

# Should the climate tail wag the policy dog?

Gerard H. Roe

*Department of Earth and Space Sciences,*

*University of Washington, Seattle, WA.*

and

Yoram Bauman

*Program on the Environment,*

*University of Washington, Seattle, WA.*

and *Institute of Global Low-Carbon Economy,*

*University of International Business and Economics, Beijing*

July 18, 2011

# Abstract

The small but stubbornly unyielding possibility of a very large long-term response of global temperature to increases in atmospheric carbon dioxide can be termed the fat tail of high climate sensitivity. Recent economic analyses suggest that the fat tail should dominate a rational policy strategy if the damages associated with such high temperatures are large enough. The conclusions of such analyses, however, depend on how economic growth, temperature changes, and climate damages unfold and interact over time. In this paper we focus on the role of two robust physical properties of the climate system: the enormous thermal inertia of the ocean, and the long timescales associated with high climate sensitivity. Economic models that include a climate component, and particularly those that focus on the tails of the probability distributions, should properly represent the physics of this slow response to high climate sensitivity, including the correlated uncertainty between present forcing and climate sensitivity, and the global energetics of the present climate state. If climate sensitivity in fact proves to be high, these considerations prevent the high temperatures in the fat tail from being reached for many centuries. A failure to include these factors risks distorting the resulting economic analyses. For example, we conclude that fat-tail considerations will not strongly influence economic analyses when these analyses follow the common—albeit controversial—practices of assigning large damages only to outcomes with very high temperature changes and of assuming a significant baseline level of economic growth.

Keywords: climate change; climate sensitivity; fat tail; policy; discount rate

# 1 The fat tail of climate sensitivity

Climate sensitivity—the long-term response of global-mean, annual-mean surface temperature to a doubling of carbon dioxide above pre-industrial values—has long been a benchmark by which to compare different estimates of the planet’s climatic response to changes in radiative forcing. A doubling of carbon dioxide increases the radiative forcing by about  $\Delta R_{2\times} \approx 4 \text{ Wm}^{-2}$ . In a remarkable piece of analysis, Svante Arrhenius (1896) made the first quantitative estimate that, in response, the equilibrium global-mean temperature would increase by  $\Delta T_{2\times} \approx 5 \text{ }^\circ\text{C}$ . A major reassessment came in 1979 with the National Research Council Charney Report. Reviewing the intervening advances in science, Charney (1979) estimated  $\Delta T_{2\times}$  between 1.5 and 4.5  $^\circ\text{C}$  (described as the ‘probable error’), a range that has changed only incrementally ever since.

Modern estimates provide an answer in terms of more precise probability distributions, or probability density functions (PDFs). The Intergovernmental Panel on Climate Change (IPCC, 2007) report gives a ‘likely’ range (2-in-3 chance) of  $\Delta T_{2\times}$  lying between 2 and 4.5  $^\circ\text{C}$ , and concludes it is ‘very unlikely’ ( $< 1\text{-in-}10$ ) to be less than 1.5  $^\circ\text{C}$ . This summary is consistent with a peculiarity of a large number of other studies, namely the small but non-trivial possibility of  $\Delta T_{2\times}$  being very much larger than the canonical 1.5 to 4.5  $^\circ\text{C}$  range. One of the important achievements in recent climate science has been to establish great confidence in the lower bound on climate sensitivity, but the upper bound has proven much less tractable (e.g., Knutti and Hegerl, 2008). Allen et al. (2006) have argued that this fat tail of possibly-high climate sensitivity is a fundamental feature of  $\Delta T_{2\times}$  estimates from observations, and Roe and Baker (2007) have argued that it is a fundamental feature of  $\Delta T_{2\times}$  estimates from numerical climate models. It will require improbably large reductions in uncertainties about the radiative forcing the planet has experienced—or, equivalently, in our uncertainty about physical feedbacks in the climate system—to substantially remove the skewness.

Some studies (e.g., Annan and Hargreaves, 2006) have combined multiple estimates in a Bayesian framework and argued this can yield narrower and less skewed distributions for  $\Delta T_{2\times}$ . However the answers derived from such an approach are critically sensitive to how independent the different estimates are, and what the Bayesian prior assumptions are. Both factors are fiendishly elusive to pin down objectively. Knutti and Hegerl (2008) and Knutti et al. (2010) provide a good review of the scientific issues involved. Efforts continue in this direction (Annan and Hargreaves, 2009) but, overall, it seems prudent to assume that estimates of  $\Delta T_{2\times}$  will not change substantially for the foreseeable future.

Several macroeconomic analyses of the costs of climate change, and of the correct willingness-to-pay to avoid it, have argued that there is a sting in the fat tail of  $\Delta T_{2\times}$  (e.g., Weitzman, 2009a,b, 2010). This is seen as follows: if, as is reasonable to assume, climate damages increase nonlinearly with temperature, then the tail of the PDF is weighted more strongly than the middle of the PDF in calculations of the expected damages. Moreover, if damages are highly nonlinear with temperature, then even the very smallest chance of absolutely apocalyptic consequences might properly dominate a rational policy. A close analogy is the St. Petersburg paradox, a coin-flipping wager with an expected value dominated by low-probability, high-valued outcomes. The extent to which this “fat-tail” argument is a worry depends in part on how rapidly the global mean temperature might rise towards its equilibrium value. This is both because impacts and costs of adaptation depend on the rapidity of change and because economic analyses implicitly discount future damages. The message in this paper is not that high temperatures cannot be reached on timescales that can impact society—if the system is forced strongly enough they are guaranteed; nor is it that the damages associated with climate change cannot be high—one can always pick a damage function that will achieve that. Rather it is to emphasize that such high temperatures and

damages would not primarily be due to the fat tail of climate sensitivity. This fact depends only on the physics of the climate system and is therefore independent of the particular choice of economic model.

## **2 The transient evolution of possible future climates**

A key player in the physical system is the enormous thermal inertia represented by the deep ocean. The whole climate system cannot reach a new equilibrium until the deep ocean has also reached equilibrium. In response to a positive climate forcing (i.e., a warming tendency), the deep ocean draws heat away from the surface ocean, and so buffers the surface temperature changes, making them less than they would otherwise be. The deep ocean is capable of absorbing enormous amounts of heat, and not until this reservoir has been exhausted can the surface temperatures attain their full, equilibrium values.

A second key player is the inherent relationship between feedbacks and adjustment time scales in physical systems. If it transpires that we do, in fact, live on a planet with a high climate sensitivity, it will be because we live on a planet with strong positive feedbacks. In other words, the net effect of all of the dynamic processes (clouds, water vapor, ice reflectivity, etc.) is to strongly amplify the planet's response to radiative forcing. In this event, it would mean that we live on a planet that is inefficient in eliminating energy perturbations: a positive feedback reflects a tendency to retain energy within the system, inhibiting its ultimate emission to space, and therefore requiring a larger temperature response in order to achieve energy equilibrium. Moreover, it is generally true that, all else being equal, an inefficient system takes longer to adjust than an efficient one. A useful rule-of-thumb is that the relevant response time of the climate system is given by the effective thermal

inertia of the deep ocean multiplied by the climate sensitivity parameter (defined as  $\Delta T_{2\times}/\Delta R_{2\times}$ , see, e.g., Roe, 2009). This behavior is absolutely fundamental and widely appreciated (e.g., Hansen et al., 1985; Wigley and Schlesinger, 1985). As time progresses, more and more of the ocean abyssal waters become involved in the warming, and so the effective thermal inertia of the climate system increases. Hansen et al. (1985) solve a simple representation of this effect and show that the adjustment time of climate is proportional to the square of climate sensitivity. In other words, if it takes 50 yrs to equilibrate with a climate sensitivity of 1.5 °C, it would take 100 times longer, or 5,000 yrs, to equilibrate if the climate sensitivity is 15°C. Although Nature is of course more complicated than this (see, e.g., Gregory, 2000), the basic picture described here is reproduced in models with a more realistic ocean circulation. In particular see Held et al. (2010) for results from fully-coupled global climate models.

In the context of the PDF of climate sensitivity, the relationship with the climate response time has been reviewed in Baker and Roe (2009), whose climate model we use in the present study, and which represents both atmospheric feedbacks and the uptake of heat by the upper and deep ocean. A schematic illustration of the model is shown in Figure 1, and is very briefly described here. The evolution of global-mean, annual-mean surface temperature, as represented by the ocean mixed layer, is governed by:

$$\rho C_p h \frac{dT}{dt} + \frac{T(1 - f_a)}{\lambda_0} - \kappa \frac{\partial T_{th}}{\partial z} \Big|_{z=0} = \Delta R_F(t), \quad (1)$$

where  $C$  is the heat capacity of water,  $h$  is the mixed-layer depth,  $f_a$  are the total atmospheric feedbacks.  $\kappa$  is the thermal conductivity governing the heat flux across the thermocline (at  $z = 0$ ), which also depends on the vertical temperature gradient. Temperature in the deep ocean,  $T_{th}$ , is

governed by a balance between diffusion and upwelling:

$$\frac{\partial T_{th}}{\partial t} = \chi \frac{\partial^2 T_{th}}{\partial z^2} - w \frac{\partial T_{th}}{\partial z}, \quad (2)$$

where  $w$  is the upwelling rate and  $\chi$  is the diffusivity. The model is virtually identical to a host of other equivalent models, which have been shown to be fully capable of emulating more complete numerical models, and also historical observations at the global scale (e.g., IPCC, 2007). Such models are regularly used to make long-term climate predictions. Their flexibility and computational efficiency means that parameters can be easily varied and that uncertainties can be fully explored. The equations and parameters follow those of Baker and Roe (2009), which matches best current estimates of heat uptake and climate feedbacks. The only change from Baker and Roe (2009) is a  $\chi$  of  $2.5 \times 10^{-4} \text{m}^2 \text{s}^{-1}$  instead of  $2.0 \times 10^{-4} \text{m}^2 \text{s}^{-1}$ , as it was found to produce slightly better agreement for historical trends. Results are insensitive to plausible variations in these climate parameters.

## 2.1 A numerical example: a doubling of CO<sub>2</sub>

We begin by presenting the climate response to an idealized, instantaneous doubling of carbon dioxide. In such a scenario the envelope of possible temperature responses must ultimately evolve in time into the equilibrium climate sensitivity distribution. The reasons for our initial focus on this scenario are two-fold. Firstly, climate sensitivity has been identified as a key uncertainty in integrated assessment models (IAMs) (e.g., IAG 2010) and this idealized step-function forcing is a particularly clear illustration of the long timescales involved in the development of the fat tail of climate sensitivity. This clarity can be obscured when looking at more complicated emissions scenarios. Secondly, it has been used itself as a scenario in some economic evaluations (e.g., Weitzman,

2010). In a subsequent section we extend the analysis to a more realistic forcing scenario.

The dramatically long timescale for the development of the fat tail is shown in Figure 2. It presents—for an instantaneous doubling of carbon dioxide—the evolution in time of the envelope of possible climate trajectories of global-mean temperature. Note the change to a logarithmic time axis after 500 years. The shading in Figure 2 represents the one-, two-, and three-standard deviation ranges for climate sensitivity, encompassing 68.3, 95.5, and 99.7% of possibilities, respectively.

The uncertainty ranges used are fully consistent with observed estimates of climate sensitivity (e.g., Allen et al., 2006), with the multi-thousand model experiments of the innovative *climateprediction.net* program (Stainforth et al., 2005), and with the IPCC ranges of uncertainty in  $\Delta T_{2\times}$  (e.g.,  $\sim 75\%$  of possible  $\Delta T_{2\times}$ s lie at less than  $\sim 4.5$  °C). Reflecting uncertainties in observations, uncertainties in feedbacks are Gaussian distributed (e.g., Roe and Armour, 2011). This leads to uncertainties in the evolving climate response that are also Gaussian initially, but which become highly skewed over time (e.g., Baker and Roe, 2009). At the extreme limit of the range we consider, values of climate sensitivity at  $3\sigma$  would actually imply that the climate system is slightly unstable, in other words a runaway greenhouse effect.<sup>1</sup> Figure 2 therefore illustrates that even a planet that is formally headed to oblivion can take a very long time to get there because of the ocean’s capacity to absorb heat. The consensus of the climate science community would be that the likelihood we are actually on a runaway greenhouse trajectory is vanishingly small (e.g., Solomon et al., 2010), although the objective basis for this belief is subtler and less well-established than is commonly appreciated. It is nonetheless true that almost all climate scientists would view the uncertainties in Figure 2 to represent an overestimate. In other words, we are effectively considering a maximum

---

<sup>1</sup>We use a model estimate for the climate feedback parameter of  $f_a = 0.65$ , and a standard deviation of  $\sigma = 0.13$ , meaning the  $3\sigma$  value is  $f_a = 1.04$ . Values of  $f > 1$  means the system is formally unstable. See Baker and Roe (2009) for details.



possible upper bound on the probabilities residing within the fat tail. Because of these physical limits, we truncate the PDF of feedbacks beyond this  $\pm 3\sigma$  range. Climate feedbacks exceeding the upper bound of this range should not be seriously entertained, and climate feedbacks less than this lower bound make no impact of any consequence on the analyses.

Let  $h(\Delta T(t))$  stand for the PDF of possible future global mean temperatures,  $\Delta T(t)$ , as they evolve over time. The abiding impression from Figure 2 is that trajectories at the low end of  $h(\Delta T)$  rapidly adjust to their equilibrium values over a few decades or a century, whereas those at the high end take thousand of years even to approach their equilibrium values. For the reasons given above, climate trajectories stay relatively tightly bunched over the course of the first few centuries, diverging only slowly thereafter, and all of that divergence occurs toward higher temperatures. Baker and Roe (2009) show, at any given time  $t$ , that the flux (i.e., growth) of probabilities past a given temperature,  $\Delta T_1$ , is given by the product of  $h(\Delta T_1)$  and the rate of warming on the trajectory passing through  $\Delta T_1$  at time  $t$ . In this light, the fat tail is walloped twice: firstly,  $h(\Delta T_1)$  is small if it is in the tail; secondly, the rate of warming is generally declining with time (e.g., Figure 2, and recalling that the time axis becomes logarithmic after year 500). Therefore the fat tail can fill up only very gradually.

Time slices of the envelope of  $h(\Delta T)$  are shown in Figure 3, together with their skewness. At  $t = 0$  the shape of  $h(\Delta T)$  is actually Gaussian; it acquires skewness only gradually over time (e.g., Baker and Roe, 2009). At the theoretical equilibrium ( $t = \infty$ ), the skewness of the distribution is infinite, but even 1000 years after doubling  $\text{CO}_2$ , the skewness of the PDF remains less than two. An important point to note is that, for any finite time, there is an upper limit on the temperature set by how much energy has been accumulated within the system to that time, which can be found by integrating Equations (1) and (2). For the highest value of  $f_a$  that we consider, 1.04 (which as

already noted would imply infinite climate sensitivity, and a slightly unstable climate system), this upper limit is the upper edge of the outermost shaded regions in Figures 2a, 4a, and 5a. Therefore, by the formal statistical definition (that the PDF asymptotes as a power law in the limit of  $T \rightarrow \infty$ ) there can be, in fact, no fat tail to the temperature distribution at any finite time.

### 3 Economic implications

How does the evolution of the climate tail affect calculations of the cost of damages associated with climate change? The analysis that follows draws heavily on Weitzman (2010), which provides a simple and incisive framework for calculating the costs of avoiding or insuring against climate change. In order to explore the significance of the fat tail, Weitzman suggests comparing two different “climate damages” functions. The first is a quadratic function of global mean temperature:

$$C_Q = \frac{1}{1 + \alpha \Delta T^2}, \quad (3)$$

$C_Q$  represents the “welfare equivalent” consumption as a fraction of what consumption would have been were  $\Delta T = 0$ . Weitzman takes the rather precise value of  $\alpha = 2.388 \times 10^{-3}$  in order to match the damage function of Nordhaus (2008), for which  $C_Q(2^\circ\text{C}) \approx 99\%$ . He also considers a “reactive” damages function with an extra term:

$$C_R = \frac{1}{1 + \alpha \Delta T^2 + \beta_2 \Delta T^\gamma}. \quad (4)$$

By choosing  $\beta = 5.075 \times 10^{-6}$  and  $\gamma = 6.754$ , Weitzman thus creates a damages function that has

significant consequences for higher global-mean temperatures ( $C_R(6^\circ\text{C}) = 50\%$ , and  $C_R(12^\circ\text{C}) = 1\%$ ). The functional choices and parameters are chosen purely for illustrative purposes.<sup>2</sup>

What should be our “willingness to pay” to avoid the climate damages associated with either  $C_Q$  or  $C_R$ , or to insure against them? Weitzman defines utility, or welfare, as a simple function of consumption,  $U(C)$ , and also assumes an underlying long-term exponential economic growth rate,  $g$ . In a world experiencing climate change, consumption accelerates because of the long-term growth rate but is also retarded by the climate damages that accompany warming. Thus for every possible trajectory of future temperature,  $\Delta T(t)$ , there are accompanying trajectories for consumption,  $C(t)$ , and welfare,  $U(t)$ .

Weitzman then compares this world to one in which consumption is reduced by a constant fraction,  $\hat{C}$ , with the remainder potentially used to avoid or insure against climate changes. In this world, this fractionally reduced consumption still grows at rate  $g$ , but does not suffer climate damages.  $\hat{C}$  can be found by calculating the integrated welfare of these two worlds up to a decision horizon ( $\equiv t_h$ , the duration of time into future that is relevant for decision making), and equating them:

$$\int_0^{t_h} U(\hat{C}e^{gt})dt = E \left[ \int_0^{t_h} U(C_{Q,R}(\Delta T(t))e^{gt})dt \right], \quad (5)$$

where  $E[\ ]$  denotes the statistical expectation operator. In principle then,  $1 - \hat{C}$  ought to be the upper bound on our willingness to pay, in terms of a fraction of consumption, in order to avoid or insure against climate damages. Following reasonably standard practices, Weitzman sets  $U(C) = C^{1-\eta}/(1-\eta)$ , and chooses  $\eta = 3$  and  $g = 2\% \text{ yr}^{-1}$ . The rate of pure time preference,  $\rho$ , is

---

<sup>2</sup>Note that in all cases  $\Delta T$  is measured in  $^\circ\text{C}$ , and so the units of  $\alpha$  and  $\beta$  depend on the choice of exponents.

This could be avoided by dividing  $T$  by a normalizing temperature.

set to zero, meaning that the welfare of future generations are given equal weight to the current. In an optimizing framework, the effect of these choices is that future reductions in consumption are discounted at an exponential rate of  $\eta g = 6\% \text{ yr}^{-1}$  because future generations are richer than the current generation and therefore suffer a smaller loss in utility from reduced consumption at the margin.

### 3.1 Comments on the economics model

The focus of this paper is on the physical constraints that affect the transient evolution of climate. These physical aspects of the climate evolution are obviously independent of the economic model chosen, and our use of Weitzman's framework is for its convenience and clear linkage to the PDFs of the climate response. Though welfare under climate change is compared to a no-loss case assuming a constant growth rate, the analysis can be cast in terms of an effective growth rate that is affected by climate. In a conventional formulation, an exponentially growing consumption has the form  $dC/dt = \tilde{g}C$ , where the growth rate  $\tilde{g}$  may vary with climate. In the Weitzman framework consumption is modified by climate damages:  $C(t) = C_{R,Q}(T(t))\exp(gt)$ . Combining these two gives:

$$\tilde{g}(t) = g + \frac{1}{C} \frac{dC}{dt} = g + \frac{1}{C} \frac{dC}{dT} \frac{dT}{dt} \quad (6)$$

Since  $dC/dT$  is negative, the effective growth rate  $\tilde{g}$  is diminished relative to  $g$ . Thus, though not explicitly parameterized, the effective growth rate can be regarded as being endogenously determined within the model. The implications of our results for more complex economic models are addressed in the discussion.

### 3.2 A numerical example: the impact of a CO<sub>2</sub> doubling

Section 2 reviewed the strong physical constraints on how  $h(\Delta T(t))$  evolves with time. From Equations (3) to (5) the evolving PDFs of  $C_{Q,R}$  can be calculated, and are shown in Figure 4b and d. Because the median of  $h(\Delta T)$  remains below 3°C over the whole millennium, there is little difference between the medians of  $C_Q$  and  $C_R$ . At the far end of the fat tail of possibilities, the  $3\sigma$  trajectory reaches 6°C at around 250 years which, were it to transpire, would cause both  $C_Q$  and  $C_R$  to be a significant fraction of consumption (for instance,  $C_R(3\sigma) = 47\%$  at year 250).

When cast in terms of welfare-equivalent consumption (i.e.,  $\hat{C}_{Q,R}$ ), we reach two important conclusions. Firstly, we see from Figure 4c and e that there is only a small difference between  $\hat{C}_Q$  and  $\hat{C}_R$ . Because of the exponential discounting of the future, both  $\hat{C}_Q$  and  $\hat{C}_R$  asymptote to a constant value, and change little for decision horizons beyond 150 years (reflecting the exponential decay of the weighting given to future damages).

The second and arguably more important conclusion is that willingness-to-pay is under 1% for even the most extreme climate possibilities. This is largely because long periods of time are required to reach high temperatures regardless of the long-run climate sensitivity. Saying that  $C_R(3\sigma) = 47\%$  at year 250 sounds impressive, but with consumption growing at 2% per year this simply means that climate change will reduce consumption from 141 times current consumption to only 53 times current consumption. Under the assumption of diminishing marginal utility of consumption it is not surprising that intertemporal optimization comes to the conclusion that current generations should be unwilling to sacrifice more than 1% of consumption for the sake of future millionaires.

## 4 A more realistic scenario

Up to this point we’ve used a doubling of  $\text{CO}_2$  as an illustrative climate forcing because of its long history as a benchmark measure of climate change, and because of its relationship to climate sensitivity. We next consider a more realistic forcing scenario that, while still idealized, captures important aspects of the real situation: 1) past climate forcing is approximately consistent with past observations and their uncertainty, 2) future forcing is based on a do-nothing, business-as-usual scenario for the 21st century, similar to the IPCC A2 scenario, 3) economic analyses and climate damages are calculated using the present day (i.e., 2011) as a baseline, and 4) the climate model is correctly initialized for the current state of the climate system, accounting for the correlated uncertainty in forcing and climate sensitivity. This last aspect, which is omitted in all IAMs that we are aware of, contributes to how soon high temperatures can be realized in the future. Again, the crucial result is that the fat tail of climate sensitivity does not play a strong role in climate projections on economically relevant time scales.

In order to describe this more realistic scenario, we must first detour briefly into the causes of uncertainty in predictions of future climate. The main points of this paragraph are summarized in Roe (2010) and Armour and Roe (2011), and are based on, and entirely consistent with, the IPCC AR4. Future climate is uncertain because of two principal factors: 1) uncertainty in future emissions, and 2) uncertainty in how the climate will respond to those emissions. For the latter, the community uses climate models that have been calibrated against past observations of global climate which, in turn, must satisfy the observed global energy balance. Many parts of this global energy balance are well constrained, including the ocean heat uptake, the global-mean temperature change, and the radiative forcing due to increases in  $\text{CO}_2$  and other greenhouse gases. The two largest sources of uncertainty are the aerosol radiative forcing and the climate sensitivity. Anthropogenic aerosols—

airborne particles, either solid or liquid—have a direct impact on the Earth’s radiative budget and also an indirect impact via their influence on clouds. The effect of aerosols is almost certainly negative, but the uncertainty is large: IPCC AR4 states the aerosol forcing as  $\sim -1.3 \pm 0.5(1\sigma)\text{Wm}^{-2}$  for the change since the pre-industrial. This negative forcing masks a significant portion of the positive forcing from the better constrained greenhouse gases ( $\sim 2.9 \pm 0.2(1\sigma)\text{Wm}^{-2}$ ). The key point is that uncertainties in climate sensitivity and climate forcing are inherently equivalent: if aerosol forcing has been strongly negative, the net climate forcing has been weak, and therefore the climate sensitivity must be high in order for us to have experienced the warming we have; on the other hand, if aerosol forcing has been weak, net forcing has been more strongly positive and a low value of climate sensitivity is implied. Thus climate forcing and climate sensitivity must be treated as complementary pairs: if one is high the other is correspondingly low.

Figure 5a shows a linear approximation to past forcing from 1850 to the present, including forcing uncertainty. The net climate forcing for the present day (radiative forcing minus ocean heat uptake) is assumed to be Gaussian distributed. For each possible forcing trajectory, a complementary climate sensitivity is chosen (or equivalently,  $f_a$  in Equation 1) such that the climate model reproduces the well-constrained temperature history from 1850 to the present (Figure 5b). Between 2011 and 2100, the climate forcing is chosen to emulate the IPCC AR2 scenario, which implies a strong increase to about  $8 \text{ Wm}^{-2}$  by 2100. Thereafter a decline in forcing is assumed at a rate which approximately stabilizes the temperature for the median trajectory. Calculations are continued through to 2300. Uncertainty in aerosol forcing is assumed to be constant though, in principle, future improvements in understanding would enable us to narrow the uncertainty in forcing and hence in the accompanying climate sensitivity. Figure 5b shows the temperature response to this scenario. Consistent with the relationship between climate sensitivity and adjustment time, tra-

jectories with low climate sensitivity respond rapidly to changes in forcing, and those with high sensitivity respond much more slowly. The median temperature is 4°C at 2100, but uncertainty in climate sensitivity leads to a diverging envelope of possible future temperatures after 2011, with a  $3\sigma$  range of about 3 to 8°C in 2100.

Should the temperatures towards the upper part of this range be considered as part of a “fat tail” of the climate response? Strictly speaking, the term fat tail has a formal statistical definition, which we’ve already shown cannot apply at any finite time. However the term is also used more colloquially as short-hand for the possibility of very large temperature changes to anthropogenic forcing. In this regard it is very important to clearly distinguish two aspects of the problem. As already noted “climate sensitivity” is defined as the equilibrium response to the (comparatively weak) climate forcing of a CO<sub>2</sub> doubling. For such a scenario it is reasonable, for instance, to refer to temperatures above the IPCC “likely” range (i.e., > 4.5°C) as lying in the “fat tail”. In contrast, the business-as-usual scenarios that lead to the possibility of similarly high temperatures by 2100 in Figure 5b are very different: they are the near-term response to a very different and much larger climate forcing. Furthermore, as can be seen in Figure 5b and Figure 6, the PDFs applying to such scenarios are not particularly skewed. For the term “fat tail” to retain its value, it should be reserved for, at minimum, much more strongly skewed PDFs.

In terms of climate damages, the quadratic damages function,  $C_Q$ , grows relatively slowly. The  $3\sigma$  trajectory of climate damages reaches 20% of consumption (i.e.,  $C_Q = 0.8$ ) at around 2120. This slow growth and the exponential discounting of future welfare produces a  $\hat{C}_Q$  of 99.7%, or in other words a 0.3% willingness to pay. For the reactive damages function, damages obviously grow much quicker, with the  $3\sigma$  trajectory reaching 20% damages forty years earlier, at about 2080. The possibility of continued warming after 2100 (despite declining forcing) occurs along



high sensitivity trajectories because of a continuing energy imbalance, and damages grow commensurately. These large damages are given strong weighting in the welfare function, and so even though the median estimate for climate damages never exceed about 10% (i.e.,  $C_Q = 0.9$ ) the expected welfare-equivalent consumption is 60%, or a willingness-to-pay of around 40% of present consumption, which is obviously large (though not unbounded).

It is important to drive home the point that such large willingness-to-pay does not come from the fat tail of high climate sensitivity, which cannot physically be realized on these timescales. Rather, it comes from the combination of a strongly forced climate system where high temperatures are inevitable, and the choice of a highly nonlinear damages function. This is illustrated in Figure 6, which plots time slices of the evolving PDFs on a log scale to highlight the behavior in the tails. The skewness of the temperature distribution reaches about 3 at 2300 (recall the skewness would be infinite at steady-state). The PDF of quadratic damage function,  $C_Q$ , amplifies this skewness a little. The PDF of the reactive damage function,  $C_R$ , is controlled by its highly nonlinear functional form: initially its PDF is much more skewed than that of temperature, and it actually gets less skewed towards the end of the calculation, because probabilities are beginning to pile up near  $C_R = 0$  (Figure 6b).

#### **4.1 The importance of correctly initializing the climate state for economic calculations**

Economic analyses for policy purposes typically take the present day as the baseline for decision making. In other words, choices made now are based on how things might change into the future, and not by the “sunk cost” changes that have already occurred. The purpose of this short section is to highlight that, in projecting possible future climate changes, it is very important that economic

models are correctly initialized with the current climate state and its uncertainties. The planet's climate change experiment is already more than a hundred years old, and two factors are of particular significance: 1) the ocean warming lags behind the surface, and the resulting thermodynamic imbalance is driving heat into the ocean, and 2) as noted above, we are uncertain about climate sensitivity because we are uncertain about the current energy balance of the planet.

We now demonstrate the effect of incorrectly initializing economic models by repeating the calculations shown in Figure 5, but neglecting to represent these two factors. The effect is to cause the range of possible future temperatures to spread out more quickly than should be the case. The reasons are as follows. Firstly, in the present-day climate state heat is currently being drawn into the ocean because it has yet to come into balance with the current forcing. To do so would, in fact, take centuries to millennia. If this imbalance is ignored in setting up the initial conditions of a model integration starting at the present day, it is equivalent to assuming that this ocean adjustment has already happened, meaning that heat which should be going into the ocean is instead available to heat the atmosphere. This leads to more rapid warming than should be the case. Secondly, and most importantly, is the fact that if climate sensitivity is high then climate forcing has been low. If this complementary pairing is neglected, it means that all trajectories (representing all possible climate sensitivities) are given the same starting point for present-day climate forcing. Hence, higher temperatures on high sensitivity trajectories are realized more quickly than is physically possible.

Figure 7 shows the effect of neglecting these two factors. For the calculation using  $C_Q$ , there is an increase of about 25% in willingness to pay (i.e.,  $1-\hat{C}_Q$ ), though it is still small, at around 0.5%. There is an even bigger impact for  $\hat{C}_R$ . The  $3\sigma$  trajectory reaches 20% damages at 2050 (about thirty years earlier than it should), and willingness-to-pay increases from 40% to 80%. These particular numbers are not the main point, but rather it is to illustrate that when the focus is on the

high damages that might accompany the possibility of large temperature changes, it is important to correctly capture the physics of how such high temperatures might be attained.

Furthermore, in fully-coupled global climate models there is a complex spatial pattern to how quickly different regions warm, and to how quickly different regions come into equilibrium with the climate forcing. Since economic analyses often amalgamate regional assessment into a global damages function (e.g., Nordhaus, 2008), IAMs ought to consider the impact of these differential regional warming rates. As of now, they do not do so.

## 5 Discussion

The primary point of this study is to emphasize that important geophysical constraints exist that have an important impact on economic analyses of climate change policy. While the equilibrium climate sensitivity is characterized by a highly skewed uncertainty distribution (i.e., a fat tail), if it does transpire that we live on a planet with an extreme climate sensitivity, that fact will not become manifest for many centuries because of the enormous thermal inertia of the ocean.

Even in the absurd limit of an infinite climate sensitivity, the amount of warming that can occur is limited by the rate at which the system can accumulate energy (found by setting  $f_a = 1$  and integrating Equation 1). Therefore, from the standpoint of the formal statistical definition—that the tail declines as a power in the limit of  $T \rightarrow \infty$ —there is no fat tail to the climate response at finite time. This is not an esoteric point. As seen in Figure (6), simple analytical damage functions can be highly sensitive to the integration limits. In idealized calculations of economic impacts, it is important to confront the physical realism of those limits.

Recent economic analyses have explored how uncertainty in the physical climate predictions affect policy guidance. Because such analyses invariably include discounting, the results can depend dramatically on how quickly this uncertainty grows into the future: the possibility of large, relatively near-term damages matters more than if those same damages occur in several centuries. We've isolated three physical factors that matter in this regard: the quasi-diffusive nature of ocean heat uptake, the correlated uncertainty between climate forcing and climate sensitivity, and the correct representation of the current climate state. Proper representation of each of these factors tends to reduce the rate at which uncertainty grows with time, and our example calculations with a simple economics model suggest that the impact on economic analyses is significant. In one example of an application of such analyses, none of the three IAMs used in a recent interagency report on the social cost of carbon (IAG, 2010) represent any of these three physical effects.

Many economic models represent a more complicated interplay between growth rate, climate damages, and emissions than we have considered here, and the exact impacts of the effects we have presented will depend on the details of the formulation of the economic equations. Nonetheless, the inferences drawn using such models still depend on fundamental trade-offs between the rate that future marginal consumption is discounted (because of economic growth), and the rate that uncertainty in climate projections grows and becomes more skewed with time. Though the magnitude of the climate response depends on the magnitude of the climate forcing, Fig. 5 illustrates how physical uncertainty in the climate system affects the shape of the uncertainty in the climate response, no matter what the economic assumptions.

High temperatures can of course be reached if the climate system is forced strongly enough, and the illustrative  $C_R$  damage function would predict the distinct possibility of apocalyptic consequences at century's end for a do-nothing scenario (Fig. 5). Such highly skewed distributions can always

be created by an economist making an appropriate choice of damage function, but we emphasize that the skewness is a primarily result of that economist’s choice and not the physical uncertainty in the climate system.

The welfare impacts of climate change ultimately depend on the interplay between three factors: the temperatures resulting from a given carbon emissions profile, the damage functions resulting from those temperatures, and baseline economic growth. Fix any two of these and vary the third, and the results are likely to span the range of policy prescriptions from “don’t worry about it” to “drastic action required immediately”. Discussions of the fat tail of climate sensitivity have focused attention on the uncertainties surrounding the first of these three factors, but in fact the uncertainties surrounding the other two factors are arguably even more profound. The most pressing questions are not about the likelihood of 10 or 20 °C temperature increases, but about the impacts of 3-6 °C temperature increases and about the ability of economic growth to cushion climate impacts in the centuries ahead.

## **Acknowledgements:**

We are very grateful to Marcia Baker and Marshall Baker for conversations and comments that are central to this work; and without implying agreement, to Marty Weitzman and Antony Millner for guidance and advice, to Simon Dietz, Charles Mason, Steve Newbold, and Alex Marten for instructive feedback.

## References

- Allen, M.R., and coauthors, 2006: Observational constraints on climate sensitivity. In *Avoiding Dangerous Climate Change*, ed H.J. Schellnhuber, Cambridge University Press, Cambridge, U.K.
- Archer, D., 2007: Methane hydrate stability and anthropogenic climate change. *Biogeosciences* **4**, 521-544.
- Annan, J.D., and J.C. Hargreaves, 2006: Using multiple observationally-based constraints to estimate climate sensitivity. *Geophys. Res. Lett.*, **33**, L06704, doi:10.1029/2005GL025259 .
- Annan, J. D. and Hargreaves, J. C., 2009: On the generation and interpretation of probabilistic estimates of climate sensitivity, *Clim. Change*, doi:10.1007/s10584-009-9715-y. ility of the CMIP3 ensemble. *Geophys. Res. Lett.*, **37**, L02703, doi:10.1029/2009GL041994.
- Armour, K., and G.H. Roe, 2011: Climate commitment in an uncertain world, *Geophys. Res. Lett.*, **38**, doi:10.1029/2010GL045850.
- Arrhenius, S., 1896: On the influence of carbon acid in the air upon the temperature of the ground. *Philosoph. Mag.*, **41**, 237-276.
- Baker, M.B., and G.H. Roe, 2009: The shape of things to come: why is climate change so predictable? *J. Climate*, **22**, 4574-4589.
- Battisti, D.S. and R.L. Naylor, 2009: Historical Warnings of Future Food Insecurity with Unprecedented Seasonal Heat. *Science*, **323**, no 5911, 240-4.
- Charney, JG, 1979: Carbon dioxide and climate: a scientific assessment. National Academy of Science, 22 pp.

- Diffenbaugh, N.S. and M. Ashfaq, 2010: Intensification of hot extremes in the United States, *Geophysical Research Letters*, L15701, doi:10.1029/2010GL043888.
- Hansen, J., G. Russell, A. Lacis, I. Fung, D. Rind, and P. Stone, 1985: Climate Response Times: Dependence on climate sensitivity and ocean mixing. *Science*, **229**, 857859.
- Hansen, J., Mki. Sato, P. Kharecha, D. Beerling, R. Berner, V. Masson-Delmotte, M. Pagani, M. Raymo, D.L. Royer, and J.C. Zachos, 2008: Target atmospheric CO<sub>2</sub>: Where should humanity aim? *Open Atmos. Sci. J.*, **2**, 217-231, doi:10.2174/1874282300802010217.
- Held, I. M., Winton, M., Takahashi, K., Delworth, T., Zeng, F., and Vallis, G. K., 2010: Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing, *J. Climate*, **23**, 2418-2427.
- Interagency Working Group on Social Cost Carbon, 2010: Technical support document: social cost of carbon for regulatory impact analysis under executive order 12866. Available at [www.epa.gov/otaq/climate/regulations/scc-tsd.pdf](http://www.epa.gov/otaq/climate/regulations/scc-tsd.pdf)
- Knutti, R., G.C. Hegerl, 2008: The equilibrium sensitivity of the Earth's temperature to radiation changes. *Nature Geoscience*, **1**:11, 735-743.
- Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, G.A. Meehl, 2010: Challenges in combining projections from multiple climate models. *J. Climate*, **23**, 2739-2758.
- IPCC, 2007: *Physical Science Basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge Univ. Press.
- Maddison, A., 2007: *Contours of the World Economy, 1-2030 AD*. Oxford University Press, 448pp.
- Nordhaus, W.D., 2008: *A Question of Balance: Economic Modeling of Global Warming*, Yale Press.

- Orr, J.C., and co-authors, 2005: Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature*, **437**(7059), 681-686.
- Roe, G.H. and M. B. Baker, 2007: Why is climate sensitivity so unpredictable? *Science*, **318**, 629-632.
- Roe, G.H., 2009: Feedbacks, time scales, and seeing red. *Annual Reviews of Earth and Planetary Sciences*, 37: 93-115.
- Roe, G.H., 2010: Knowability and no ability in climate change projections. National center for environmental economics report no. 0564. Available at <http://yosemite.epa.gov/ee/epa/erm.nsf/vwAN/EE-0564-117.pdf>.
- Roe, G.H., and K.C. Armour, 2011: How sensitive is climate sensitivity? *Geophysical Research Letters*, *in press*
- Seager, R., M. Ting, I. M. Held, Y. Kushnir, J. Lu, G. Vecchi, H.-P. Huang, N. Harnik, A. Leetmaa, N.-C. Lau, C. Li, J. Velez, and N. Naik, 2007: Model projections of an imminent transition to a more arid climate in southwestern North America, *Science*, **316**, 1181-1184.
- Solomon, S. and fourteen others, 2010: Climate Stabilization Targets: Emissions, Concentrations and Impacts over Decades to Millennia. National Research Council, National Academy of Sciences.
- Stainforth, D.A., T. Aina, C. Christensen, M. Collins, N. Faull, D.J. Frame, J.A. Kettleborough, S. Knight, A. Martin, J.M. Murphy, C. Piani, D. Sexton, L.A. Smith, R.A. Spicer, A. J. Thorpe, and M.R. Allen, 2005: Uncertainty in predictions of the climate response to rising levels of greenhouse gases, *Nature*, **433**, 403-406.



- Torn, M.S., and J. Harte, 2006: Missing feedbacks, asymmetric uncertainties, and the underestimation of future warming, *Geophys. Res. Lett.*, **33**, L10703, doi:10.1029/2005GL025540.
- Weitzman, M.L., 2009a: Additive Damages, Fat-Tailed Climate Dynamics, and Uncertain Discounting. *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 3, 2009-39.
- Weitzman, M.L., 2009b: On Modeling and Interpreting the Economics of Catastrophic Climate Change, *Review of Economics and Statistics*, **91**, 1-19.
- Weitzman, M.L., 2010: GHG Targets as Insurance Against Catastrophic Climate Damages, *in preparation*
- Wigley, T.M.L., and M.E. Schlesinger, 1985: Analytical solution for the effect of increasing CO<sub>2</sub> on global mean temperature. *Nature*, **315**, 649-652.

## Figures

**Figure 1:** Schematic illustration of the simple global-mean climate model of Baker and Roe (2009).

**Figure 2:** It takes a very long time to realize the full equilibrium PDF of climate sensitivity. The figure shows the response to an instantaneous doubling of  $\text{CO}_2$  at  $t = 0$ , and for standard parameters. The shading reflects the feedback ranges as explained in the text. Note the change to a logarithmic  $x$ -axis after  $t = 500$  yr, and see Figure 3 for the full shape of the distribution.

**Figure 3**  $h(\Delta T(t))$  at different times, in response to an instantaneous doubling of  $\text{CO}_2$  at  $t = 0$ . The skewness of the distributions are also shown. Note that, as described in the text, the upper bound on possible temperatures is finite at finite time, limiting the skewness of the distributions.

**Figure 4:** A closer look at the next thousand years, for an instantaneous doubling of  $\text{CO}_2$  at  $t = 0$ : (a) Evolution of the PDF of possible future climates,  $h(\Delta T(t))$ ; (b) evolution of the PDF of possible climate damages, using the quadratic damages function,  $h(C_Q)$ ; (c) welfare-equivalent consumption,  $\hat{C}_Q$ , as a function of the decision horizon. For all curves,  $g = 2\%\text{yr}^{-1}$  and calculations using  $\eta = 2, 3, 4$  are shown as dotted, solid, and dashed lines, respectively. Panels (d) and (e) are the same as panels (b) and (c), but using the reactive damages function,  $C_R$ . Even for the  $3\sigma$  possibility of climate sensitivity, and for the reactive damages function, the willingness-to-pay remains at less than 1% of equivalent consumption. Shading denotes confidence intervals as in Figure 2.

**Figure 5:** As for Figure 4, but for a more realistic climate forcing scenario. Starting at 1850, past forcing (i.e., pre-2011) is consistent with observations given IPCC AR4 (IPCC, 2007). Future forcing through 2100 is approximately consistent with the IPCC A2 scenario (IPCC, 2007). Thereafter climate forcing declines at a rate chosen to approximately stabilize temperatures for the median

value of climate sensitivity. Note that uncertainty in past forcing must be matched to uncertainty in climate feedbacks, in order to match the well-constrained temperature record. Because of this complementary uncertainty in forcing and climate sensitivity, a lower wedge of forcing uncertainty matches with an upper wedge of uncertainty in climate response. See text for more details. Damages in panels c) and d) are calculated with respect to a baseline of 2011. Shading denotes confidence intervals as in Figure 2.

**Figure 6:** PDFs at particular time slices for the realistic forcing scenario of Figure 5. The  $y$ -axis is logarithmic to highlight the behavior in the tails. a) PDFs of temperature. Note that, as explained in Section 2.1, for physical reasons, the temperature PDFs are truncated at finite time; b) PDFs of the quadratic and reactive damage functions  $C_Q$  and  $C_R$ , respectively. Also given in the legends are the skewness of the different distributions. The skewness in  $C_R$  is due primarily to the highly nonlinear nature of the chosen damage function and not to the skewness of the temperature PDF.

**Figure 7:** The impact of an incorrectly initialized climate model. As for Figure 5, but neglecting the relationship between forcing uncertainty and climate sensitivity, and also the thermodynamic imbalance between the surface and the deep ocean. As a result, temperatures (panel b) spread out more quickly than in Figure 5b. Large damages associated with  $C_R$  are realized more quickly, leading to a significantly reduced  $\hat{C}_R$ . The figure illustrates the importance of correctly initializing the climate state at the starting time of an economic analyses.

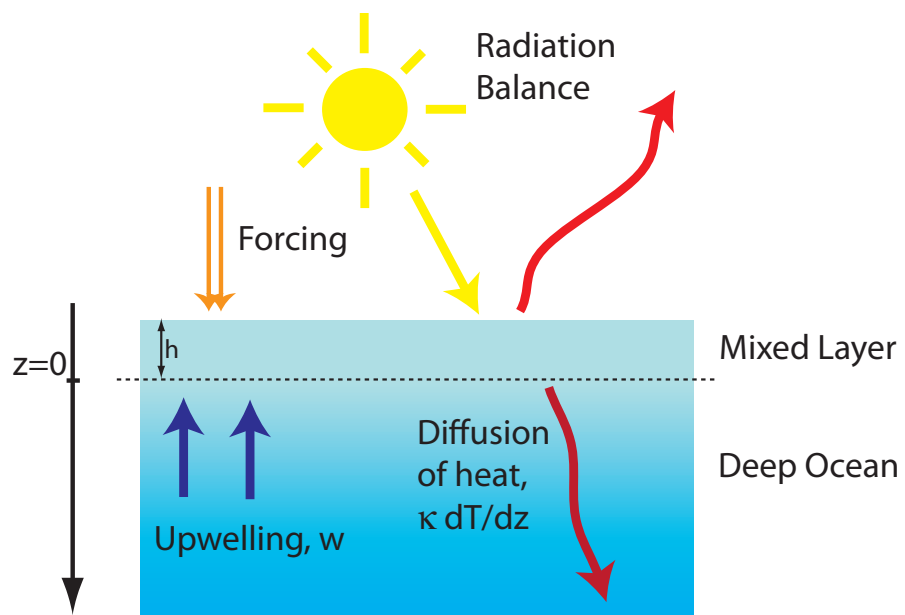


Figure 1: Schematic illustration of the simple global-mean climate model of Baker and Roe (2009).

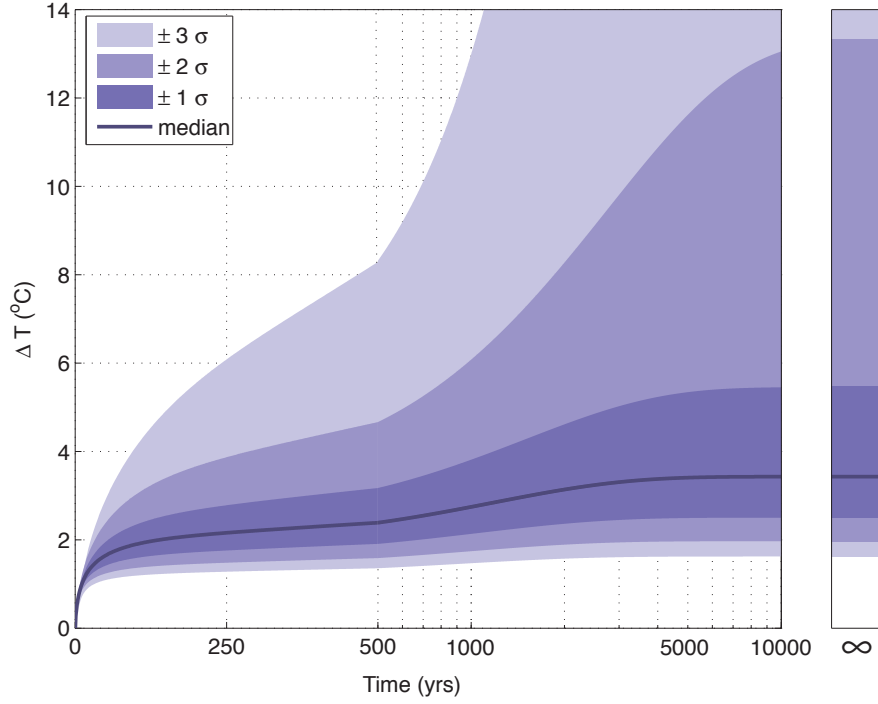


Figure 2: It takes a very long time to realize the full equilibrium PDF of climate sensitivity. The figure shows the response to an instantaneous doubling of  $\text{CO}_2$  at  $t = 0$ , and for standard parameters. The shading reflects the feedback ranges as explained in the text. Note the change to a logarithmic  $x$ -axis after  $t = 500$  yr, and see Figure 3 for the full shape of the distribution.

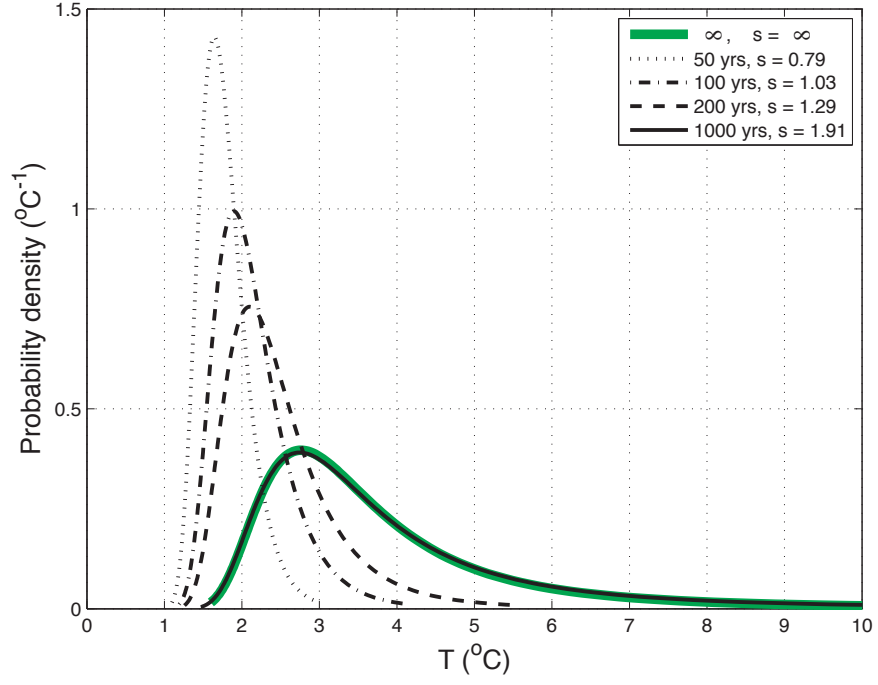


Figure 3:  $h(\Delta T(t))$  at different times, in response to an instantaneous doubling of  $\text{CO}_2$  at  $t = 0$ . The skewness of the distributions are also shown. Note that, as described in the text, the upper bound on possible temperatures is finite at finite time, limiting the skewness of the distributions.

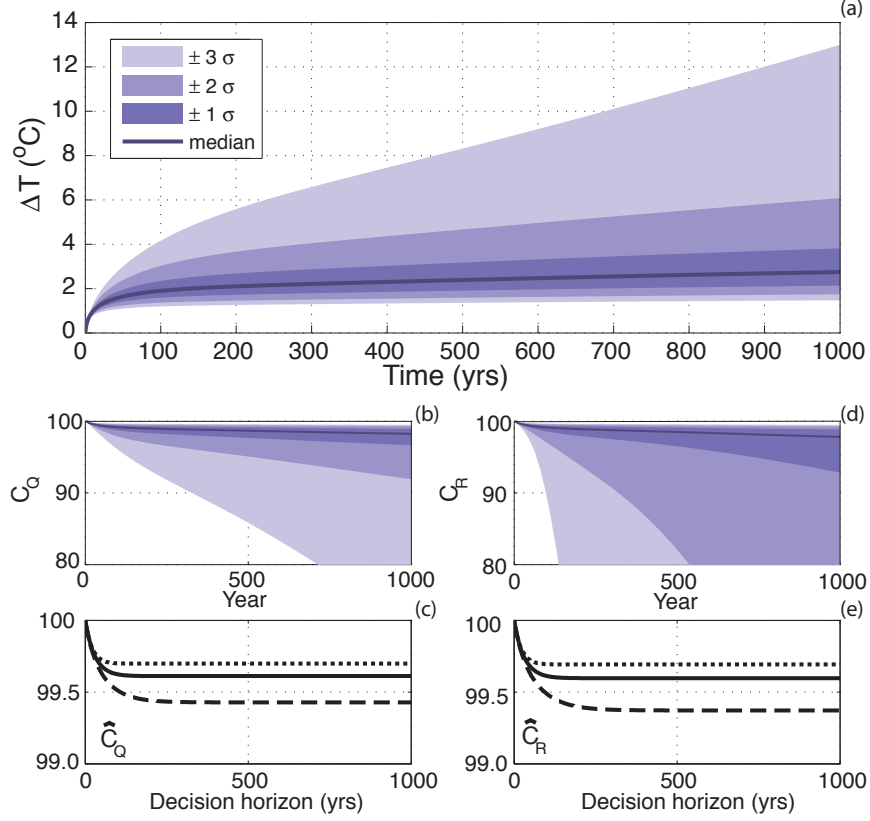


Figure 4: A closer look at the next thousand years, for an instantaneous doubling of  $\text{CO}_2$  at  $t = 0$ : (a) Evolution of the PDF of possible future climates,  $h(\Delta T(t))$ ; (b) evolution of the PDF of possible climate damages, using the quadratic damages function,  $h(C_Q)$ ; (c) welfare-equivalent consumption,  $\hat{C}_Q$ , as a function of the decision horizon. For all curves,  $g = 2\% \text{yr}^{-1}$  and calculations using  $\eta = 2, 3, 4$  are shown as dotted, solid, and dashed lines, respectively. Panels (d) and (e) are the same as panels (b) and (c), but using the reactive damages function,  $C_R$ . Even for the  $3\sigma$  possibility of climate sensitivity, and for the reactive damages function, the willingness-to-pay remains at less than 1% of equivalent consumption. Shading denotes confidence intervals as in Figure 2.

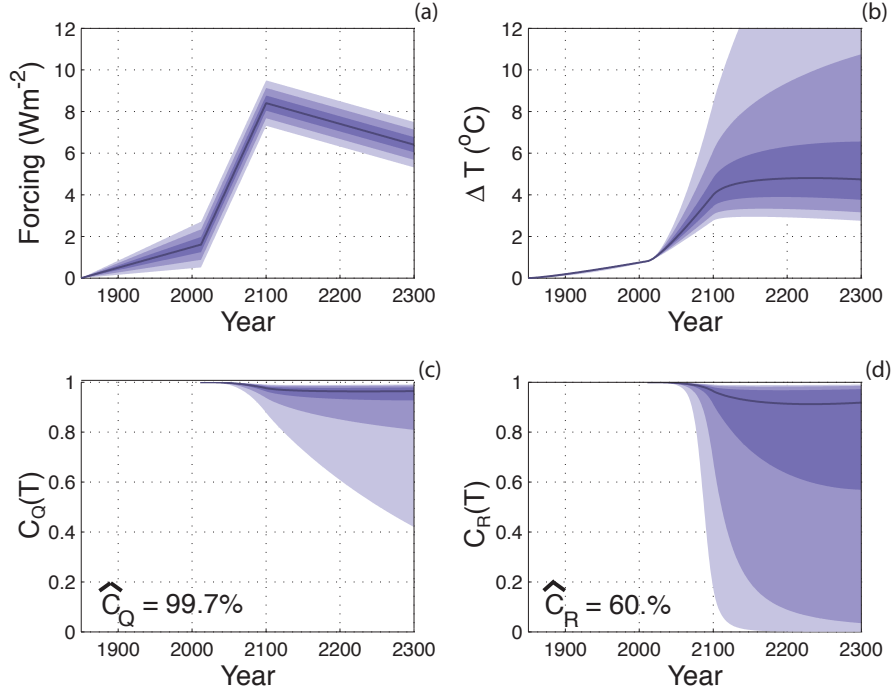


Figure 5: As for Figure 4, but for a more realistic climate forcing scenario. Starting at 1850, past forcing (i.e., pre-2011) is consistent with observations given IPCC AR4 (IPCC, 2007). Future forcing through 2100 is approximately consistent with the IPCC A2 scenario (IPCC, 2007). Thereafter climate forcing declines at a rate chosen to approximately stabilize temperatures for the median value of climate sensitivity. Note that uncertainty in past forcing must be matched to uncertainty in climate feedbacks, in order to match the well-constrained temperature record. Because of this complementary uncertainty in forcing and climate sensitivity, a lower wedge of forcing uncertainty matches with an upper wedge of uncertainty in climate response. See text for more details. Damages in panels c) and d) are calculated with respect to a baseline of 2011. Shading denotes confidence intervals as in Figure 2.



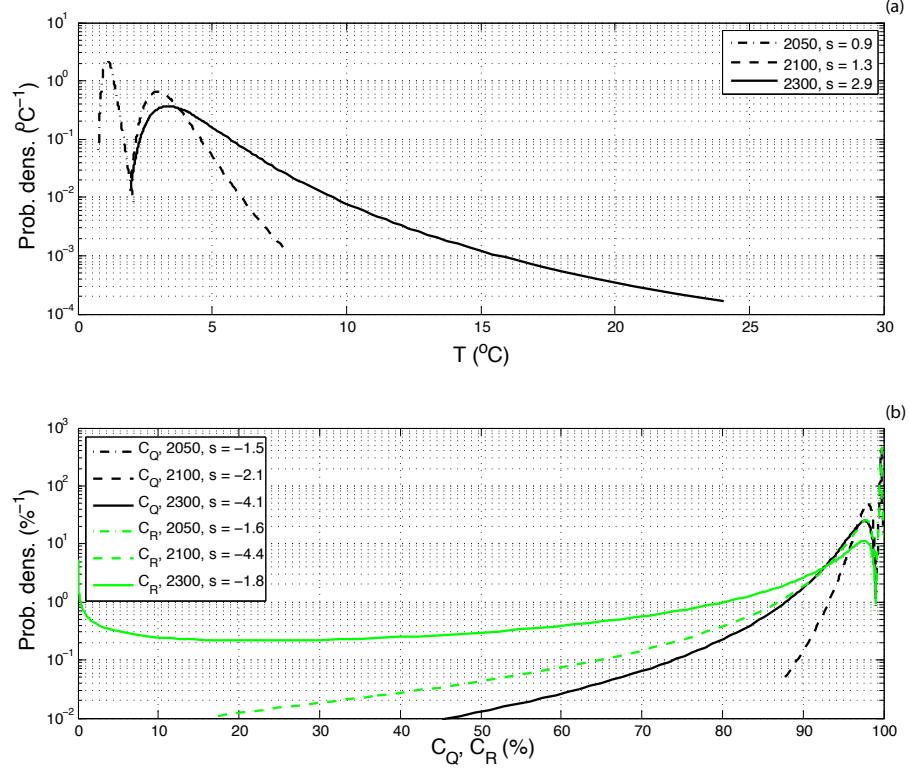


Figure 6: PDFs at particular time slices for the realistic forcing scenario of Figure 5. The  $y$ -axis is logarithmic to highlight the behavior in the tails. a) PDFs of temperature. Note that, as explained in Section 2.1, for physical reasons, the temperature PDFs are truncated at finite time; b) PDFs of the quadratic and reactive damage functions  $C_Q$  and  $C_R$ , respectively. Also given in the legends are the skewness of the different distributions. The skewness in  $C_R$  is due primarily to the highly nonlinear nature of the chosen damage function and not to the skewness of the temperature PDF.

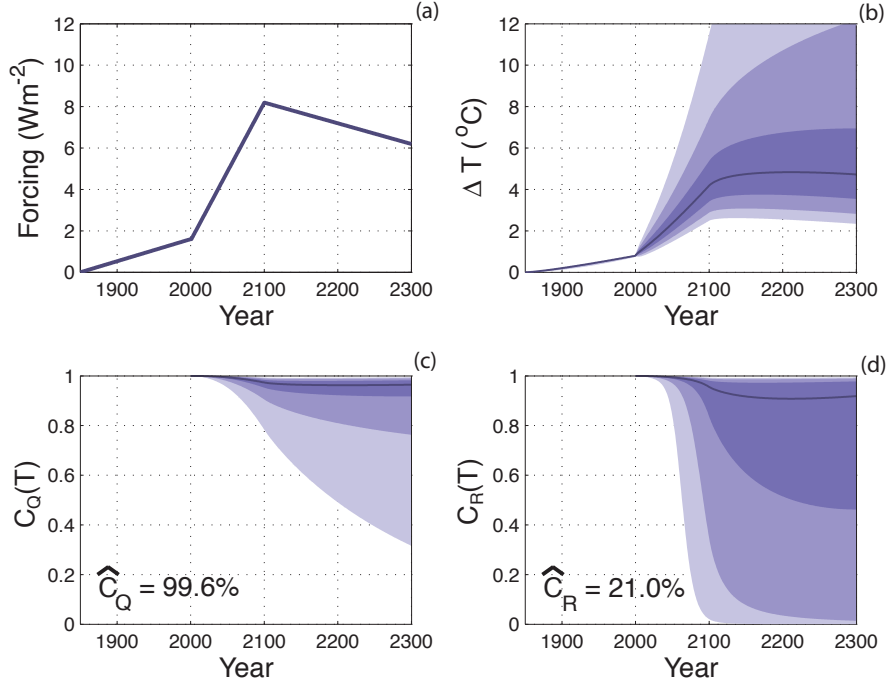


Figure 7: The impact of an incorrectly initialized climate model. As for Figure 5, but neglecting the relationship between forcing uncertainty and climate sensitivity, and also the thermodynamic imbalance between the surface and the deep ocean. As a result, temperatures (panel b) spread out more quickly than in Figure 5b. Large damages associated with  $C_R$  are realized more quickly, leading to a significantly reduced  $\hat{C}_R$ . The figure illustrates the importance of correctly initializing the climate state at the starting time of an economic analyses.