GHG Targets as Insurance Against Catastrophic Climate Damages

Martin L. Weitzman
Harvard University

The climate system is an angry beast and we are poking it with sticks.[†]

Abstract

A critical issue in climate change economics is the specification of the so-called "damages function" and its interaction with the unknown uncertainty of catastrophic outcomes. This paper asks how much we might be misled by our economic assessment of climate change when we employ a conventional quadratic damages function and/or a thin-tailed probability distribution for extreme temperatures. The paper gives some numerical examples of the indirect value of various greenhouse gas (GHG) concentration targets as insurance against catastrophic climate change temperatures and damages. These numerical exercises suggest that we might be underestimating considerably the welfare losses from uncertainty by using a quadratic damages function and/or a thin-tailed temperature distribution. In these examples, the primary reason for keeping GHG levels down is to insure against high-temperature catastrophic climate risks.

1. Introduction

An important question often asked about climate change is: how bad might it get? Catastrophic climate-change damages are characterized by deep

structural uncertainties in the science combined with severe constraints on the ability to evaluate meaningfully the welfare losses from high temperatures. The critical centerpiece of any credible economic analysis of climate change must be its extreme uncertainty. Values of key future variables—temperatures, climate, comprehensive damages, overall welfare, and so forth—cannot be known now. They must be conceptualized instead as random variables, yet to be drawn from some probability density function (PDF). How bad might it get? The answer must ultimately be expressed in the language of tail probabilities.

This paper concentrates on the appropriate way to represent uncertain global warming and uncertain damages. The "damages function" is a notoriously weak link in the economics of climate change, because it is difficult to specify *a priori* and because, as will be shown, the results from a cost-benefit analysis (CBA) or an integrated assessment model (IAM) can be quite sensitive to its specification at the upper end of extreme impacts. Another notoriously weak link in the economics of climate change is the estimation of tail fatness for PDFs associated with extreme warmings. These problems are especially acute at catastrophically high temperatures, because huge uncertainties surround any estimates of extreme damages or probabilities of climate-change disasters.[1]

The paper investigates what might happen to an economic analysis of climate change with a significantly more reactive damages function than the standard quadratic and with PDFs having tails of varying degrees of fatness. The purpose is not to propose a valid detailed model of climate change, but rather to show on a toy model that basic conclusions of such a model can rely strongly on the assumptions of the model. The paper attempts to give some extremely rough ballpark estimates of the differences in steady-state temperature PDFs and damages as a function of greenhouse gas (GHG) target concentration levels. These differences can vary greatly according to the specification, but on the whole they are substantial enough to suggest that in some situations—especially when catastrophic damages interact with fat-tailed uncertainty—we might be underestimating welfare losses considerably. With the examples being considered in this paper, the primary reason for keeping target GHG levels down is to insure against high-temperature catastrophic climate risks.[2]

Climate change is so complicated, and it involves so many sides of so many different disciplines and viewpoints, that no analytically-tractable model or paper can aspire to illuminate more than but a facet. Because

---

[1] These problems have not much been studied. One of the very few IAM-based studies of extreme climate change is by Dietz (2011), which also contains references to earlier work in this area.

[2] This is not exactly "insurance" in the traditional sense because the probability of a catastrophe is being reduced by prevention measures, rather than being compensated by making the planet whole again after the catastrophe.

the climate change problem is so complex, there is frequent reliance on sophisticated numerical computer simulations. These can be indispensable, but sometimes they do not provide a simple intuition for the processes they are modeling. In this paper I go to the opposite extreme by focusing on relatively tractable comparative-steady-state solutions. What I am presenting here is a kind of "stress test" approach to grasping intuitively the robustness of modeling highly uncertain extreme damages.

2. Uncertain Equilibrium Warmings

There are so many sources of uncertainty in climate change that a person almost does not know where or how to begin cataloging them. For specificity, I focus especially on the uncertainty of so-called "equilibrium climate sensitivity."

Equilibrium climate sensitivity (hereafter denoted S) is a key macro-indicator of the eventual temperature response to GHG changes. It is defined as the global average surface warming that follows a sustained doubling of atmospheric carbon dioxide (CO_2), after the climate system has reached a new equilibrium.[3] Calculating the actual time trajectory of temperatures is a complicated task that requires sophisticated computer modeling based on general circulation models with hundreds of parameters and variables. The human mind being what it is, however, there is a compelling need to reduce and relate such a complicated dynamic reality to an aggregate indicator—like S. This is a simplistic reduction that overlooks important spatial and temporal aspects of climate change. Nevertheless, the concept is still very useful for capturing the "big picture"—perhaps because the more complicated simulation models find that several aspects of climate change seem to scale approximately linearly with S.[4] As just one example of an application of this convenient reductionism, the GHG concentrations that would prevent so-called "dangerous anthropogenic interference"—however it is defined—are often made by back-of-the-envelope calculations based on S. But because S is uncertain, the uncertain temperature changes induced by a given GHG concentration can only be described in terms of probabilities. This paper follows closely the spirit and assumptions (and drawbacks) of the S-reductionist approach.

It should be clearly understood that under the rubric of "equilibrium climate sensitivity" I am trying to aggregate together an entire suite of

---

[3] In scientific jargon, S is a so-called "fast equilibrium" concept based upon "fast feedbacks" (geologically speaking). The concept omits slower-acting feedbacks, such as changes in albedo, changes in biological sinks or sources, temperature-induced releases of carbon from clathrates, and the like. So-called "earth system sensitivity" includes slower-acting feedbacks and is presumably larger, perhaps significantly so. For a time horizon on the scale of 150 years or so, it is not implausible that "earth system sensitivity" might be the more relevant concept. Greater details are available, e.g., in Hansen *et al.* (2008).

[4] See, e.g., Knutti and Hegerl (2008).

uncertainties, including some non-negligible unknown unknowns. So climate sensitivity is to be understood here as a prototype example or a metaphor, which is being used to illustrate much more generic issues in the economics of highly uncertain climate change. The insights and results of this paper are not intended to stand or fall on the narrow issue of accurately modeling uncertain climate sensitivity *per se*. Whatever its source, greater uncertainty typically strengthens the case I am trying to make in this paper. Empirically, it is not the fatness of the tail of the climate sensitivity PDF *alone* or the reactivity of the damages function to high temperatures *alone*, or the degree of relative risk aversion *alone*, or the rate of pure time preference *alone*, or any other factor *alone*, that counts, but rather the *interaction* of all such factors in determining the upper-tail fatness of the PDF of the relevant measure of overall expected welfare.

The Intergovernmental Panel on Climate Change in its IPCC-AR4 (2007) Executive Summary explains S this way: "The equilibrium climate sensitivity is a measure of the climate system response to sustained radiative forcing. It is not a projection but is defined as the global average surface warming following a doubling of carbon dioxide concentrations. It is *likely* to be in the range 2 to 4.5°C with a best estimate of 3°C, and is *very unlikely* to be less than 1.5°C. Values substantially higher than 4.5°C cannot be excluded, but agreement of models with observations is not as good for those values." The IPCC defines "likely" as a probability above 66% but below 90%, which would mean that the probability that climate sensitivity is greater than 4.5°C (Prob[$S \geq 4.5°C$]) is between 5% and 17%. A more recent average estimate of 14 leading climate scientists is Prob[$S \geq 4.5 °C$] = 23%.[5] In this paper I choose 70% as defining "likely" and I calibrate all upper-tail probability distributions so that P[$S \geq 3 °C$] = 50% and P[$S \geq 4.5 °C$] = 15%.[6]

The upper-half tail of the probability distribution is the region $S > S_M$, whose total probability mass is 0.5, where the climate-sensitivity median is taken to be $S_M = 3°C$. I use three two-parameter PDFs to represent this upper-half tail of climate sensitivity: (1) the Normal distribution, which has a thin upper tail; (2) the Pareto (or Power) distribution, which has a fat upper tail; (3) the Lognormal distribution which has an upper tail on the borderline between fat and thin. There is some wiggle room in the definition of what constitutes a fat-tailed PDF or a thin-tailed PDF, but almost everyone agrees that probabilities declining exponentially or faster (like the Normal) are thin tailed, while probabilities declining polynomially or slower (like the Pareto) are fat tailed. The intermediate-tailed Lognormal is an interesting borderline case because the probabilities in its upper tail decline slower than

---

[5] Zickfield, Morgan, Frame, and Keith (in review).

[6] I lean more toward P[$S \geq 4.5 °C$] $\leq$ 17% than toward P[$S \geq 4.5 °C$] $\geq$ 5% because, for a time horizon of a century and a half or so, it is plausibly the more inclusive "earth system sensitivity" that matters more than the "fast equilibrium sensitivity" that IPCC-AR4 refers to. See also footnote 1.

exponentially but faster than polynomially.[7] For all three PDFs I calibrate the two parameters so that $P[S \geq 3] = 0.5$ and $P[S \geq 4.5] = 0.15$. A major goal of this paper is to experiment with different PDFs above the median value of $S_M = 3°C$. For the purposes of this paper, very little depends on the exact form of the PDF for the 50% of probability *below* the median. By contrast, we are forced to speculate and extrapolate wildly concerning the PDF for the 50% of probability *above* the median, and, as we shall see, this can have major consequences.

The notation $f_I(S)$ refers to the PDF of climate sensitivity S. The subscript $I = L$ refers to a **Lognormal** PDF, the subscript $I = N$ refers to a PDF whose upper-half tail is **Normal**, and the subscript $I = P$ refers to a distribution whose upper-half tail is **Pareto** (or **Power**).

I begin with the base case of the Lognormal, whose upper-half PDF here is

$$f_L(S) = \frac{1}{0.3912 \sqrt{2\pi}\, S} \exp\left(-\frac{(\ln S - 1.099)^2}{2(0.3912)^2}\right) \tag{1}$$

for all $S \geq 3$. As can readily be confirmed, the parameter values in Equation (1) have been calibrated so that $P[S \geq 3°C] = 0.5$ and $P[S \geq 4.5\,°C] = 0.15$. I also consider two other possibilities for the upper-half tail: a fat-tailed Pareto PDF and a thin-tailed Normal PDF.

My upper-half-tail Pareto PDF is also specified by its parameters being set so that simultaneously $P_P[S \geq 3] = 0.5$ and $P_P[S \geq 4.5] = 0.15$. It is readily confirmed that the corresponding upper-half-tail Pareto PDF is

$$f_P(S) = 38.76\, S^{-3.969}. \tag{2}$$

My upper-half-tail Normal PDF is again specified by its two parameters being set so that simultaneously $P_N[S \geq 3] = 0.5$ and $P_N[S \geq 4.5] = 0.15$. It is readily confirmed that the corresponding upper-half-tail Normal PDF is

$$f_N(S) = \frac{1}{1.447 \sqrt{2\pi}} \exp\left(-\frac{(S-3)^2}{2\,(1.447)^2}\right). \tag{3}$$

The following table gives some values for the three cumulative distributions.

I think that not many climate scientists would quibble about the "big picture" of the PDF of climate sensitivity given by Table 1 for low values of climate sensitivity. For what it is worth, the median upper 5% probability level over all 22 climate-sensitivity PDFs cited in IPCC-AR4 is 6.4°C, which is consistent with the Pareto PDF above.[8] Notice that the absolute probabilities of very high values of S are quite small. Even so, the *relative* probabilities of

---

[7] The moment generating function of a Lognormal PDF is infinite, although every moment is finite.

[8] Details in Weitzman (2009a).

Table 1: $P[S \geq \widehat{S}]$ for the three probability distributions used in this paper

$\widehat{S} =$	3°C	4.5°C	6°C	8°C	10°C	12°C	15°C	20°C
$P_{\mathrm{P}}[S \geq \widehat{S}]$	0.5	0.15	0.064	0.027	0.014	0.008	0.004	0.002
$P_{\mathrm{L}}[S \geq \widehat{S}]$	0.5	0.15	0.038	0.006	0.001	2×10^{-4}	2×10^{-5}	6×10^{-7}
$P_{\mathrm{N}}[S \geq \widehat{S}]$	0.5	0.15	0.019	0.003	7×10^{-7}	3×10^{-10}	6×10^{-17}	4×10^{-32}

high S are extremely dependent on whether the upper tail of the relevant PDF is fat, thin, or intermediate.

The next step is to convert PDFs of equilibrium climate sensitivity S into PDFs of equilibrium temperature change T, as a function of given stable greenhouse gas (GHG) target concentrations. Let G stand for atmospheric GHGs as measured in parts per million (ppm) of carbon dioxide equivalent (CO_2e). Climate sensitivity corresponds to the equilibrium temperature change eventually induced by a sustained doubling of CO_2e. Let $\Phi(G)$ represent the "forcing factor" as a function of the steady-state GHG level G, with $\Phi(G)$ normalized by making $\Phi(560) \equiv 1$. An atmospheric concentration of $G = 560$ ppm represents a doubling of the pre-industrial-revolution level of $G = 280$. As is well known, the forcing factor Φ increases linearly in the *logarithm* of CO_2e concentrations.[9] With normalization $\Phi(560) \equiv 1$, the precise formula is

$$\Phi(G) = \frac{\ln{(G/280)}}{\ln 2}. \tag{4}$$

Therefore, a given constant level of GHGs G and a given equilibrium climate sensitivity S translates into a steady-state temperature change of

$$T = \Phi(G) \times S. \tag{5}$$

If $f_{\mathrm{I}}(S)$ is the relevant PDF of climate sensitivity, then, from the standard Jacobean formula for change of random variables, the relevant PDF of temperatures T for a given level of G is

$$\psi_{\mathrm{I}}(T \mid G) = \frac{f_{\mathrm{I}}(T/\Phi(G))}{\Phi(G)}. \tag{6}$$

To anchor the upper tail of extreme warmings, I arbitrarily focus on multiples of 6°C. Six degrees of extra warming is about the upper limit of what the human mind can envision for how the state of the planet might change. It serves as a routine upper bound in attempts to communicate what the most severe global warming might signify, including the famous "burning embers" diagram of the IPCC and several other popular expositions.[10] Twelve degrees

---

[9] See, e.g., Archer (2007).
[10] See, e.g., IPCC-AR4 (2007) and Lynas (2007). One recent study (Kriegler *et al.*, 2009) asked 52 experts for their subjective probability estimates of triggering a "tipping point of

of global warming is used here as an example of a round number ($12°C = 2 \times 6°C$) that transcends our ability to imagine, with any reasonable measure of accuracy, what the earth might be like for super-high temperature increases. Throughout the numerical examples that follow, I arbitrarily take $18°C$ ($= 3 \times 6°C$) to be an upper bound beyond which temperatures are not allowed to go—by fiat. Thus, for all calculations of expected values, damages are capped at $18°C$ and probabilities of such damages are calculated as $P[T \geq 18°C]$. In this sense $18°C$ might be envisioned as something akin to a global "death temperature."

The issue of how to deal with the deep structural uncertainties in climate change would be completely different and immensely simpler if systemic inertias, like the time required for the system to naturally remove extra atmospheric CO_2, were short, as is the case for many airborne pollutants like particulates, sulfur dioxide, and ozone. Then an important component of an optimal strategy might be along the lines of "wait and see." With strong reversibility, an optimal climate-change policy should logically involve (among other elements) waiting to learn how far out on the bad fat tail the planet might end up, followed by midcourse corrections if we seem to be headed for a disaster. Alas, the problem of climate change seems bedeviled almost everywhere by significant stock-accumulation inertias—in atmospheric CO_2, in the absorption of heat or CO_2 by the oceans, in the uptake of CO_2 by the biosphere, in albedo changes, in the wildcard reaction to warming temperatures of the enormous stocks of methane clathrates trapped in permafrost and continental shelves, and in many other relevant physical and biological processes that are extremely slow to respond to attempts at reversal.

Take atmospheric carbon dioxide as a prime example. Solomon *et al.* (2009) calculated how concentrations of CO_2 would be expected to fall off over time if *all* anthropogenic emissions were to cease immediately, following a future 2% annual growth rate of emissions up to peak concentrations of 450, 550, 650, 750, 850, and 1,200 ppm. As the authors state: "The example of a sudden cessation of emissions provides an upper bound to how much reversibility is possible, if, for example, unexpectedly damaging climate changes were to be observed." Results differed for different trajectories and scenarios, but a crude rule of thumb seemed to be that approximately 70% of the peak enhancement level over the preindustrial level of 280 ppm persevered after 100 years of *zero* emissions, while approximately 40% of the peak enhancement level over the preindustrial level of 280 ppm persevered after 1,000 years of *zero* emissions. In the Solomon *et al.* study, were atmospheric CO_2 concentrations to peak at 800 ppm, followed forever thereafter

---

major changes" in each of five possible categories: (1) the Atlantic meridional overturning circulation; (2) the Greenland Ice Sheet; (3) the West Antarctic Ice Sheet; (4) the Amazon rainforest; (5) the El Niño/Southern Oscillation. For what it is worth, at an average temperature increase of $T \approx 6°C$ the *expected* (probability weighted) number of such expert-assessment tipping points was three (out of a possible five).

by zero emissions, then atmospheric concentrations would be ≈650 ppm after 100 years and ≈500 ppm after 1,000 years. A different recent study by Archer *et al.* (2009) reached essentially the same conclusions. These numbers do not look to me like evidence supporting "wait and see" policies. The capacity of the oceans to take up atmospheric heat, the saturation of carbon sinks, the loss of albedo, and many other relevant mechanisms tell a similar story of long stock-accumulation irreversibilities relative to the time it takes to filter out and act upon meaningful signals of impending disasters.

In Table 2, the first row represents steady-state atmospheric stocks of GHG concentrations G (measured in ppm of CO_2e). The second row below it gives the median equilibrium temperature T_M as a function of stabilized GHG stocks. The rows starting just below T_M give the probabilities of achieving at least the steady-state temperature increase represented by the entries in the table (6°C or 12°C or 18°C) for each of the three chosen PDFs (**P**areto = fat tail, **L**ognormal = intermediate tail, **N**ormal = thin tail).

The thing that seems so striking about Table 2 is how relatively rapidly the probabilities of high temperatures increase as a function of GHG concentrations – and how dependent these high temperatures can be on the assumed fatness of the upper tail of the PDF of climate sensitivity. Throughout Table 2, the target level of GHG concentrations influences strongly the probabilities of especially high temperatures. One can readily see in shorthand form what are the ultimate temperature consequences of moving from lower to higher steady-state GHG concentrations. Of course these ultimate temperature consequences are expressible only as probabilities. It can be quite misleading to look just at measures of central tendency, like the median. What to me is far more alarming than the relatively moderate rise of T_M as a function of G is what is happening in the upper reaches of the various PDFs, where the really catastrophic outcomes are concentrated. The higher levels of GHGs seem especially worrisome to me because they are pushing temperature probabilities towards the upper tail at an uncomfortably rapid rate.

To see things most sharply, notice at the two opposite extremes that 400 ppm of G here effectively blocks temperatures from rising much above 6°C, whereas 1,000 ppm of G here assigns a probability of ≈41% to P[$T \geq$ 6°C] and ≈1%–5% to P[$T \geq$ 12°C], depending on the assumed tail fatness. Notice too how the differences between the three different PDFs (with three different degrees of fatness in their tails) are manifested for various GHG concentrations. Throughout most of Table 2 there is a disturbingly non-robust dependence of outcomes on the presumed fatness of the upper tail of the PDF, which we simply cannot know. The thin-tailed normal distribution effectively excludes the really hotter temperatures, while the fat-tailed Pareto distribution presents a much more worrisome picture. This awkward dependence upon presumed tail fatness is more pronounced the deeper one penetrates into the extreme tail of the underlying PDF of climate sensitivity. At the higher concentrations of GHGs, say ≥650 ppm of CO_2e, a temperature

Table 2: Probabilities of exceeding $T = 6°C$, $T = 12°C$, $T = 18°C$, for given G = ppm of CO_2e

G :	400	450	500	550	600	650	700	750	800	900	1000
T_M :	$1.5°$	$2.1°$	$2.5°$	$2.9°$	$3.3°$	$3.6°$	$4.0°$	$4.3°$	$4.5°$	$5.1°$	$5.5°$
$P_P[T \geq 6]$	0.9%	2%	4%	6%	8%	11%	15%	18%	22%	30%	39%
$P_L[T \geq 6]$	10^{-4}	0.3%	1%	3%	6%	10%	14%	19%	24%	33%	41%
$P_N[T \geq 6]$	10^{-9}	10^{-5}	0.2%	1%	4%	9%	14%	20%	25%	35%	43%
$P_P[T \geq 12]$	0.11%	0.26%	0.48%	0.75%	1.1%	1.5%	1.9%	2.3%	2.8%	3.8%	5.0%
$P_L[T \geq 12]$	10^{-7}	10^{-6}	10^{-5}	0.02%	0.05%	0.12%	0.23%	0.41%	0.65%	1.4%	2.3%
$P_N[T \geq 12]$	10^{-45}	10^{-24}	10^{-15}	10^{-10}	10^{-8}	10^{-6}	10^{-5}	0.02%	0.03%	.22%	.73%
$P_P[T \geq 18]$	0.03%	0.08%	0.14%	0.23%	0.32%	0.44%	0.56%	0.69%	0.84%	1.2%	1.5%
$P_L[T \geq 18]$	10^{-10}	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-5}	0.01%	0.01%	.02%	.06%	.12%
$P_N[T \geq 18]$	10^{-108}	10^{-58}	10^{-37}	10^{-26}	10^{-20}	10^{-16}	10^{-13}	10^{-11}	10^{-10}	10^{-7}	10^{-6}

increase of 6°C is sufficiently close to the middle-body range of all three climate-sensitivity PDFs that tail fatness *per se* does not matter so much in determining P[$T \geq 6$]. On the other hand, tail fatness always matters a lot for determining P[$T \geq 12$], even for higher GHG concentrations ≥650, because this part of the range of temperatures is well into the extreme tail of the underlying climate-sensitivity PDFs. In more colorful language, very different tails may be appended to animals having roughly similar bodies.

I want to state clearly and emphatically here, once and for all, that very high atmospheric temperature changes like T = 12°C will likely take millennia to attain. The higher the limiting temperature, the longer it takes to achieve atmospheric equilibrium because the oceans will first have to absorb the enormous amounts of heat being generated. Alas, building up such enormous amounts of heat in the ocean is like compressing a very powerful coiled spring. It could set in motion nasty surprises, such as long-term methane clathrate releases from the continental shelves, whose possibly horrific consequences are essentially irreversible and would have to be dealt with later. For the toy model of this paper, overall damages generated by equilibrium T = 12°C are best conceptualized as associated with *being on the trajectory* whose asymptotic limiting atmospheric temperature change is T = 12°C. It is important to bear this interpretation in mind, even though a discrete date will be assigned for the "as if" transition to higher temperatures in the toy model.

3. Uncertain Damages From Climate Change

From the very outset, the representation of damages from climate change presents some severe conceptual and practical problems. Here I merely follow most of the literature by postulating that damages from increased temperatures are manifested in reduced form *as if* they impair output multiplicatively.[11] In this version of a just-so story, all losses from climate change are interpreted as if they literally translate into a welfare-equivalent fractional loss of consumption. My own preferred specification leans more towards being additive in the utilities of consumption and temperature change, rather than multiplicative, which can make a huge difference because an additively-separable specification does not allow one to substitute increased consumption for increased temperature changes nearly so readily as the standard multiplicatively-separable specification.[12] However, in this paper I show that *even with* a high-substitutability multiplicatively-separable formulation, results can be sensitive to the specification of the damages function for extreme temperature changes. This paper examines only the damages side, and that very simplistically. I do not try to explicitly estimate costs of achieving various

---

[11] This multiplicative story could be challenged, and it makes a difference. See Weitzman (2009b).

[12] See Weitzman (2009b) and Weitzman (2010) for some specific examples.

GHG targets, much less attempt to determine an optimal policy by explicitly balancing the costs of achieving a given GHG target against its benefits.

So what should the damages function be for very high temperatures? No one knows, of course. The predominant approach attempts to calculate what the world would be like for a given small increase in global average temperatures. The climate-change economist tries to quantify such things as net damages (after subtracting out adaptation costs) from changes in: agricultural productivity, life styles, population movements, rising oceans, hurricanes, and so forth. This is a constructive approach that probably represents the best we can do for small temperature changes. But I am uneasy when this approach is extended to large changes in global average temperatures. Taking an extreme example, suppose for the sake of argument that average global warming were to increase by the extraordinary amount of $12°C$ (with an extraordinarily low probability, of course). It is true that people live very well in places where the average temperature is $12\ °C$ higher than in Yakutsk, Siberia. However, I do not think that these kinds of analogies can justify using such a comparative geography approach for estimating welfare-equivalent damages from an average *planetary* temperature change of $12°C$, which translates into extraordinarily variable *regional* climate changes. Global mean temperatures involve a double averaging: across space and over time. A "damages function" is a reduced form representing global welfare losses from global average temperatures, which subsumes a staggering amount of regional, seasonal, and even daily weather heterogeneity. Regional and seasonal climate changes are presumably much more unpredictable than global average surface temperatures. There is just too much structural uncertainty and too much heterogeneity to put trustworthy bounds on the unprecedented almost-unimaginable changes to planetary welfare from average global temperatures increasing by $12°C$. When there is such big uncertainty about catastrophic damages, and when the damages function for high temperature changes is so conjectural, the relevant degree of risk aversion, yet another important unknown here, will tend to play a significant role in an economic analysis of climate change.

Of course I have no objective way to determine the magnitudes of high-temperature damages. The last time the world witnessed periods where global average temperatures were very roughly $10°C$ or so above the present was during the Eocene epoch $≈55–34$ million years ago. During these warming periods the earth was ice free while palm trees and alligators lived near the North Pole. The Eocene was also the last epoch in which there were geologically rapid increases in mean global temperatures of magnitude $≈5°C$ or so above an already warm background. Such hyperthermal events occurred over an average period of very roughly $≈100,000$ years or so, which is extremely gradual compared to current worst-case anthropogenically-induced trajectories. It is unknown what exactly triggered these Eocene temperature increases, but they were accompanied by equally striking atmospheric carbon increases. One likely culprit is the strong-feedback release of large amounts

of methane hydrates from clathrate deposits, which is a non-negligible possibility over the next century or two if current GHG emissions trends are extrapolated.[13] The major point here is that relatively rapid changes of global average temperatures $\approx 6°C$ above present values are extremely rare events and are extraordinarily far outside the scope of human experience. For huge temperature increases like $T \approx 12°C$, the planetary effects are even more difficult to imagine. To find a *geologically instantaneous* increase in average global temperatures of magnitude $T \approx 12°C$, one would have to go back hundreds of millions of years.

For me, 12°C has a vivid image as a kind of a reference point from a recent study, which estimated that global average temperature increases of $\approx 11–12°C$ (with, importantly, accompanying humidity in the form of high wet-bulb temperatures) would cross an absolute thermodynamic limit to metabolic heat dissipation. Across this threshold, represented by a wet-bulb temperature of 35°C, more than half of today's human population would be living in places where, at least once a year, there would be periods when death from heat stress would ensue after about six hours of exposure.[14] (By contrast, today the highest wet bulb temperature anywhere on Earth is about 30°C). The authors of this study furthermore point out: "This likely overestimates what could practically be tolerated: Our (absolute thermodynamic) limit applies to a person out of the sun, in a gale-force wind, doused with water, wearing no clothing and not working." Even at much lower wet-bulb temperatures than 35°C, life would become debilitating and physical labor would be unthinkable. The massive unrest and uncontainable pressures this might bring to bear on the world's human population are almost unimaginable. The Earth's ecology, whose valuation is another big uncertainty, would be upended. A temperature change of $\approx 12°C$ therefore represents an extreme threat to human civilization and global ecology as we now know it, even if, conceivably, it might not necessarily mean the end of *Homo sapiens* as a species. I don't think that a person needs accurate specific stories about what might happen for $T \approx 12 °C$ to imagine truly ruinous catastrophes undoing the planet and severely undermining the security of human civilization – at the very minimum.

Let T represent the *change* in future worldwide average surface temperature, always measured in degrees Centigrade. Let $\widetilde{C}(T)$ represent welfare-equivalent consumption as a *fraction* of what potential consumption would be in the absence of any climate change (at that imaginary imputed time when it is "as if" the random variable T materializes). Remember, in this artificial story consumption changes discretely when T materializes all at once at an imaginary pre-specified time, with $\widetilde{C}(T)$ at that time representing a hypothetical shock to welfare-equivalent consumption along the entire future dynamic trajectory.

---

[13] For more about methane clathrates, see Archer (2007) or the recent paper by Shakova *et al.* (2010).
[14] Sherwood and Huber (2010).

The most popular single formulation of a damages function in the literature is the quadratic form $\widetilde{C}_Q(T) = 1/\left[1 + (T/\alpha)^2\right]$, where α is a positive temperature-scaling parameter calibrated to give some "reasonable" values of $\widetilde{C}_Q(T)$ for relatively small warmings, say up to $T \approx 2.5°$C. Standard estimates of α in the literature are more or less similar, although I hasten to add that such calibrations were intended by the authors to capture low-temperature damages and were never intended to be extrapolated to very high temperature changes, which is just what I will be doing here. For the sake of having a specific prototype example, I calibrate α in Equation (7) to conform with the damages function in the latest version of the well known DICE model of William Nordhaus (2008), where he effectively used $\alpha = 20.46$.[15] In this case, welfare equivalent consumption is given by the formula

$$\widetilde{C}_Q(T) = \frac{1}{1 + \left(\dfrac{T}{20.46}\right)^2},\tag{7}$$

where the natural scaling factor for T is the rather large temperature $\alpha = 20.46°$C.

The results in terms of welfare-equivalent relative consumption levels for this quadratic case are given by $\widetilde{C}_Q$ in the second row of Table 3. (The third-row variable $\widetilde{C}_R$ will be discussed presently.)

I do not find such numbers as $\widetilde{C}_Q(T)$ in Table 3 at all convincing for high temperatures. At the mind-bending average global temperature change of $T = 18°$C, the welfare-equivalent damage as a fraction of consumption *at that time* (when $T = 18°$C materializes and conventional consumption is presumably much higher than today) is "only" 44%. The implied welfare-equivalent consumption damages of 35% for $T = 15°$ and 19% for $T = 10°$ also seem to me to be far too low for doing a credible analysis of the consequences of catastrophic losses from extreme climate change. My tentative conclusion is that the quadratic form (7), which was never intended to be applied for temperature changes beyond a few degrees centigrade, is not appropriate for assessing the welfare impacts of disastrously high temperature changes. The quadratic "welfare equivalent" damages function (expressed as a fraction of what potential consumption would be if $T = 0$), which is enumerated

Table 3: Welfare-equivalent consumption $\widetilde{C}_Q(T)$ and $\widetilde{C}_R(T)$

T	2°C	3°C	4°C	5°C	6°C	7°C	8°C	9°C	10°C	12°C	15°C	18°C
$\widetilde{C}_Q$	99%	98%	96%	94%	92%	90%	87%	84%	81%	74%	65%	56%
$\widetilde{C}_R$	99%	97%	91%	75%	50%	27%	13%	7%	3%	1%	0.2%	0.1%

---

[15] Nordhaus's DICE model is perhaps the most famous IAM in the economics of climate change. The quadratic coefficient $1/\alpha^2 = 0.00239$ was used to generate his Figure 3-3 on page 51. Nordhaus wisely does not try to project beyond $T = 6°$, which is suitable for his purposes but unsatisfactory for mine.

as $\widetilde{C}_Q$ in the second row of Table 3, is pre-ordained to make extreme climate change look empirically negligible almost no matter what else is assumed.

All damage functions are made up—especially for extreme situations. Therefore, neither I nor anyone else has an objective basis for determining magnitudes of high-temperature damages. I now want to "give the devil his due" by characterizing very roughly two points on a much more reactive global damages function, which seems to me far more plausible than the quadratic and which attributes far bigger welfare-equivalent damages to higher temperatures. Of course no one knows how to estimate welfare-equivalent damages for very high temperature changes. I anchor my "give the devil his due" damages function on the two "iconic" (if arbitrary) global-average temperature changes: $6°C$ and $12°C$. What these two iconic global warmings might mean for the human condition and for the rest of the planet has already been sketched out. At $6°C$ I propose welfare-equivalent consumption of $\widetilde{C}_R(6°C) = 50\%$ (at that time), while for $12°C$ I propose welfare-equivalent consumption of $\widetilde{C}_R(12°C) = 1\%$. I do not consider such estimates to be extreme when interpreted as welfare-equivalent damages to a fictitious agent representing the entire planet, although others may disagree and are free to substitute their own guesstimates.

Some IAMs and CBAs recommend a "climate policy ramp" gradualism that would approach atmospheric CO_2 levels of ≈ 700 ppm, which would arguably make GHG CO_2e levels be ≈ 750 ppm. From Table 2, GHG concentrations of 750 ppm would eventually result in temperature increases $\geq 6°C$ with probability $\approx 19\%$ and temperature increases $\geq 12°C$ with average probability $\approx 1\%$ (depending very much upon how fat-tailed is the relevant PDF). Using the proposed reactive specification of damages $[\widetilde{C}_R(6°C) = 50\%$ and $\widetilde{C}_R(12°C) = 1\%]$, I calculated for the lognormal PDF that at $G = 750$ ppm of CO_2e there is $\approx 19\%$ chance of damages greater than 50% and $\approx 1\%$ chance of damages greater than 99%. With the quadratic damages function (7) shown in Table 3, at $G = 750$ ppm I calculated for the lognormal PDF that the probability of damages $\geq 50\%$ is $\approx 0.1\%$, while the probability of damages $\geq 99\%$ is $\approx 10^{-8}$. With these kinds of numbers, it is no wonder that a quadratic damages function is fearless about attaining CO_2e concentrations of 750 ppm—or even much higher!

The third row of Table 3 adds a term to the denominator of Equation (7) making it have the polynomial form $\widetilde{C}_Q(T) = 1/[1 + (T/\alpha)^2 + (T/\beta)^\gamma]$, where (as before) $\alpha = 20.46°C$, while I calibrated the temperature-scaling factor β and the exponent γ so that $\widetilde{C}_R(6°C) = 50\%$ and $\widetilde{C}_R(12°C) = 1\%$. The relevant parameter values are $\beta = 6.081$ and $\gamma = 6.754$. For this case, in place of the "non-reactive" (7) we have instead a "reactive" damages function of form

$$\widetilde{C}_R(T) = \frac{1}{1 + \left(\dfrac{T}{20.46}\right)^2 + \left(\dfrac{T}{6.081}\right)^{6.754}}. \tag{8}$$

The reason I kept the quadratic term from Equation (7) in Equation (8) with the same value of α as in Equation (7) is to make Equation (8) behave like the standard damages function (7) for small values of T. Notice, however, that when temperature changes having a scaling factor of 6.081°C are exponentiated to such a high power as 6.754 in Equation (8), the consequence is something like a tipping point where the damages function changes dramatically around the "iconic" global warming level of ≈ 6°C. Not surprisingly, $\widetilde{C}_R(T)$ in Table 3 is indeed more "reactive" than $\widetilde{C}_Q(T)$ to higher temperature changes. As mentioned, global average temperatures are arbitrarily forbidden from going above $\overline{T} = 18$°C, which corresponds in Table 3 to $\widetilde{C}_R(18°C) \approx 0.1\%$. I readily admit the artificiality of a damages function whose specification has been created primarily to show how an extrapolation from small to large temperature changes has the potential to exert undue influence on model conclusions.

4. Welfare Effects of Uncertain Climate Change

From Equation (5), steady-state global warmings T (given steady-state GHG levels G), are equal to the forcing function $\Phi(G)$ [defined by expression (4)] times climate sensitivity S. Since S is a random variable with some postulated PDF, then (for any given G) T is a random variable with PDF given by Equation (6). And then, given some postulated damages function of temperature [namely, here, $\widetilde{C}_Q(T)$ or $\widetilde{C}_R(T)$], welfare-equivalent consumption in that steady state is itself a random variable. I now manufacture an artificial numerical example of welfare impacts by linking the uncertain-temperature methodology of Section 2 with the damages functions of Section 3.

Suppose a constant relative risk aversion utility function (of consumption) having the form

$$U(C) = \frac{C^{1-\eta}}{1-\eta}, \tag{9}$$

where η is the coefficient of relative risk aversion. With r being the interest rate, ρ being the rate of pure time preference (or "utility discount rate"), and g being the growth rate of per-capita consumption, the fundamental Ramsey equation is

$$r = \rho + \eta g. \tag{10}$$

Following what Ramsey originally proposed, I take the rate of pure time preference (or the so-called "utility discount rate") throughout this paper to be zero [i.e., $\rho = 0$ in Equation (10)]. As Ramsey famously put the issue, "it is assumed that we do not discount later enjoyments in comparison with earlier ones, a practice which is ethically indefensible and arises merely from the weakness of the imagination." Several other (but far from all) famous economists concur with this Ramsey interpretation of intergenerational equity.[16] Taken together, quotations from these "famous economists" sound to

---

[16] Pigou: (pure time preference) "implies ... our telescopic faculty is defective." Harrod: "pure time preference (is) a polite expression for rapacity and the conquest of reason by

me much more like a normative judgement about intergenerational ethics than a description of short-run individual behavior. I think that the Ramsey case of zero discounting of future utilities is the appropriate abstraction for a normative analysis of climate change. Ethically or morally, the Ramsey abstraction treats the utility of different generations equally, while taking full account of the fact that economic growth will make future generations richer and less needy than the present generation. My base-case CRRA coefficient is $\eta = 3$, which corresponds to an eminently plausible degree of risk aversion that I believe is close to the median "best estimate" of economists. My base-case future growth rate of per capita consumption is $g = 2\%$ per year. These base-case values imply from Equation (10) an interest rate of $r = 6\%$ per year, and therefore the numerical results to follow cannot in any way be ascribed to assuming an unrealistically low discount rate.

Were η to be changed substantially, then r and g would not mesh quite so nicely with past reality. If $\eta = 2$ and $r = 6\%$, then Equation (10) with $\rho = 0$ implies $g = 3\%$—probably too high. If $\eta = 4$ and $g = 2\%$, then Equation (10) with $\rho = 0$ implies $r = 8\%$, also probably too high. So I think it is fair to say that this proposed "package" of base-case point-estimate values ($\rho = 0$, $r = 6\%$, $\eta = 3$, $g = 2\%$) looks more or less realistic, is internally consistent, and is immune from the criticism that discounting of climate change is being marginalized.[17]

For my base case I use the lognormal PDF of S (with its intermediate tail fatness), as given by Equation (1). I assume a particularly simplistic time scenario. Let G be the GHG target. For the next τ years, consumption grows at annual rate $g = 2\%$ and GHG levels build up to (and stay at) G. My base case is $\tau = 150$ years. Then, suddenly, at time $\tau = 150$ years from now, consumption is reduced by a fraction corresponding to the realization of T (given G), and the assumed damages function (namely $\widetilde{C}_Q$ or $\widetilde{C}_R$). Remember, the atmospheric equilibrium for large values of T may occur many centuries, and even millennia, into the future. However, the welfare *impact* of the damages generated by equilibrium T is conceptualized as associated with *being on the trajectory* whose asymptotic limiting atmospheric temperature change is T, and is being normalized as if it occurs at time τ. After this "as if" permanent shock to the level of consumption $\tau = 150$ years in the future, growth continues thereafter at annual rate $g = 2\%$. In other words, there are no damages whatsoever until time $\tau = 150$ years from now, when the sky is allowed to artificially fall all at once. The growth rate in this just-so story is never impacted,

---

passion." Koopmans: "(I have) an ethical preference for neutrality as between the welfare of different generations." Solow: "in solemn conclave assembled, so to speak, we ought to act as if the social rate of pure time preference were zero." [All quotes are taken from Arrow (1999).] I think it should be clear that the above citations refer to a normative or prescriptive, rather than a positive or descriptive, view of the world.
[17] These values of $\rho = 0$, $\eta = 3$, and $g = 2\%$ per year are close to those that were proposed by Dasgupta (2008), and were considered fully acceptable by Nordhaus (2008, pp. 61 and 187).

either before or after the output shock at $\tau = 150$. This is a very primitive formulation, but I think it is good enough to make the point that a reactive damages function and a tail of intermediate fatness suffice to dominate the effects of discounting, even at the very high interest rate of $r = 6\%$ and with climate changes localized as if occurring a century and a half from now. At an interest rate of 6%, the relevant discount *factor* for goods and services a century and a half hence is $\exp(-0.06 \times 150) = \exp(-9) = 0.01\%$, which is a very low number.

Without loss of generality, present consumption at time $t = 0$ is normalized at $C(0) = 1$. Let $\widehat{C}$ represent the deterministic-equivalent value of $C(0)$ that would give the same welfare relative to there being zero climate change. Then $\widehat{C}$ is the implicit solution of the equation

$$\int_0^\infty U(\widehat{C}\exp(0.02t))\,dt = \int_0^{150} U(\exp(0.02t))\,dt$$
$$+ E\left[\int_{150}^\infty U(\widetilde{C}_J\exp(0.02t))\,dt\right], \tag{11}$$

where $E(\cdot)$ is the expected-value operator and $J = Q$ or $J = R$.

Now substitute the lognormal PDF (1), (6) (for $\mathbf{I} = \mathbf{L}$), the utility function (9) (for base case $\eta = 3$), and the damages functions (7) and (8) into Equation (11). Solving for $\widehat{C}$ then yields

$$\widehat{C}_J = (1 - e^{-6} + e^{-6}E[(\widetilde{C}_J)^{-2}])^{-0.5}, \tag{12}$$

where $J = Q$ or $J = R$. The following table gives rounded-off values of $\widehat{C}_Q$ and $\widehat{C}_R$ as a function of G.

With the quadratic damages function (7), there is essentially the same welfare-equivalent consumption level of $\approx 100\%$ independent of GHG concentrations G. This is because the expected present discounted welfare impact of quadratic damages incurred a century and a half from now, evaluated at an interest rate of $r = 6\%$, is essentially zero. Thus, with a standard quadratic damages function, in this formulation GHG concentrations of 1,000 ppm of CO_2e are essentially no worse than GHG concentrations of 400 ppm of CO_2e when discounted at rate $r = 6\%$ per year. There is only a miniscule willingness to pay (WTP) now to avoid significantly higher GHG concentrationsa century and a half from now. No wonder, then, that optimal IAM trajectories derived from a quadratic damages function encourage gradualist climate-policy-ramp CO_2e levels approaching 750 ppm!

Table 4: Certainty-equivalent consumption for base-case lognormal PDF, $\eta = 3, \overline{T} = 18$

G:	400	450	500	550	600	650	700	750	800	900	1000
$\widehat{C}_Q$	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
$\widehat{C}_R$	100%	99.95%	99.4%	97.6%	92.4%	82.7%	70.2%	57.9%	47.2%	32.7%	24.0%

Table 5: Sensitivity analysis giving $\widehat{C}_R$ for $G = 750$ ppm

Case	CRRA η	PDF	Impact year τ	Growth g	Temp bd $\overline{T}$	Welfare
base	3⇒58%	L⇒58%	150⇒58%	2%⇒58%	18°⇒58%	$\widehat{C}_R$⇒58%
alt1	2⇒88%	N⇒81%	200⇒89%	3%⇒95%	12°⇒91%	$\widehat{C}_Q$⇒100%
alt2	4⇒22%	P⇒27%	100⇒25%	1%⇒16%	24°⇒38%	n.a.

The welfare-equivalent certainty-equivalent "as if" consumption levels $\widehat{C}_R$ enumerated in Table 4 are each expressed *relative* to an artificial norm of $G = 280$ ppm, $T = 0$. In other words these numbers represent the WTP, in terms of certainty-equivalent consumption now and forever, to eliminate *all* climate change. The various "as if" consumption levels $\widehat{C}_R$ (as a function of steady-state G) are difficult to interpret in absolute terms, and should be compared with each other as fractions or multiples. For example, the welfare-equivalent fractional loss of as-if-deterministic consumption accompanying a change in GHG concentrations from 550 ppm to 750 ppm by Table 4 is $(0.976 - 0.579)/0.976 = 0.407$—i.e., keeping target GHG levels at 550 ppm rather than letting them rise to 750 ppm is worth spending up to 40.7% of present certainty-equivalent consumption at $G = 550$. Or, to take another example, the welfare-equivalent fractional loss of as-if-deterministic consumption accompanying a change in GHG concentrations from 600 ppm to 650 ppm in Table 4 is $(0.924 - 0.827)/0.924 = 0.105$. In other words, a modest coefficient of relative risk aversion of 3 is big enough to make it worth spending 10.5% of consumption, now and forever, to avoid the higher (but still very small) probabilities of bad warmings 150 years from now (which are indicated in Table 2) by keeping target GHG levels at 600 ppm rather than letting them rise to 650 ppm.

Notice that the WTP to keep GHG concentrations below ≈500 ppm of CO_2e are small because such low concentrations effectively wall off the higher temperature changes—and discounting moderate events happening 150 years from now at an effective interest rate of 6% (or an effective discount factor of 0.01%) takes care of the lower temperature changes. Above ≈550 ppm of CO_2e, however, the danger of higher temperatures accelerates greatly the WTP now in order to avoid bad climate change outcomes a century and a half from now, overriding even a discount rate of 6% per year. Such high WTP levels are testimony to the power of combining risk aversion with fat tails and a reactive damages function. At greater and greater GHG concentrations, risk aversion to the possibility of taking a catastrophic "hit" to consumption becomes more and more the dominant force in WTP calculations. Again, I think the inescapable conclusion is that GHG policy is most accurately viewed as an insurance policy against disastrous outcomes.

I now mention briefly a few results from some primitive sensitivity experiments. In order to compress these results, I report them only for 750 ppm of CO_2e. I pick 750 ppm of CO_2e for two reasons. First, 750 ppm

of CO_2e is an upper limit on optimal GHG concentrations that is approached by some optimizing IAMs and CBAs. Second, under non-optimal business-as-usual scenarios, concentrations of 750 ppm of CO_2e are all too conceivable as early as the end of this century.

Generally speaking, outcomes are very dependent on how the extreme tail damages and the extreme tail probabilities are formulated. In other words, results are *not* robust to how catastrophic outcomes are modeled and specified. In this sense, the main robust finding of the paper is *non-robustness to stress tests*.

As was already shown, the standard quadratic damages function never produces a significant enough welfare impact to matter very much in determining policy. The very first form of non-robustness to report on, therefore, is the sensitivity of results to the form of the damages function, already discussed previously. The "devil's advocate" reactive damages function paints a very different picture in Table 4 than the standard non-reactive quadratic damages function.

I next examine what happens for different values of the coefficient of relative risk aversion (still keeping $\rho = 0$ in the background—more on this later). From Table 4, welfare-equivalent consumption for $\eta = 3$ is 58%. For $\eta = 4$, welfare-equivalent consumption at 750 ppm is 22%. For $\eta = 2$, welfare-equivalent consumption at 750 ppm is 88%. So quantitative values of the WTP to avoid a GHG concentration of 750 ppm vary widely with the assumed degree of risk aversion, and so too do corresponding policy recommendations. What is extremely interesting here is the strong reversal of the commonly-accepted role of η, which comes from thinking too much in terms of the purely deterministic case. Through the deterministic Ramsey formula, a higher value of η is traditionally associated with a higher value of the discount rate (here $r = \eta g$), for any given growth rate of consumption g. This higher value of η then translates into a lower weighting for distant-future events, like climate change. As an example, with an annual growth rate $g = 2\%$ the relevant discount factor for converting benefits a century and a half from now into today's currency for $\eta = 4$ is $\exp(-0.02 \times 150 \times 4) = 6.1 \times 10^{-6}$; for $\eta = 3$ it is $\exp(-0.02 \times 150 \times 3) = 1.2 \times 10^{-4}$; and for $\eta = 2$ it is $\exp(-0.02 \times 150 \times 2) = 2.5 \times 10^{-3}$. This is a very wide relative range for discount factors, although all of these numbers are extremely low in absolute terms. However, in the presence of uncertainty, higher values of η also indicate higher relative risk aversion, which can have an even more powerful effect in the opposite direction for a reactive damages function combined with a semi-fat upper tail of the temperature-change PDF. Thus, the damages function (8) combined with even an intermediate-fatness tail like the lognormal is sufficient to reverse the traditional role of η, because the effect of aversion to catastrophic uncertainty here outweighs the effect of time discounting.

If the Pareto fat-upper-tail PDF (2) is substituted for the lognormal (1) in the range of climate sensitivity above the median $S_M = 3$, then

welfare-equivalent consumption at 750 ppm is 27% instead of 58% in Table 4. If the normal thin-upper-tail PDF (3) is substituted for the lognormal (1) above the median $S_M = 3$, then welfare-equivalent consumption at 750 ppm is 81% instead of 58% in Table 4. The WTP to avoid the uncertain consequences of a GHG concentration of 750 ppm of CO_2e is thus highly dependent on the assumed fatness of the upper tail of the PDF of climate sensitivity.

In Table 4, I assumed that the global warming impacts of being on a trajectory that eventually asymptotes to temperature change T can be assessed "as if" they arrive at $\tau = 150$ years from now. If the time of arrival for global warming is artificially pinpointed instead at $\tau = 200$ years from now, then welfare-equivalent consumption at 750 ppm is 89% instead of 58% in Table 4. If the time of arrival for global warming is artificially pinpointed at $\tau = 100$ years from now, then welfare-equivalent consumption at 750 ppm is 25% instead of 58% in Table 4.

For the base case enumerated in Table 4, I projected an annual growth rate of consumption $g = 2\%$. If the annual growth rate of consumption is instead $g = 1\%$, then welfare-equivalent consumption at 750 ppm is 16%, instead of 58% in Table 4. If the annual growth rate of consumption is instead $g = 3\%$, then welfare-equivalent consumption at 750 ppm is 95%, instead of 58% in Table 4.

Finally, I examine the artificially imposed upper bound cutoff $\overline{T}$, beyond which global average temperatures are arbitrarily not allowed to go. In Table 4, I assumed an upper-bound temperature cutoff of $\overline{T} = 18°C$. If the upper-bound temperature cutoff is arbitrarily made 6° higher, so that $\overline{T} = 24°C$, then welfare-equivalent consumption at 750 ppm is 38% instead of 58% in Table 4. If the upper-bound temperature cutoff is artificially made 6° lower, so that $\overline{T} = 12°C$, then welfare-equivalent consumption at 750 ppm is 91% instead of 58% in Table 4.

In the following Table 5, I summarize briefly the results of the above sensitivity experiments.

Readers can form their own judgements, but for me Table 5 seems to be indicating a disturbing lack of robustness with respect to parameter values that are extremely difficult to know with any degree of accuracy. Many researchers promote alternative specifications that do not imply nearly such extreme outcomes as do some of my specifications. I do not claim that their formulations are wrong, or even implausible. I merely point out that they are unlikely to be robust with respect to assumptions about extreme catastrophic climate change under uncertainty, and therefore they fail a reasonable stress test.[18]

---

[18] The study of Costello *et al.* (2010) finds a remarkably low WTP to avoid climate change. However, when their model is re-run with my base-case specifications of the damages and utility functions, then WTP to avoid climate change for an upper bound of $\overline{T} = 18°C$

To test parametric sensitivity with respect to the rate of pure time preference ρ, consider the base case ($\rho = 1.5\%$, $\eta = 2$, $g \approx 2\%$) in Nordhaus (2008), which from Equation (10) corresponds to a discount rate $r \approx 5.5\%$ per year. When Nordhaus runs instead through his DICE model "my" base case ($\rho \approx 0$, $\eta = 3$, $g \approx 2\%$), which implies $r = 6\%$ per year, there is no substantive difference in outcomes.[19] However, DICE is (essentially) a deterministic formulation in the spirit of an optimal control problem featuring a relatively non-reactive quadratic loss function. When lognormal uncertainty of form (1) is introduced, then there is a tremendous difference between the two base cases. With quadratic losses (7), as we have come to expect, $\widehat{C}_Q \approx 100\%$. Even with my reactive damages function (8), if I plug the Nordhaus base-case specification ($\rho = 1.5\%$, $\eta = 2$, $g = 2\%$) into my simplistic model I get $\widehat{C}_R \approx 99\%$ for $G = 750$ ppm of CO_2e. In other words, the WTP now to avoid altogether the consequences of $G = 750$ at future time $\tau = 150$ goes from $\approx 42\%$ to $\approx 1\%$ for two specifications that would otherwise have near-identical consequences in a deterministic world. The reason for such a dramatic difference is that when pure time discounting is as high as $\rho = 1.5\%$ per year, the risk aversion effect is overcome by the discounting effect. Once again, readers can form their own judgements about the implied robustness of policy implications under stress-test uncertainty—here with respect to various combinations of ρ and η.

5. Discussion

I think that several themes emerge from this paper.

The paper suggests that an economic analysis of climate change might be very sensitive to uncertainties about such things as the fatness of PDF tails for temperature changes, the specification of the damages function, cutoff bounds, relative risk aversion, rates of pure time preference, growth rates, concentrations of GHGs, and so forth. When relatively fat-tailed PDFs are combined with a reactive damages function, then seemingly modest changes in target levels of GHGs can sometimes have very big welfare consequences. In such conditions, the primary purpose of keeping down GHGs is to prevent large damages from extreme warmings in the "bad" tail, which is a much more powerful incentive to target low GHG levels than trying to keep down the relatively modest damages from median temperatures. But the exact quantitative extent to which changes in target levels of GHGs can cause these very big welfare consequences depends sensitively on how the extremes are modeled and specified. While conclusions from some plausible formulations

---

becomes $\approx 95\%$—i.e., welfare equivalent consumption becomes $\approx 5\%$. (I am indebted to Christopher Costello for re-running the Costello *et al.* model with my specifications.) A tentative conclusion is that the results of their model are even less robust than the results of the model of Nordhaus (2008) to the "stress test" numerical exercises of this paper.

[19] This is essentially "Run 3" reported on page 187 of Nordhaus (2008).

seem relatively immune to being represented by a measure of central tendency like the median, conclusions from some other formulations, which appear equally if not more plausible to me, are extraordinarily far from being captured by median values and seem to be highly dependent on a variety of underlying uncertainties. Thus, we might be in an unfortunate position where results from an economic analysis of climate change have a wide range of possible policy recommendations, which depend upon barely knowable assumptions well beyond the realm of ordinary experience. While I do not think that this feature nullifies climate-change CBAs or IAMs, I do think it should make us especially cautious about the ability of economic analysis to give robust policy advice for the specific application of such methods to catastrophic climate change. The moral of the story may be that, under extreme tail uncertainty, seemingly casual decisions about functional forms and parameter values for catastrophic outcomes might well dominate CBA of climate change.

Another suggestion coming out of the paper is that the standard quadratic damages function simply cannot register, and therefore will not react to, the possibility of catastrophic climate change. Once the usual quadratic damages specification is made, an optimal policy will not get alarmed by high values of GHG concentrations, and almost inevitably it will recommend relatively mild mitigation measures. The climate-policy-ramp gradualism that emerges from many IAMs may be rooted in the fact that, even when uncertainty is introduced in the form of Monte Carlo simulations, the usual quadratic damages function never really allows the model to get very far away from effectively plugging median values into a deterministic climate-change CBA and then discounting away the consequences.

Needless to say, a very large number of caveats apply to the toy model of this paper.

The main omission is the lack of realistic dynamics in the toy model. For simplicity, the toy model of this paper essentially analyzes and compares steady states, with only the most primitive cause-and-effect dynamics. I think that this simplification allows some transparent insights that can get obscured by much more complicated dynamic models, but it comes at a price by omitting nuanced considerations of growth, discounting, how long it takes to approach a steady state, and so forth. A drawback of my toy model approach is that I could be missing some critical dynamic interactions that are unable to be captured by the crudeness of such a simplistic comparative-steady-state formulation. So conclusions of this paper are at most suggestive and may need to be modified in the light of performing numerical simulations from much more complicated dynamic computer models.[20]

---

[20] Absent numerical integration over PDFs with given functional forms to obtain exact expected values (as was done in this paper), and even with stratified sampling, an enormous number of simulations may be required in order to penetrate far enough into the "bad" tail to derive valid expected values for very-low-probability very-high-impact situations.

That having been said, the suggestive comparative-steady-state numerical outcomes of this toy model seem to me as if they might be sufficiently powerful that an appropriately muffled version would likely survive a fully dynamic treatment. Remember, throughout this paper the default fatness that *really* matters concerns the bad tail of the reduced-form PDF of some overall measure of "welfare." Nobody knows the tail fatness of the PDF of "welfare"—which is the very feature driving the lack of robustness to specifications of climate extremes. As a consequence, I somehow doubt very much that robustness will miraculously be restored by introducing more sophisticated dynamics. Actually, the various "stress tests" of this paper do not strike me as being particularly "stressful" at all. Therefore, I suspect rather strongly that it may be difficult to dislodge altogether the verdict that a CBA of climate change can be quite sensitive to assumptions about extreme tail events—and that the primary reason for keeping GHG levels down may be mainly to insure against high-temperature catastrophic climate damages.

6. Conclusion

If a particular type of idiosyncratic uncertainty affects only one small part of an individual's or a society's overall portfolio of assets, exposure is naturally limited to that specific component and bad-tail fatness is not such a paramount concern. However, some very few but very important real-world situations have potentially *unlimited* exposure due to structural uncertainty about their potentially open-ended catastrophic reach. Climate change potentially affects the whole worldwide portfolio of utility by threatening to drive all of planetary welfare to disastrously low levels in the most extreme scenarios. With global climate change, diversification is limited because all eggs are inherently in one basket. The comparative-steady-state toy model of this paper suggests that the results of climate change CBA can sometimes depend non-robustly on seemingly casual decisions about functional forms and parameter values associated with extreme tails. The findings of this paper may be a warning that the results of climate change CBA can be largely driven by the "fear factor" associated with low-probability high-impact catastrophes, which is difficult to model robustly.

References

ARCHER, D. (2007) *Global Warming*. Oxford: Blackwell.

ARROW, K. J. (1999) Discounting, morality, and gaming, in *Discounting and Intergenerational Equity*, P. R. Portney and J. P. Weyant, eds., pp. 13–21. Washington, DC: Resources for the Future.

COSTELLO, C., M. NEUBERT, S. POLASKY, and A. SOLOW (2010) Bounded uncertainty and climate change economics, *Proceedings of the National Academy of Sciences* **107**, 8108–8110.

DASGUPTA, P. (2008) Discounting climate change. *Journal of Risk and Uncertainty* **37**, 141–169.

DIETZ, S. (2009) High impact, low probability? An empirical analysis of risk in the economics of climate change, *Climate Change* **103**, 519–541 .

HANSEN, J., M. SATO, P. KARECHA, D. BEERLING, V. MASSON-DELMOTTE, M. PAGANI, M. RAYMO, D. L. ROYER, and J. C. ZACHOS (2008) Target atmospheric CO_2: Where should humanity aim?, *Open Atmospheric Science Journal* **2**, 217–231.

IPCC-AR4 (2007) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge: Cambridge University Press, 2007 (available from: http://www.ipcc.ch).

KNUTTI, R., and G. C. HEGERL (2008) The equilibrium sensitivity of the Earth's temperature to radiation chages, *Nature Geoscience* **1**, 735–743.

KRIEGLER, E., J. W. HALL, H. HELD, R. DAWSON, and H. J. SCHELLNHUBER (2009) Imprecise probability asessment of tipping points in the climate system, *Proceedings of the National Academy of Sciences* **106**, 5041–5046.

LYNAS, M. (2007) *Six Degrees: Our Future on a Hotter Planet*. New York: HarperCollins.

NORDHAUS, W. D. (2008) *A Question of Balance*. New Haven, CT: Yale University Press.

SHAKHOVA, N., I. SEMILETOV, A. SALYUK, V. YUSUPOV, D. KOSMACH, and O. GUSTAFSSON (2010) Extensive methane venting to the atmosphere from sediments of the East Siberian Arctic shelf, *Science* **327**, 1246–1250.

SHERWOOD, S. C., and M. HUBER (2010) An adaptability limit to climate change due to heat stress, *Proceedings of the National Academy of Sciences* **107**, 9552–9555.

SOLOMON, S., G.-K. PLATTNER, R. KNUTTI, and P. FRIEDLINGSTSTEIN (2009) Irreversible climate change due to carbon dioxide emissions, *Proceedings of the National Academy of Sciences* **106**, 1704–1709.

WEITZMAN, M. L. (2009a) On modeling and interpreting the economics of catastrophic climate change, *Review of Economics and Statistics* **91**, 1–19.

WEITZMAN, M. L. (2009b) Additive damages, fat-tailed climate dynamics, and uncertain discounting, *Economics e-journal* **3**, 2009–39.

WEITZMAN, M. L. (2010) What is the "damages function" for global warming—and what difference might it make? *Climate Change Economics* **1**, 57–69.

ZICKFIELD, K., M. G. MORGAN, D. FRAME, and D. W. KEITH (in review) Expert judgments about transient climate response to alternative trajectories of radiative forcing, *Proceedings of the National Academy of Sciences*.