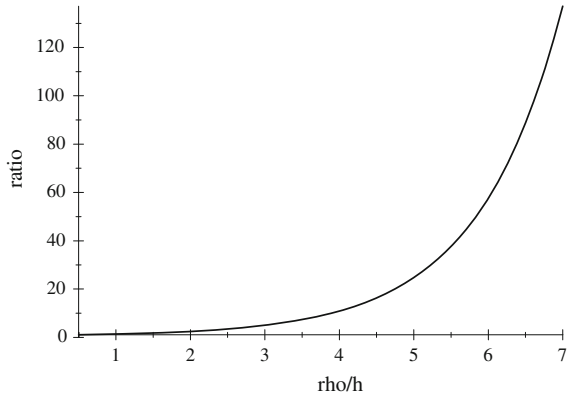
For general T , with $h = \frac{1}{T}$, the ratio of the maximum premium in the stochastic compared to the deterministic event time is

$$\text{premium ratio: } \frac{x'}{x} = \frac{\exp(\frac{\rho}{h})}{1 + \frac{\rho}{h}} > 1.$$

Figure 2 shows the graph of this ratio of premiums as a function of the ratio $\frac{\rho}{h}$. For example, with an annual discount rate $\rho = 0.02$, $\frac{\rho}{h} = 0.5$ corresponds to an expected event time of $T = 25$ years and a premium ratio of 1.1; $\frac{\rho}{h} = 7$ corresponds to an expected event time of 350 years and a premium ratio of 137. Thus, for low probability events, moving from a deterministic to a stochastic model increases the maximum premium by two orders of magnitude. “Low probability” means that the hazard rate is small relative to the PRTP.

Although the maximum premium is higher in the stochastic compared to the deterministic setting, the premium in the former is much less sensitive to the PRTP. The (absolute value) elasticity of x (in the deterministic case) with respect to ρ

Fig. 2 The ratio of the maximum premium under stochastic to the maximum premium under deterministic event time $\left(\frac{x'}{x}\right)$ as a function of $\frac{\rho}{h}$



is $\phi(x) = \rho T = \frac{\rho}{h}$, which is linear in ρ ; and the elasticity of x' (in the stochastic case) with respect to ρ is $\phi(x') = \frac{\rho}{h+\rho}$, which decreases in ρ . The ratio of these elasticities is

$$\text{elasticity ratio: } \frac{\phi(x')}{\phi(x)} = \frac{\frac{\rho}{h+\rho}}{\frac{\rho}{h}} = \frac{h}{h+\rho},$$

which is small for “low probability events”.

This example illustrates the two features described above: moving from the deterministic to a stochastic setting increases the maximum premium, and also makes it less sensitive to the PRTP.

2.2 Positive Consumption Growth

If per capita income is expected to grow, future generations will be richer than current generations. With decreasing marginal utility of income, growth makes people today less willing to sacrifice to avoid future damages. In the deterministic setting, the Ramsey formula gives the social discount rate (SDR), r , as a function of the pure rate of time preference, ρ , the growth rate, g , and the elasticity of marginal utility (the inverse intertemporal elasticity of substitution), η : $r = \rho + \eta g$. With zero growth or infinite intertemporal elasticity of substitution, the social discount rate equals the pure rate of time preference. Positive growth and finite intertemporal elasticity of substitution increases the social discount rate.

I normalize consumption at $t = 0$ to 1, and assume that growth is constant, g , so potential consumption at time $t > 0$ prior to the event is e^{gt} . The event results in a permanent $\Delta \times 100\%$ reduction in potential consumption flow, so at a post-event time t , consumption is $c = e^{gt} (1 - \Delta)$. The insurance premium is deducted from poten-

tial consumption (not investment), so payment of the premium does not change the growth rate. The premium needed to eliminate the risk is proportional to the value-at-risk, Δe^{gt} , with proportionality factor X . If society pays the premium, consumption is $c = e^{gt} (1 - \Delta X)$. Utility (u) is isoelastic in consumption: $u(c) = \frac{c^{1-\eta}-1}{1-\eta}$.²

In order for the premia to be easily compared with the zero-growth analogs, I present them as a percent of Δ . In the deterministic case, the maximum premium (as a percent of Δ) that society is willing to pay for perfect insurance is

$$y = \frac{1 - \left((1 - (1 - \Delta)^{1-\eta}) \left(1 - e^{-\frac{(\rho+(\eta-1)g)}{h}} \right) + (1 - \Delta)^{1-\eta} \right)^{\frac{1}{1-\eta}}}{\Delta} 100. \tag{1}$$

In the exponentially distributed case, the maximum premium that society is willing to pay for perfect insurance is

$$z = \frac{1 - \left((1 - (1 - \Delta)^{1-\eta}) \left(1 - \frac{h}{(\rho+g(\eta-1)+h)} \right) + (1 - \Delta)^{1-\eta} \right)^{\frac{1}{1-\eta}}}{\Delta} 100. \tag{2}$$

As a consistency check, note that for $g = 0 = \eta$, y and z collapse to their deterministic and stochastic analogs in Sect. 2.1. Although the social discount rate, r , equals the PRTP if either $g = 0$ or $\eta = 0$, we need both of those equalities to hold in order for the premia in this section to equal their analogs in Sect. 2.1. For example, with $\eta > 0$, utility is nonlinear in consumption, even if there is no growth.

Figure 3 graphs the premia y and z as functions of ρ , and Fig. 4 graphs the elasticity of these premia with respect to ρ .³ With $g = 0 = \eta$, the elasticity in the deterministic case is linear, equal to $\frac{\rho}{h} = \rho T$. With $g > 0$ and $\eta > 0$ the elasticity with respect to ρ is approximately linear in ρ under the deterministic event time, ranging between 1 and 4 as ρ ranges from 0.005 to 0.02 (for $T = 200$). By this measure, it appears that optimal policy is very sensitive to discounting assumptions. However, the elasticity of the premium with respect to ρ under stochastic event time is small and insensitive to ρ , reaching only about 0.5 at $\rho = 0.02$ and scarcely increasing thereafter.

Comparing these figures to Fig. 1 shows that introducing growth ($g > 0$) and measuring costs in consumption rather than utility units ($\eta > 0$) leaves unchanged the qualitative comparisons between deterministic and stochastic event times discussed above: The maximum premium under the stochastic event time is much larger, but

²The model in this section is based on Karp and Tsur (2008). That paper considers non-constant PRTP, whereas here the PRTP is constant. In the model here, society can eliminate the risk, whereas in the Karp and Tsur model society is able only to prevent the hazard from increasing. Neither model is nested in the other. Appendix 1 derives Eqs. 1 and 2.

³Figures 11, and 12 in Appendix 1 shows graphs of the premia as functions of η , and g . Those parameters have the same qualitative effect as ρ on the two premia (under deterministic and stochastic event time), although of course the magnitudes of their effects are different.

Fig. 3 Maximum premium (as a percent of cost of event) as a function of ρ , under certain event time (*solid*) and random event time (*dashed*) for $T = 200 = \frac{1}{h}$, $\Delta = 0.3$, $g = 0.01$ and $\eta = 2$

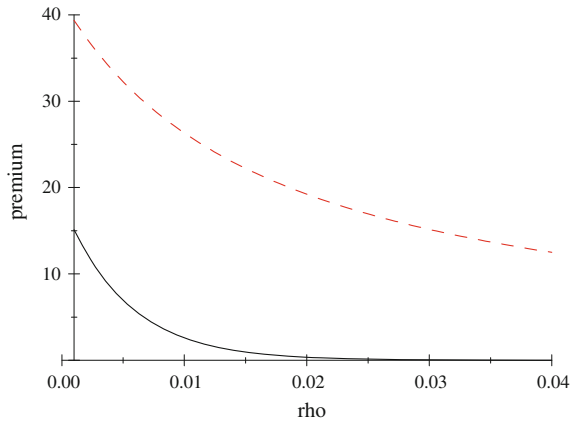
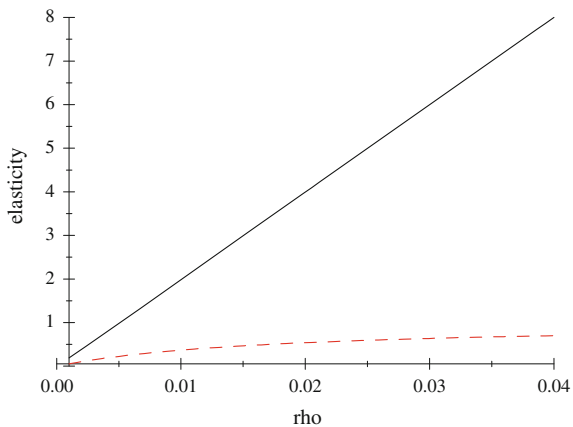


Fig. 4 The elasticity of the maximum premium under certain event time (*solid*) and under random event time (*dashed*) for $T = 200 = \frac{1}{h}$, $\Delta = 0.4$, and $\eta = 2$



also much less sensitive to the PRTP (and other discounting parameters) compared to the premium under the deterministic event time.

Arrow (2007) examines the effect of the discount rate on our willingness to avoid climate change, posing the question in terms of growth rates rather than levels of damages. His examples suggest that the benefits of significant climate policy outweighs the cost to such a large extent that the cost-benefit ratio is not sensitive to discounting assumptions.

3 Optimization Models

The comparison in Sect. 2 is based on a cost-benefit exercise. That material shows that moving from a deterministic to random event time increases the maximum acceptable insurance premium and makes that premium less sensitive to discount-

ing assumptions. This section considers the sensitivity of policies in an optimizing framework. I first consider two analytic examples, which (together with the results above) illustrate circumstances where greater nonlinearity in the model makes the optimal policy less sensitive to the discount rate. I then assess this relation using several climate policy models.

3.1 Analytic Examples

Two familiar and tractable renewable resource models illustrate the relation between the steady state and the discount rate. In the simplest fishery model, the change in the stock, S , equals a growth function, $f(S)$, minus harvest, h ; and the utility flow, $u(h)$, depends on harvest but not on the stock. The growth equation is $\frac{dS}{dt} = \dot{S} = f(S) - h$ and the payoff, evaluated at time 0, is the infinite stream of discounted utility, $\int_0^\infty e^{-\rho t} u(h_t) dt$, where the discount rate is ρ . The optimal (interior) steady state is the solution to $f'(S) = \rho$. The elasticity of the steady state, with respect to the discount rate, is

$$-\frac{\rho}{S_\infty} \frac{dS_\infty}{d\rho} = \frac{-1 f'(S_\infty)}{S_\infty f''(S_\infty)}.$$

The elasticity of the steady state has the same form as the inverse of the Arrow-Pratt risk aversion, but here applied to the growth (instead of utility) function. As the growth function becomes more steeply curved (evaluated at the steady state) the steady state becomes less sensitive to the discount rate.

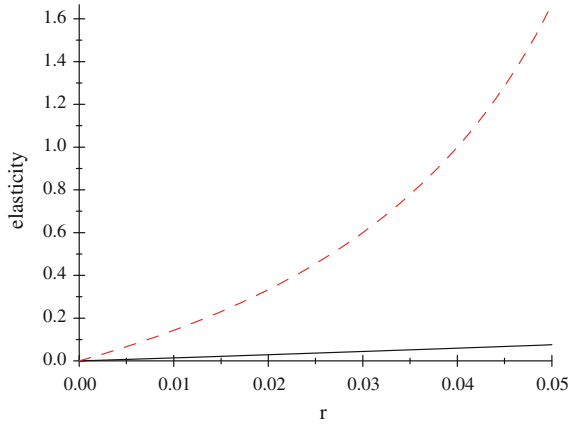
For the logistic growth function, $f = \varphi S \left(1 - \frac{S}{k}\right)$; φ is the intrinsic growth rate and k is the carrying capacity. For this function, the elasticity of the steady state with respect to the discount rate is

$$\frac{-1 f'(S_\infty)}{S_\infty f''(S_\infty)} = \frac{\rho}{\varphi - \rho}.$$

A higher intrinsic growth rate, φ , increases the curvature of the growth function, and makes the steady state less sensitive to the discount rate. Figure 5 graphs the elasticity as a function of the PRTP for estimates of the intrinsic growth rate $\varphi = 0.71$ for Pacific Halibut and $\varphi = 0.08$ for Antarctic fin-whale (Clark 1975). Even for the slow-growing fin-whale, the elasticity is less than 1 for reasonable discount rates. For $\rho = \varphi$, where it is optimal to drive the stock to extinction, the elasticity is infinite.

The second simplest fishery model allows harvest costs (and thus the utility flow) to depend on the stock, but assumes that the flow of benefits due to harvest (like the growth function) is linear in the harvest. If the price per unit of harvest, p , is constant and the harvest costs linear in harvest, h , then the flow of benefit equals $(p - c(S_t)) h_t$. The payoff is the discounted stream of benefits, $\int_0^\infty e^{-\rho t} (p - c(S_t)) h_t dt$. In this case,

Fig. 5 The elasticity of the steady state with respect to the ρ in the zero-extraction cost model, for Pacific Halibut (*solid*, with $\varphi = 0.71$) and Antarctic Fin-whale (*dashed*, with $\varphi = 0.08$)

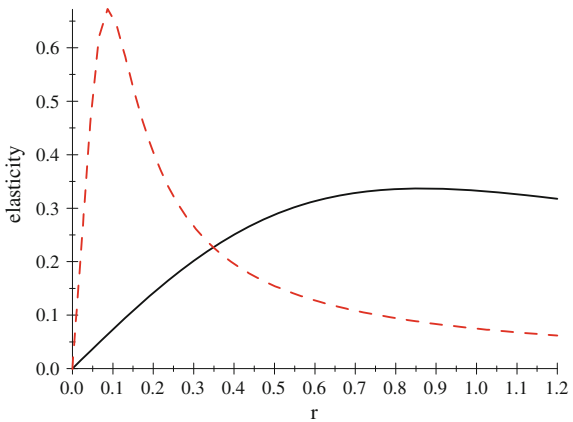


the optimal harvest policy is bang-bang: the harvest is set to its maximum feasible level if the stock is above the steady state, and the harvest is set to 0 if stock is below the steady state; when the stock is at the steady state, the harvest maintains it at that level: $h_\infty = f(S_\infty)$. The steady state is the solution to

$$\rho = f'(S) - \frac{c'(S)f(S)}{p - c(S)}. \tag{3}$$

Using the logistic growth model and $c(S) = \frac{c}{S}$, with parameter values taken from Clark (1975), Fig. 6 (taken from Ekeland et al. 2012) shows the elasticity of the steady state stock for Pacific halibut (solid graph) and Antarctic fin-whale (dashed graph). The elasticities are non-monotonic in the discount rate, but both are well below 1. For reasonable values of ρ (i.e., values much less than 0.2, or 20% per year), the steady state is much less sensitive to the discount rate

Fig. 6 Elasticity of steady state with respect to ρ for Pacific halibut (*solid graph*, using $K = 80.5 \times 10^6$ and $\frac{c}{p} = 17.7 \times 10^6$ kg) and for Antarctic fin-whale (*dashed graph*, where $K = 400,000$ whales and $\frac{c}{p} = 40,000$ whales). Parameter values from Clark (1979)



for the fast-growing halibut, compared to the slow growing whales. Again, a higher intrinsic growth rate implies a more non-linear growth function, and for reasonable parameter values causes the optimal policy to be less sensitive to the discount rate.

3.2 *Climate-Related Models*

The Stern Review (2006) (hereafter SR) is perhaps the most widely discussed document on climate policy during the past decade. Several economists focused on the SR's discounting assumptions. SR chose a PRTP of $\rho = 0.001$, an elasticity of marginal utility of $\eta = 1$, and a growth rate of 0.013, implying a social discount rate (SDR) of $r = 0.014$ (or 1.4 %).

Nordhaus (2007) illustrates the importance of these discounting assumptions by comparing the results of three runs of the DICE model. Two of these runs use combinations of the PRTP and η consistent with a SDR of about 5.5 %, almost four times the level in the SR. With the Nordhaus values, the optimal carbon tax in the near term is approximately \$35/ton Carbon (or \$9.5/ ton CO₂), and the optimal level of abatement in the near term about 14 % of Business as Usual (BAU) emissions. A third run, using the SR's values of ρ and η (together with the DICE assumptions about growth) led to a carbon tax of \$350/ton and a 53 % level of abatement, close to the level that the SR recommends. Thus, the carbon tax increases by a factor of 10 and abatement increases by a factor of $\frac{53}{14} = 3.8$ with the decrease in the SDR. Because abatement costs are convex, the percent change in the tax (caused by changes in the discounting assumptions) is larger than the percent change in abatement. Based on these numbers, an estimate of the elasticity of the tax with respect to the discount rate is approximately $\frac{10}{4} = 2.5$ and the elasticity of abatement with respect to the tax is approximately $\frac{3.8}{4} = 0.95$. These values are in the range of elasticities of the maximum premium, in the deterministic cost-benefit setting shown in Fig. 4.

In a different context (focused on the effect of catastrophic damages) Nordhaus (2009) compares optimal policy under a PRTP of 0.015 and 0.001, holding other DICE parameters (including the elasticity of marginal utility) at their baseline levels. He reports that the reduction in PRTP increases the optimal carbon tax from \$42/tC to \$102/tC. This 2.4-fold increase in the optimal tax is much less than the 10-fold increase reported in Nordhaus (2007), where both the PRTP and the elasticity are changed.⁴

Karp (2005) uses a linear-quadratic model, calibrated to reflect abatement costs and climate-related damages that are of the same order or magnitude as in DICE. In this stationary, partial equilibrium model, there is no growth, so the pure rate

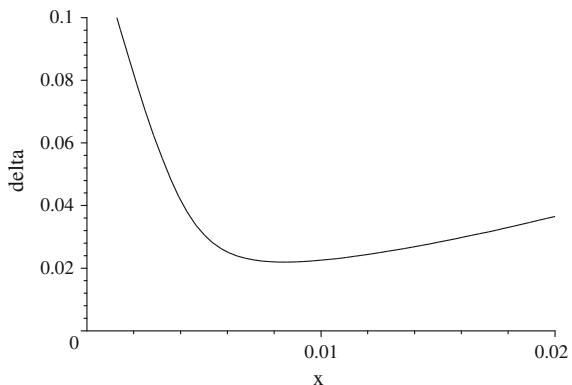
⁴The increase in the tax from \$42/tC to \$102/tC leads to a fall in per capita income (during the period when it is lowest, presumably the first period) from \$6,801 to \$6,799, i.e. about 0.03 %. In view of parameter uncertainty, a 0.03 % reduction in per capita income, equivalently a \$15 billion increase in aggregate abatement costs (at Gross World Product—GWP—of \$50 trillion), is close to rounding error.

of time preference equals the social discount rate. A decrease in the discount rate from 3 to 1 % increases abatement in the first period by a factor of 2.5. Nordhaus's (2007) experiments, described above, reduce the pure rate of time preference by approximately 1.5 %, increasing abatement by a factor of 3.8. By this measure, the sensitivity of policy to the discount rate is of the same order of magnitude in the two models.

Fujii and Karp (2008) provide a more involved analysis of the role of discounting, using a one-state variable model calibrated to approximate the costs and benefits underlying the SR recommendations. In that model, Δ_t is the consumption loss due to mitigation expenditures and remaining climate-related damage, as a fraction of the no-damage no-control scenario (i.e. where there is no potential for climate damage). Reducing the discount rate increases abatement expenditures and reduces the trajectory of the damages, as expected. However, the magnitude of those induced changes was much smaller than the previous studies led us to expect. This insensitivity is probably not due to a peculiarity of our climate model, because Fujii and Karp (2006) find a similar relation using a standard renewable resource model. The analytic examples in Sect. 3.1 provide some insight into this result.

The highly nonlinear relation between expenditures and damages may explain this insensitivity. Figure 7, taken from Fujii and Karp (2008), shows the graph of the steady state climate-related costs, Δ , as a function of steady state expenditures (expressed as a fraction of income) x . This graph reaches a global minimum where climate related expenditure is 0.845 % of consumption. This is the optimal steady state level of expenditures under zero discounting. Total costs fall rapidly for smaller values of x . This graph implies that small increases in expenditure, below the global optimum, achieve significant reductions in total costs. Therefore, a very low SDR achieves nearly the global minimum, and even substantially larger SDRs take us close to the global minimum. Because initial expenditures (compared to steady state expenditures) are even less sensitive to the SDR, the entire trajectory is “relatively insensitive” to the discount rate.

Fig. 7 The graph of steady state costs, Δ , as a function of steady state expenditures, x



Gerlagh and Liski (2012), using an extension of Golosov et al. (2013), calibrate a model to reflect climate-related damages and abatement costs similar to those in Norhaus (2009). They find that the optimal tax is very sensitive to discounting assumptions. The logarithmic utility is the inverse of the exponential damages, causing the flow payoff to be linear in the cumulative emissions (a state variable); the state dynamics are also linear in emissions. Their model is thus linear in the stock of emissions. I conjecture that this linearity contributes to the sensitivity of the tax, with respect to the discount rate(s).

This selective review shows that the discount rate matters a great deal in some, but not all, dynamic optimization models. All of this evidence comes from specific models or specific ways of presenting the tradeoff between abatement costs and avoided damages, and therefore it cannot lead to general conclusions. It provides some (but certainly not conclusive) examples where optimal policy tends to be less sensitive to the PRTP when the model is highly nonlinear.

4 Do Catastrophes Swamp Discounting?

Weitzman (2009) examines the effect of parameter uncertainty on the social discount rate. Using a two period model, representing the current period and the distant future, he calculates the marginal expected value of transferring the first unit of certain consumption from the present into an uncertain future. His chief result is open to several interpretations. In my view, a “modest” interpretation is correct and useful. A controversial interpretation is that the result undermines our ability to sensibly apply cost-benefit analysis to situations where there is uncertainty about catastrophic events. A “corollary” to this interpretation is that the recognition of catastrophic events makes discounting a second order issue. I think that both the controversial interpretation and the corollary to it are incorrect.

In order to explain these points, I consider a simplified version of his model. Let c be the known current consumption, c' the random future consumption, v the number of certain units of consumption transferred from the current period to the future, β the utility discount factor, and u the utility of consumption.⁵ The social discount factor for consumption, i.e. the marginal rate of substitution between “the first” additional certain unit of consumption today and in the future, is

$$\Gamma = -\beta E_{c'} \left(\frac{\frac{du(c'+v)}{dv}}{\frac{du(c-v)}{dv}} \right)_{|v=0} . \tag{4}$$

⁵This model assumes that the sacrifice of one unit of consumption today makes one unit available in the future period. If instead, the sacrifice of v units makes $\theta(v)$ units available the future, the numerator of the right side of Eq. 4 becomes $\frac{du(c'+\theta(v))}{dv} \theta'(v)$. These and other extensions can be easily treated, but the simple model that I use is adequate for my purposes.

The chief result is that, under the assumptions of the model, $\Gamma = \infty$; Weitzman dubs this result “the dismal theorem”.

The model includes a number of important features, including: (i) the uncertainty about c' is such that there is a “significant” probability that its realization is 0; (ii) the marginal utility of consumption at $c' = 0$ is infinite; and (iii) it is possible to transfer a certain unit of consumption into the future. Features (ii) and (iii) are assumptions, but (i) is an implication of the assumption that the variance of c' is an unknown parameter, and the decision-maker’s subjective distribution for this parameter has “fat tails”.

Any of these assumptions can be criticized, but in my view, a more fundamental issue involves the interpretation of the dismal theorem. A modest interpretation is that uncertainty about the distribution of a random variable can significantly increase “overall uncertainty” about this random variable, leading to a much higher risk premium (and therefore a much higher willingness to transfer consumption from the present into the future) relative to the situation where the distribution of the random variable is known. This modest interpretation is not controversial.

An extreme interpretation is that under conditions where the dismal theorem holds, society should be willing to make essentially any sacrifice to transfer a unit of certain consumption into the future. That interpretation is also not controversial, because it is so obviously wrong. Even with 0 discounting (in this two period stationary model with the same utility function in both periods), we would never be willing to transfer to the future more than half of what we currently have.

The controversial interpretation is that the dismal theorem substantially undermines our ability to sensibly apply cost-benefit analysis to situations with “deep uncertainty” about catastrophic risks. The basis for this claim is that in order to use the social discount rate given in Eq. (4), we need to modify the model so that Γ is finite. Weitzman suggests ways of doing this, such as truncating a distribution or changing an assumption about the utility function or its argument, in order to make Γ finite. The alleged problem is that the resulting Γ is extremely sensitive to the particular device that we use to render it finite. Because we do not have a consensus about how to achieve this finite value, we do not have a good way to select from the many extremely large and possibly very different social discount rates. In this setting, it is difficult to use cost-benefit analysis.

This controversial interpretation is not persuasive. Horowitz and Lange (2008) identify clearly the nub of the misunderstanding; I rephrase their explanation. Nordhaus (2008) also identifies this issue, and he provides numerical results using DICE to illustrate how cost-benefit analysis can be used even when damages are extremely large.⁶ Millner (2013) contains a thorough discussion of the Dismal Theorem, together with extensions.

⁶These numerical results are interesting, but someone who accepts the controversial interpretation of the dismal theorem will not regard them as a convincing counter-argument to that interpretation. All of the numerical experiments arise in a deterministic context, and in that respect they do not really confront the dismal theorem, which makes sense only in a setting with risk and uncertainty. For this reason, I think that Horowitz and Lange’s very simple treatment of the problem is particularly helpful.

The problem with the controversial interpretation is that the value of Γ in Eq. (4) is essentially irrelevant for cost-benefit analysis. This expression, which is evaluated at $v = 0$, gives the value of the “first” marginal unit transferred. The fact that the derivative may be infinite does not, of course, imply that the value of transferring one (non-infinitesimal) unit of sure consumption is infinite. If we want to approximate the value of a function, it makes no sense to use a Taylor approximation evaluated where that function’s first derivative is infinite. We would make (essentially) this mistake if we were to use Eq. (4) as a basis for cost-benefit analysis with climate policy. The only information that we obtain by learning that the value of the derivative evaluated at $v = 0$ is infinite is that a non-infinitesimal policy response must be optimal. This fact is worth knowing, but it does not create problems for using cost-benefit analysis.

Although I think that the controversial interpretation is unsound, it has a “corollary” that is not so easy to dismiss. This corollary states that catastrophic risks swamps the effect of the pure rate of time preference. The idea is that because the expectation of the term in parenthesis in Eq. (4) is so large, the magnitude of β , and thus of the pure rate of time preference, is relatively unimportant. Norhaus (2009), despite his criticism of the controversial interpretation of the dismal theorem, endorses this view (emphasis added.):

...discounting is a second-order issue in the context of catastrophic outcomes. ... If the future outlook is indeed catastrophic, that is understood, and policies are undertaken, the discount rate has little effect on the estimate of the social cost of carbon *or to the optimal mitigation policy*.

This corollary may hold in specific settings, but it would be surprising if it is a general feature of catastrophic risk. The magnitude of the expectation of the term in parenthesis in Eq. (4), evaluated at $v = 0$, can certainly swamp the magnitude of β ; but I have just noted that the former term is irrelevant for cost-benefit analysis (beyond telling us that non-infinitesimal policy is optimal).

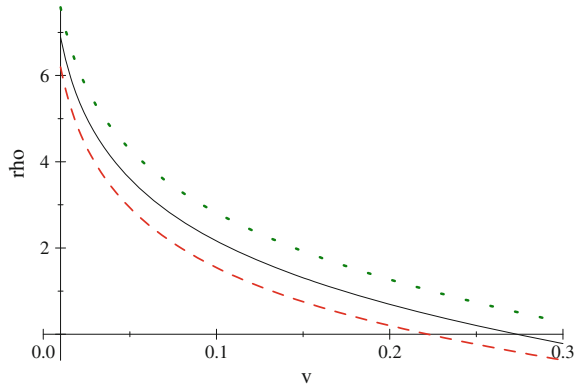
In order to get a sense of whether the corollary is likely to hold, and also to illustrate why the controversial interpretation of the dismal theorem is not persuasive, I use an example with $u = \frac{c^{1-\eta}}{1-\eta}$. Set $\beta = \exp(-\rho T)$ and choose a unit of time equal to a century. With this choice of units, ρ is the annual pure rate of time preference expressed as a percent. Suppose that c' takes the value c with probability $1 - p$ and the value 0 with probability p . For $p > 0$ the right side of Eq. (4) is infinite, as in the dismal theorem. The optimization problem is

$$\max_v \left(u(c - v) + \beta \left[(1 - p) u(c + v) + pu(v) \right] \right).$$

Normalize by setting $c = 1$, so that v equals the fraction of current consumption that we transfer into the future. With a bit of manipulation, the first order condition for the optimal v is

$$\rho = \frac{1}{T} \ln \left((1 - p) \left(\frac{1 - v}{1 + v} \right)^\eta + p \left(\frac{1 - v}{v} \right)^\eta \right). \tag{5}$$

Fig. 8 The relation between the transfer, $v \in [0.01, 0.3]$, (the x axis) and the annual percentage discount rate, ρ , for $p = 0.05$ (dashed), $p = 0.1$ (solid) and $p = 0.2$ (dotted), with $\eta = 2$ and $T = 1$ century



Using $\eta = 2$ and $T = 1$ (so that the “future” is a century from now), Fig. 8 shows the relation between the *annual* percentage pure rate of time preference, ρ , and the optimal value of v for p equal to 0.05, 0.1 and 0.2. The figure shows that the fact that the expression in Eq. (4) is infinite does not cause any problem in determining an optimal value of the transfer. It also illustrates the less-obvious point that catastrophic risk does not swamp the effect of discounting, in determining the optimal level of the transfer.

We get a bit more insight into the relative importance of the PRTP, ρ , and the probability of catastrophe, p , by taking the ratio of the elasticities of the optimal v with respect to these variables. Define the elasticities and the ratio of elasticities as:

$$\frac{dv}{dp} \frac{p}{v} = \alpha; \quad \frac{dv}{d\rho} \frac{\rho}{v} = \phi; \quad \tau = -\frac{\phi}{\alpha}.$$

For $\tau > 1$, v is more sensitive to discounting than to the probability of catastrophe, and the reverse holds for $\tau < 1$. For the isoelastic example,

$$\tau = \left(1 + \frac{v^\eta}{p((v+1)^\eta - v^\eta)} \right) \ln \left(p \left(\frac{1-v}{v} \right)^\eta + (1-p) \left(\frac{1-v}{1+v} \right)^\eta \right), \quad (6)$$

where Eq. (5) implicitly defines $v = v(\rho, p, T)$. The ratio of elasticities shown in Eq. 6 depends explicitly on p and v and it depends implicitly on ρT , because that product determines the optimal value of v , given p and η .

Figure 9 shows the graphs of τ for $T = 1$ and $p = 0.05$, i.e. a 5% chance of the catastrophe within 1 century, for $\eta = 2$ (the solid curve) and $\eta = 0.5$, the dashed curve. If $\eta = 2$, the equilibrium transfer is more sensitive to the discount rate than to the probability of catastrophe ($\tau > 1$) if $\rho > 0.74\%$; for $\eta = 0.5$, the transfer is more sensitive to discounting than to the probability of catastrophe if $\rho > 0.21\%$. Both of these critical values are much larger than the 0.1% PRTP used in the Stern Review, but they are much smaller than PRTPs used for other studies, e.g. the various

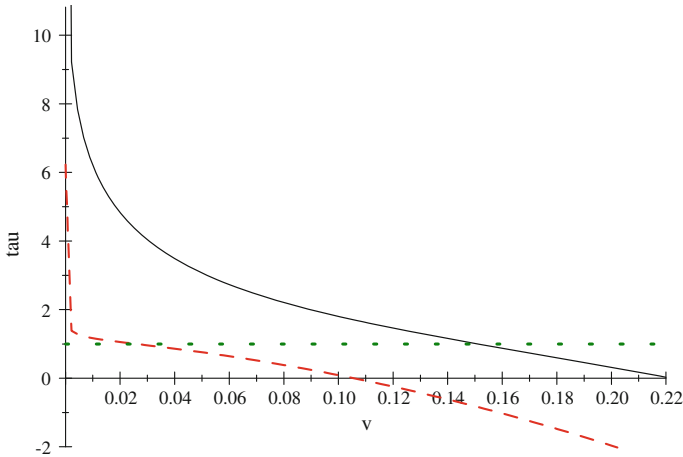


Fig. 9 The graph of τ for $\eta=2$ (solid) and $\eta=0.5$ (dashed), with $T=1$ and $p=0.05$. For $\eta=2$, ν ranges from 0.03 to 0.22 as ρ ranges from 4 to 0.1%. For $\eta=0.5$, ν ranges from 0.06 to 9×10^{-7} as ρ ranges from 4 to 0.1%. The dotted line shows $\tau = 1$

incarnations of DICE. Thus, the discount rate might be either more or less important than the probability of catastrophe in determining the optimal transfer; it is likely to be less important for discount rates commonly used.

Here I explain how to interpret Fig. 9. The dotted line shows $\tau = 1$; for values of ν where the graph lies above the dotted line, the equilibrium is more sensitive to the PRTP ρ , than to the risk of catastrophe, p . As noted above, for given values of η, T, p , the equilibrium value of ν is a decreasing function of ρ . For $\eta = 2$, $\nu \in [3 \times 10^{-2}, 0.22]$ as ρ falls from a PRTP of 4% (per year) to 0.1%. For $\eta = 0.5$, $\nu \in [9 \times 10^{-7}, 0.06]$ as ρ falls from a PRTP of 4% (per year) to 0.1%. Given p, η , Eq. 6 determines the value of ν at which $\tau = 1$. Given this value of ν , and p, T, η , Eq. 5 determines the critical value of ρ at which $\tau = 1$. Those critical values are $\rho = 0.74\%$ for $\eta = 2$ and $\rho = 0.21\%$ for $\eta = 0.5$.

5 How Do We View the Distant Future?

The sections above examine the importance, to climate policy, of the magnitude of the PRTP. This section considers the applicability of a constant PRTP in the climate context. The PRTP measures a person’s willingness to transfer utility between two points in time. Even if this person uses a constant PRTP to evaluate a utility transfer from one period to another for herself, there is no reason that society would use the same constant rate to evaluate an intertemporal transfer between two different people. If we are better able to distinguish among people who are closer to us, either in space, time, or genetically, compared to people who are further from us, then we plausibly

discount hyperbolically (with respect to space, time, or genetics), not at a constant rate. More generally, if our willingness to transfer utility between two generations depends not only on the distance between those two generations, but also on their distance from us, then we discount at a non-constant rate.

Many papers discuss the plausibility of hyperbolic (more generally, nonconstant) discounting, both in the context of individual decision problems (Phelps and Pollack 1968), (Laibson 1997), (Barro 1999), (Heal 1998), (Harris and Laibson 2001) and for societal problems such as climate change (Cropper and Laibson 1999), (Karp 2005), (Fujii and Karp 2008), (Karp and Tsur 2011), (Schneider et al. 2012), (Gerlagh and Liski 2012), (Karp 2013b). The “Weber-Fechner law” states that human response to a change in stimulus, such as sound or light, is inversely proportional to the pre-existing stimulus. For example, from the standpoint of period 0, delaying utility from period 1 to period 2 represents a much larger proportional increase in delay than does a delay from period 10 to 11, even though the absolute increase in delay is the same in the two cases. Heal (2001) invokes this observation as justification for a decreasing discount rate. Applied to discounting, the “law” is consistent with a discount factor of t^{-K} , where K is a positive constant. Heal calls this “logarithmic discounting”.

Ramsey (1928) remarked “My picture of the world is drawn in perspective. ... I apply my perspective not merely to space but also to time.” Karp (2013c) shows that perspective applied to space corresponds to a special case of logarithmic “spatial discounting.” To the extent that spatial perspective provides a useful analogy for temporal perspective, this result provides further support for the hypothesis that our view of the world corresponds to hyperbolic discounting.

A two-parameter model of preferences makes it possible to distinguish between intertemporal utility transfers for a single individual and across individuals. The PRTP, $\rho \geq 0$, measures a person’s willingness to sacrifice their own future utility in order to increase their own current utility. A second discounting parameter, denoted $\lambda \geq 0$, measures a planner’s willingness to transfer utility across *different people at different points in time*. I consider a pure public good (or bad), in which every person alive in a period has the same utility flow. In this setting, there is no reason to consider transfers between different people alive at a point in time. A richer model would be able to evaluate transfers across different people at different points in time, and across different people at the same point in time, but that is too much to ask of a two-parameter model. I assume that the planner gives equal weight to all people currently alive, and in that respect is utilitarian.

Consider the simplest constant-population OLG model in which agents lives for two periods. A “public project”, e.g. emissions of a certain amount of greenhouse gasses, increases current aggregate utility, and reduces next-period aggregate utility by one unit.⁷ What is the minimal increase in current aggregate utility needed to justify these emissions? The evaluation is complicated by the fact that it involves a utility exchange both between the current young people and their future old selves,

⁷In order to keep this example simple, suppose that current emissions affect only next period utility. Obviously, a realistic climate model has to take into account that today’s emissions persist in the atmosphere, affecting utility over a long span.

and also between the next period young people and the current old people. That is the nature of climate policy.

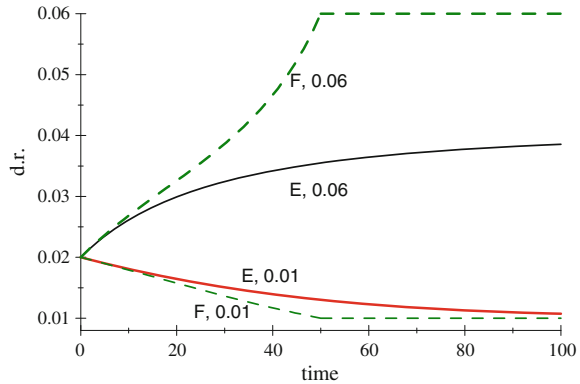
An increase in current utility of $e^{-\rho}$ leaves the current young people indifferent. If $\lambda = \rho$, the planner is willing to transfer utility between two people one period apart at the same rate as a person would transform utility between her future and current selves. If $\lambda = 0$, the planner gives equal weight to the future young and the current old people. If $\lambda = \infty$, the planner gives zero weight to the future young. It is worth emphasizing that the value $\lambda = 0$ (*not* $\lambda = \rho$) implies that the planner treats currently living and not-yet-born agents symmetrically. Readers will have different views about the “reasonable” relation between ρ and λ , but it is clear that smaller values of λ imply greater weight to the not-yet-born, and in that respect correspond to greater altruism.

In this model, half of the people are young, and half are old in any period. The planner would accept the project if and only if the current increase in aggregate utility is no less than $D(1) = 0.5(e^{-\rho} + e^{-\lambda})$, the planner’s one period discount factor. This planner would accept a project that lowers aggregate utility by one unit t periods from now if it increases current aggregate utility by no less than $D(t) = 0.5(e^{\lambda(t-1)}e^{-\rho} + e^{-\lambda t})$, the planner’s t period discount factor. The term $0.5e^{\lambda(t-1)}e^{-\rho}$ equals the present value at $t - 1$ to the agents born at $t - 1$ of the one unit loss in utility ($e^{-\rho}$) times the planner’s weight on those people’s welfare, $0.5e^{\lambda(t-1)}$. The second term, $0.5e^{-\lambda t}$, equals the weight the planner puts on the utility of people born t periods in the future. Defining $\beta = \frac{e^{-\rho} + e^{-\lambda}}{2}e^{\lambda}$ and $\delta = e^{-\lambda}$, gives $D(1) = \beta\delta$ and $D(t) = \beta\delta^t$ for $t > 1$. This particular form of discounting is known as β, δ or quasi-hyperbolic discounting (Phelps and Pollack 1968). The case usually emphasized is $\lambda < \rho$, where $\beta < 1$. This model is typically used for a single-agent decision problem, to capture present bias, whereas I use it to distinguish between transfers across time for a single agent from transfers across time between different agents.

Karp (2013b) provides the formula for the discount factor in a generalization of this model, where each agent lives for T years, and time is continuous. Ekeland and Lazrak (2012) provide the discount function for a model in which agents’ lifetime is exponentially distributed with mortality rate (hazard rate) θ . Setting $\theta = \frac{1}{T}$ makes the two models comparable. Figure 10, taken from Karp (2013b), shows the discount rates under exponentially distributed and finite lifetimes, for parameter values $\rho = 0.02 = \theta = \frac{1}{T}$, and for $\lambda \in \{0.01, 0.06\}$. The planner’s discount rate falls if $\lambda < \rho$ (as with hyperbolic discounting), and rises for $\lambda > \rho$. In both of these cases, the planner’s preferences are time inconsistent. Preferences are time consistent only for $\lambda = \rho$, where the planner makes no distinction between transferring utility across time for the same individual and between two different individuals.

For $\lambda \neq \rho$ in this two-parameter model, time consistency requires that the planner gives less weight to the old than to the young agent in a period (Karp 2013a). Obstfeld (1988) use this type of time consistent model to study fiscal policy, and (Schneider et al. 2012) use it to study climate policy. Those authors discount the utility of currently living agents back to the time of their birth. This procedure means that, for $\lambda < r$, currently living people have less weight in the planner’s objective function,

Fig. 10 Discount rates (d.r.) for $\theta = 0.02 = r = \frac{1}{7}$. *Solid curves* (labelled E) correspond to exponentially distributed lifetime and *dashed curves* (labelled F) correspond to fixed lifetime. Numerical values in label show value of λ



the older they are; with this procedure, the planner’s preferences are time consistent. However, if the planner gives all currently living people equal weight, then for $\lambda \neq r$, the planner’s preferences are time inconsistent.

6 Conclusion

Agents who discount the utility of future generations may be unwilling to make much of a sacrifice to avoid or ameliorate large damages that occur in the probably distant future. The degree of sacrifice we are willing to undertake may be very sensitive to discounting assumptions. Many prominent integrated assessment models illustrate these conclusions. I do not dispute their importance, but it is worth remembering that they are model-dependent. Moving from a deterministic to a random event time, in a cost-benefit setting, can weaken both conclusions, merely as a consequence of Jensen’s inequality. Using analytic models and a review of numerical results, I provide examples where policy prescriptions are more sensitive to discounting assumptions the more linear is the model.

Although most models emphasize the sensitivity of optimal policy to assumptions about discounting, there is a strand in the literature that claims that discounting is relatively unimportant when discussing potentially catastrophic events. In my view, the model that has been adduced to support this conclusion in fact has little if anything to say about the conclusion. In some circumstances, catastrophic risk does swamp discounting in determining optimal policy, and in other circumstances it does not. I do not think that we currently have a basis for thinking that either tendency is more plausible.

Most climate policy models use an infinitely lived agent. These models provide a sensible starting point, because they admit normative conclusions. However, intertemporal transfers across a single individual or between different individuals are conceptually distinct. Even if, in the interest of tractability, we want to use a constant

discount rate to evaluate both types of transfers, there is no logical reason for using the same constant to evaluate the two types of transfers. If we accept that different discount rates should be used to evaluate the two types of transfers, and if we also want to attach the same weight to the welfare of *currently* living agents, then the implicit social planner has time inconsistent preferences. In that case, the planner’s problem becomes a sequential game instead of an optimization problem, and we lose the straightforward normative implications of the model.

Appendix 1

If society pays the premium, thus eliminating risk, the present discounted stream of utility is

$$R(X; \Delta, g, \eta) \equiv \int_0^\infty e^{-\rho t} \frac{(e^{gt}(1-\Delta X))^{1-\eta} - 1}{1-\eta} dt = \int_0^\infty \frac{e^{-(\rho+(\eta-1)g)t} ((1-\Delta X))^{1-\eta} - e^{-\rho t}}{1-\eta} dt = \frac{1}{1-\eta} \left(\frac{((1-\Delta X))^{1-\eta}}{(\rho+(\eta-1)g)} - \frac{1}{\rho} \right)$$

If society does not pay the insurance premium, and the event occurs at time T , the payoff is

$$\begin{aligned} P(T; \Delta, g, \eta) &\equiv \int_0^T e^{-\rho t} \frac{(e^{gt})^{1-\eta} - 1}{1-\eta} dt + \int_T^\infty e^{-\rho t} \frac{(e^{gt}(1-\Delta))^{1-\eta} - 1}{1-\eta} dt \\ &= \int_0^T e^{-\rho t} \frac{((e^{gt})^{1-\eta} - 1 - ((e^{gt}(1-\Delta))^{1-\eta} - 1))}{1-\eta} dt + \int_0^\infty e^{-\rho t} \frac{(e^{gt}(1-\Delta))^{1-\eta} - 1}{1-\eta} dt \\ &= \int_0^T e^{-\rho t} \frac{e^{-g(\eta-1)t} (1 - (1-\Delta)^{1-\eta})}{1-\eta} dt + \int_0^\infty e^{-\rho t} \frac{(e^{gt}(1-\Delta))^{1-\eta} - 1}{1-\eta} dt \\ &= \frac{(1 - (1-\Delta)^{1-\eta})}{1-\eta} \frac{1 - e^{-(\rho+(\eta-1)g)T}}{(\rho+(\eta-1)g)} + \frac{1}{1-\eta} \left(\frac{(1-\Delta)^{1-\eta}}{(\rho+(\eta-1)g)} - \frac{1}{\rho} \right) \end{aligned}$$

If the event time T is exponentially distributed with hazard h , then

$$E e^{-(\rho+(\eta-1)g)T} = \int_0^\infty e^{-(\rho+(\eta-1)g)t} h e^{-ht} dt = \frac{h}{(\rho + g\eta - g + h)}. \tag{7}$$

Using this formula, we have

$$EP(T; \Delta, g, \eta) = \frac{(1 - (1-\Delta)^{1-\eta})}{1-\eta} \frac{1 - \frac{h}{(\rho+g\eta-g+h)}}{(\rho + (\eta - 1)g)} + \frac{1}{1-\eta} \left(\frac{(1-\Delta)^{1-\eta}}{(\rho + (\eta - 1)g)} - \frac{1}{\rho} \right)$$

For the certain event time, with $T = \frac{1}{h}$, the maximum premium society would pay, X , is the solution to

$$R(X; \Delta, g, \eta) = P(T; \Delta, g, \eta)$$

or

$$\frac{1}{1-\eta} \left(\frac{((1-\Delta X))^{1-\eta}}{(\rho+(\eta-1)g)} - \frac{1}{\rho} \right) = \frac{(1-(1-\Delta))^{1-\eta}}{1-\eta} \frac{1-e^{-(\rho+(\eta-1)g)T}}{(\rho+(\eta-1)g)} + \frac{1}{1-\eta} \left(\frac{(1-\Delta)^{1-\eta}}{(\rho+(\eta-1)g)} - \frac{1}{\rho} \right).$$

Solving for X

$$\left(\frac{((1-\Delta X))^{1-\eta}}{(\rho+(\eta-1)g)} \right) = (1 - (1 - \Delta)^{1-\eta}) \frac{1-e^{-(\rho+(\eta-1)g)T}}{(\rho+(\eta-1)g)} + \left(\frac{(1-\Delta)^{1-\eta}}{(\rho+(\eta-1)g)} \right) \Rightarrow$$

$$(1 - \Delta X) = \left((1 - (1 - \Delta)^{1-\eta}) \left(1 - e^{-\frac{(\rho+(\eta-1)g)}{h}} \right) + (1 - \Delta)^{1-\eta} \right)^{\frac{1}{1-\eta}} \Rightarrow$$

Solving for X and setting $y = 100X$ gives Eq. 1. I obtain Eq. 2 by taking expectations with respect to T , using Eq. 7, and repeating the steps used to obtain Eq. 1 (Figs. 11 and 12).

Fig. 11 Maximum premium (as a percent of cost of event) as a function of η , under certain event time (*solid*) and random event time (*dashed*) for $T = 200 = \frac{1}{h}$, $\Delta = 0.4$, $g = 0.01$ and $\rho = 0.01$

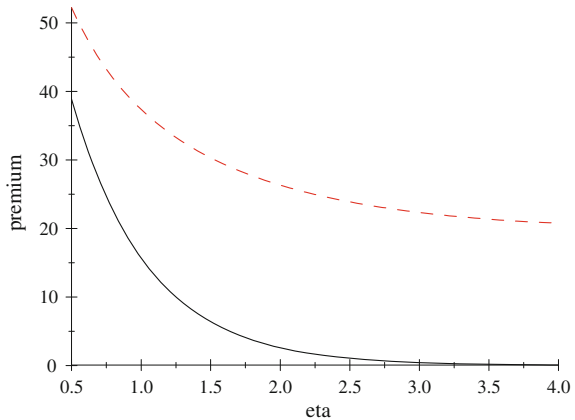
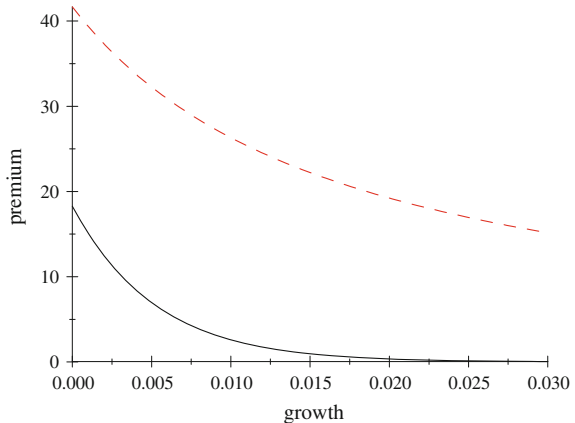


Fig. 12 Maximum premium (as a percent of cost of event) as a function of growth, under certain event time (*solid*) and random event time (*dashed*) for $T = 200 = \frac{1}{h}$, $\Delta = 0.4$, $\eta = 2$ and $\rho = 0.01$



References

- Arrow, K. (2007). Globak climate change: A challange to policy. *The Economists Voice*, 4, 1–5.
- Barro, R. (1999). Ramsey meets Laibson in the neoclassical growth model. *Quarterly Journal of Economics*, 114, 1125–52.
- Calvo, G., & Obstfeld, M. (1988). Optimal time-consistent fiscal policy with finite lifetimes. *Econometrica*, 56, 32–411.
- Clark, C. W. (1975). *Mathematical bioeconomics: The optimal management of renewable resources*. Wiley.
- Cropper, M., & Laibson, D. (1999). The implications of hyperbolic discounting for project evaluation. In P. R. Portney & J. Weyant (Eds.), *Discounting and Intergenerational Equity*. Resources for the Future.
- Ekeland, I., Karp, L., & Sumaila, R. (2012). *Equilibrium management of fisheries with altruistic overlapping generations*. University of British Columbia Working Paper.
- Ekeland, I., & Lazrak, A. (2010). The golden rule when preferences are time inconsistent. *Mathematical and Financial Economics*, 4(1), 29–55.
- Fujii, T., & Karp, L. (2008). Numerical analysis of non-constant pure rate of time preference: a model of climate policy. *Journal of Environmental Economics and Management*, Forthcoming, 56, 83–101.
- Fujii, T. (2008). Numerical analysis of non-constant pure rate of time preference: a model of climate policy. *Journal of Environmental Economics and Management*, 56, 83–101.
- Gerlagh, R., & Liski, M. (2012). *Carbon prices for the next thousand years*. Unpublished Working Paper.
- Harris, C., & Laibson, D. (2001). Dynamic choices of hyperbolic consumers. *Econometrica*, 69(5), 935–957.
- Heal, G. (1998). *Valuing the future: economic theory and sustainability*. New York: Columbia University Press.
- Heal, G. (2001). Intertemporal welfare economics and the environment. In K.-G. Maler & J. Vincent (Eds.), *The handbook of environmental economics*. Amsterdam: North Holland.
- Horowitz, J., & Lange, A. (2008). What's wrong with infinity—a note on Weitzman's dismal theorem. University of Maryland Working Paper.
- Karp, L. (2005). Global Warming and hyperbolic discounting. *Journal of Public Economics*, 89, 261–282.
- Karp, L. (2009). Sacrifice, discounting and climate policy: five questions. Giannini Foundation working paper.
- Karp, L. (2013a). Overlapping generations and environmental policy: An Introduction. Giannini Foundation Working Paper.
- Karp, L. (2013b). *Provision of a public good with altruistic overlapping generations and many tribes*. Giannini Foundation Working Paper.
- Karp, L. (2013c). *Railroad discounting*. Giannini Foundation Working Paper.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 62, 443–78.
- Nordhaus, W. D. (2007). A Review of the stern review on the economics of climate change. *Journal of Economic Literature*, 3, 686–702.
- Norhaus, W. (2009). An analysis of the dismal theorem. Cowles Foundatiuon Discussion Paper No 1686.
- Phelps, E., & Pollack, R. (1968). On second-best national savings and game: equilibrium growth. *Review of Economic Studies*, 35, 185–199.

- Ramsey, F. (1928). A mathematical theory of savings. *Economic Journal*, 38, 59–543.
- Schneider, M., Traeger, C., & Winkler, R. (2012). Trading off generations: Equity, efficiency and climate change. *European Economic Review*, 56, 1621–1644.
- Stern, N. (2007). *The economics of climate change*. Cambridge University Press.
- Weitzman, M. (2009). On modeling and interpreting the economics of catastrophic climate change. *Review of Economics and Statistics*, 91(1), 1–19.