GHG Targets as Insurance Against Catastrophic Climate Damages

Martin L. Weitzman[*]

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 GHG concentration targets as insurance against catastrophic climate-change temperatures and damages. These numerical examples 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 will 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 absolutely critical centerpiece of any credible economic analysis of climate change has to be 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 a random variable (RV), yet to be drawn from some probability density function (PDF) that has been adequately "stress tested."

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 temperatures. Another notoriously weak link in the economics of climate change is the estimation of tail fatness for the PDF of extreme temperatures. These problems are especially acute at catastrophically high temperatures, because huge uncertainties surround any estimates of extreme damages or probabilities of climate-change disasters. Ideally, one wants analytically tractable forms that capture adequately the economic reality of global warming. The existing literature on CBAs and IAMs of climate change mostly concentrates on super-moderate quadratic damages and on super-thin-tailed point-mass PDFs.

This paper addresses the issue of what might happen to an economic analysis of climate change with a significantly more reactive damages function than the quadratic and with PDFs having tails of varying degrees of fatness. 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 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. The critical question here is: *How fast* does the probability of a catastrophe decline relative to the welfare impact of the catastrophe. Even tiny probabilities can be offset by negative welfare impacts that are big enough. In such conditions the fact that the tiny probabilities are themselves unknown is, other things being equal, more troubling than if they were known precisely. 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. For situations where fat-tailed PDFs are combined with a reactive damages function, the welfare differences between various target GHG levels are typically very large and there is a much stronger case for keeping down GHG target levels than when tails are thin or damages are quadratic.

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 of this problem. Because 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" for grasping quickly and intuitively the issue of highly uncertain extreme damages, which is several steps removed from practical implementation. This paper is mostly about conceptualizing the problem of high-temperature catastrophic damages and giving some rough sense of the magnitudes involved via particular numerical examples. It is less about giving decisive numerical values for actual practical policy advice, although some crude policy implications will be apparent. The beauty of this toy-model approach is that the formulas I will use are sufficiently simple and transparent that readers can easily plug in different specifications or attach the model to other frameworks and make their own inferences. A drawback of my toy model approach is that it might be missing some critical dynamic interactions that are unable to be captured by the crudeness of what is largely an exercise in comparative steady states. So any conclusions of this paper are at most suggestive and may need to be modified in the light of performing detailed numerical simulations from much more complicated dynamic computer models. Still, I think there is an important role for baby models such as this one, which give some direct intuition that may be considerably more transparent than what emerges from simulations of much more complicated formulations.

2   Uncertain Equilibrium Temperatures

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 on the uncertainty of so-called "equilibrium climate sensitivity." This is a relatively well-defined and relatively well-studied example of known unknowns, even if the uncertainties themselves are uncertain. However, it should be clearly understood that under the rubric of "equilibrium climate sensitivity" I am trying to aggregate together an entire suite of 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.

The uncertainty in climate sensitivity translates into uncertainty about ultimate temperatures, which, at the end of a long chain of further uncertainties, translates finally into uncertainty about the loss of future welfare. The paper follows only a few crude links in this long chain and makes a few extremely rough ballpark estimates of what all of this might mean

for an economic analysis of climate change. Even ignoring the economic part altogether, determining probabilistic climate change for future emission scenarios is challenging, as it requires a synthesis of uncertainties along the cause-effect chain from emissions to climate change via uncertainties in the carbon cycle, radiative forcing, temperature responses, and so forth. When economic uncertainty about how to evaluate extreme climate-change damages is tacked on, one is dealing with an integrated assessment problem whose specification is extraordinarily unsure. Empirically, it is not the fatness of the tail of temperature PDFs *alone* or the reactivity of the damages function to high temperatures *alone*, or any other factor *alone*, that counts, but the *combination* of all such factors. In this paper, the default fatness that really matters concerns the bad tail of the reduced-form PDF of welfare-equivalent consumption (which includes damages) – *not* climate sensitivity per se.

"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 mean surface warming that follows a sustained doubling of atmospheric carbon dioxide (CO_2), after the climate system has reached a new equilibrium. 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 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.[1] 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. However, 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 very closely the spirit and assumptions (and drawbacks) of the S-reductionist approach.

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

---

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

a probability above 66%. In this paper I choose 70% as defining "likely" and I calibrate all upper-tail probability distributions so that P[$S \geq$3°C]=.5, and P[$S \geq$4.5°C]=.15.

The upper-half tail of the probability distribution is the region $S > S_\mathbf{M}$, whose total probability mass is .5, where the climate-sensitivity median is $S_\mathbf{M}$=3°C. I use three 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 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 exponentially but faster than polynomially.[2] For all three PDFs I calibrate parameters so that P[$S \geq$3]=.5 and P[$S \geq$4.5]=.15. A major goal of this paper is to experiment with different PDFs above the median value of $S_\mathbf{M}$=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 about the PDF for the 50% of probability *above* the median, and, as we shall see, this can have major consequences.

The notation $f_\mathbf{I}(S)$ refers to the PDF of climate sensitivity S. The subscript $\mathbf{I}=\mathbf{L}$ refers to a **L**ognormal PDF, the subscript $\mathbf{I}=\mathbf{N}$ refers to a PDF whose upper-half tail is **N**ormal, and the subscript $\mathbf{I}=\mathbf{P}$ refers to a distribution whose upper-half tail is **P**areto (or **P**ower). I begin with the base case of the Lognormal, whose upper-half PDF here is

$$f_\mathbf{L}(S) = \frac{1}{.391212\sqrt{2\pi}\,S}\exp\left(-\frac{(\ln S - 1.098612289)^2}{2(.391212)^2}\right) \tag{1}$$

for all $S \geq 3$. As can readily be confirmed, the parameter values in (1) have been calibrated so that P[$S \geq$3°C]=.5 and P[$S \geq$4.5°C]=.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 specified by its parameters being set so that simultaneously $P_\mathbf{P}$[S≥3]=.5 and $P_\mathbf{P}$[S≥4.5]=.15. It is readily confirmed that the corresponding upper-half-tail Pareto PDF is

$$f_\mathbf{P}(S) = 38.75958\,S^{-3.969362}. \tag{2}$$

My upper-half-tail Normal PDF is also specified by its two parameters being set so that

---

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

simultaneously $P_{\mathbf{N}}[S{\geq}3]{=}.5$ and $P_{\mathbf{N}}[S{\geq}4.5]{=}.15$. It is readily confirmed that the corresponding upper-half tail Normal PDF is

$$f_{\mathbf{N}}(S) = \frac{1}{1.4473\,\sqrt{2\pi}}\,\exp\left(-\frac{(S-3)^2}{2\,(1.4473)^2}\right). \tag{3}$$

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

$\widehat{S} =$	3°C	4.5°C	6°C	8°C	10°C	12°C	15°C	20°C
$P_{\mathbf{P}}[S \geq \widehat{S}]$	.5	.15	.064	.027	.014	.008	.004	.002
$P_{\mathbf{L}}[S \geq \widehat{S}]$	.5	.15	.038	.006	.001	$2{\times}10^{-4}$	$2{\times}10^{-5}$	$6{\times}10^{-7}$
$P_{\mathbf{N}}[S \geq \widehat{S}]$	.5	.15	.019	.003	$7{\times}10^{-7}$	$3{\times}10^{-10}$	$6{\times}10^{-17}$	$4{\times}10^{-32}$

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

I think that not many climate scientists would quibble about the "big picture" of the PDF of climate sensitivity given by Table 2 for low values of climate sensitivity. For what it is worth, the median upper five percent probability level over all 22 climate-sensitivity PDFs cited in IPCC-AR4 is 6.4°C, which fits with the Pareto PDF above.[3] Notice that the probabilities of very high values of S are quite small. Even so, the relative probabilities of high S are extremely dependent on whether the upper tail of the relevant PDF is fat, thin, or intermediate.

It is tempting to say that climate sensitivity above, say, 15°C is "impossible." I would prefer to think that anything is possible under the novel experiment of (geologically instantaneously) doubling atmospheric CO_2 concentrations in a situation where so many unknowns are so highly uncertain. Take the lognormal PDF as a base case. I am not sure how one would distinguish operationally here between a very rare event that is (almost) "impossible" and a very rare event that has a one in five thousand chance of materializing. Such fine distinctions can be ignored in most applications, and the analysis can proceed as if the event is "impossible" for all practical purposes. But in the extraordinary case of global warming, whose potential damages could engulf the entire planet, one does not have the luxury of ignoring even the lowest of low-probability events if they occur with the highest of highly negative impacts.

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

---

[3]Details in Weitzman (2009a).

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. 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 $f_{\mathbf{I}}(S)$ is the relevant PDF of climate sensitivity, then the relevant PDF of temperatures T for a given level of G is

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

To anchor the upper tail of extreme temperatures, I focus sharply on just two iconic (if arbitrary) values of extraordinarily high mean global temperatures: 6°C and 12°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.[4] One recent study[5] asked 52 experts for their subjective probability estimates of triggering a "tipping point of 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, the *expected* (probability weighted) number of such expert-assessment tipping points was three (out of a possible five) at an average temperature increase of $T \approx 6$°C.

Twelve degrees of global warming is used here as an example of a round number (12°C=2×6°C) that transcends our ability to imagine, with any reasonable measure of accuracy, what the earth might be like for super-high temperature increases. For me, 12°C is especially iconic because of a recent study, which estimated that global average temperature increases of

---

[4] See, e.g., IPCC-AR4 (2007) and Lynas (2007).
[5] Kriegler et al, (2009).

≈11-12°C would cause conditions under which 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 likely ensue after about six hours of exposure.[6] The authors of this study furthermore point out: "This likely overestimates what could practically be tolerated: our limit applies to a person out of the sun, in a gale-force wind, doused with water, wearing no clothing and not working." A temperature change of ≈12°C therefore represents an extreme threat to human civilization as we now know it, even if it does not necessarily mean the end of Homo sapiens as a species.

Throughout the numerical examples that follow, I arbitrarily take 18°C (3×6°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 like 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 ozone, sulfur dioxide, and particulates. 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 the wildcard of the behavior of methane clathrates, 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

---

[6]Sherwood and Huber (2010).

the peak enhancement level over the preindustrial level of 280 ppm persevered after 1,000 years of *zero* emissions. This means, for example, that if atmospheric CO_2 concentrations were to peak at 800 ppm, followed forever thereafter by *zero* emissions, then atmospheric concentrations would be $\approx$650 ppm after 100 years and $\approx$500 ppm after 1,000 years. These numbers do not look to me like evidence supporting "wait and see" policies. The capacity of the oceans to take up atmospheric heat, and many, 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. Under such conditions of limited learning relative to reversibility, the fact that the small probabilities of big disasters are themselves uncertain is not an excuse for delay. Just the opposite, it is a stronger call to immediate action than if the probabilities were known precisely.

In the following Table 2, the first row represents steady-state atmospheric stocks of greenhouse gas concentrations G (measured in ppm of CO_2e). The second row below it gives the median equilibrium temperature $T_{\mathbf{M}}$ as a function of stabilized GHG stocks. The rows starting just below $T_{\mathbf{M}}$ 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 (Pareto=fat tail, Lognormal=intermediate tail, Normal=thin tail).

G :	400	450	500	550	600	650	700	750	800	900	1000
$T_{\mathbf{M}}$:	1.5°	2.1°	2.5°	2.9°	3.3°	3.6°	4.0°	4.3°	4.5°	5.1°	5.5°
$P_{\mathbf{P}}[T \geq 6]$	.9%	2%	4%	6%	8%	11%	15%	18%	22%	30%	39%
$P_{\mathbf{L}}[T \geq 6]$	10^{-4}	.3%	1%	3%	6%	10%	14%	19%	24%	33%	41%
$P_{\mathbf{N}}[T \geq 6]$	10^{-9}	10^{-5}	.2%	1%	4%	9%	14%	20%	25%	35%	43%
$P_{\mathbf{P}}[T \geq 12]$	.11%	.26%	.48%	.75%	1.1%	1.5%	1.9%	2.3%	2.8%	3.8%	5.0%
$P_{\mathbf{L}}[T \geq 12]$	10^{-7}	10^{-6}	10^{-5}	.02%	.05%	.12%	.23%	.41%	.65%	1.4%	2.3%
$P_{\mathbf{N}}[T \geq 12]$	10^{-45}	10^{-24}	10^{-15}	10^{-10}	10^{-8}	10^{-6}	10^{-5}	.08%	.03%	.22%	.73%
$P_{\mathbf{P}}[T \geq 18]$	.03%	.08%	.14%	.23%	.32%	.44%	.56%	.69%	.84%	1.2%	1.5%
$P_{\mathbf{L}}[T \geq 18]$	10^{-10}	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-5}	.01%	.01%	.02%	.06%	.12%
$P_{\mathbf{N}}[T \geq 18]$	10^{-108}	10^{-58}	10^{-37}	10^{-26}	10^{-20}	10^{-16}	10^{-13}	10^{-11}	10^{-10}	10^{-7}	10^{-6}

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

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. From Table 2, the target level of GHG concentrations influences strongly the probabilities of 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 moderate rise of $T_\mathbf{M}$ 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 prevents temperatures from rising much above 6°C, whereas 1000 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 $\geq$650 ppm of CO_2e, a temperature increase of 6°C is sufficiently close to the middle range of all three climate-sensitivity PDFs that tail fatness per se does not matter so much in determining P[$T \geq$6]. On the other hand, tail fatness always matters a lot for determining P[$T \geq$12], even for higher GHG concentrations $\geq$650, because this part of the range of temperatures is well into the extreme tail of the underlying climate-sensitivity PDFs.

Table 2 gives the impression that the primary purpose of keeping down G may be to prevent possibilities of extreme temperatures in the upper range of the PDFs, and perhaps only secondarily to keep down the median temperature (although this may be important too). In this sense Table 2 is indicating indirectly how much "insurance" society is willing to buy to ward off the risk of very high temperatures by paying the "cost" of keeping GHG concentrations below various levels. I do not analyze explicitly the costs of achieving various steady-state GHG targets, being content to let them stand for themselves as proxies for less or more active mitigation measures. Only the "value" or demand side of insurance (against

high-temperature extreme damages) is being presented (and that indirectly), not its "cost" or supply side. Throughout this paper, target steady-state GHG concentrations are interpreted as an imperfect proxy for "policy."

If one wants a transparent summary of the temperature consequences of higher GHG concentrations, I think that Table 2 is fine. Perhaps the analysis should be ended here, as the table speaks for itself quite eloquently in ways that an informed citizen might understand once the possible consequences of the "iconic" values $T=6°C$ and $T=12°C$ have been explained, as they were before. But an economist is tempted to take the analysis at least one step further toward quantifying damages before succumbing to a computer simulation of a full-blown dynamic IAM with lots of non-transparent moving parts.

For any given G, and for any given PDF of S, we have derived a PDF of T (via (4) and (6)), some numerical values of which are displayed in Table 2. Of course such type of analysis ignores all kinds of dynamics to concentrate on the more-easily-understandable long run value of T associated with a sustained stationary level of G. In the coming extension to potential damages, I will follow closely the spirit and assumptions of the approach to temperature PDFs of this section. Thus, throughout this paper the simplistic methodology looks primarily at the "big picture" of a still photograph of damages in a steady state, with only the most primitive story (later) about dynamics. I do not investigate seriously the important subject of the motion picture describing the dynamics of getting to the steady state. I only push the comparative-steady-state envelope slightly further by letting in the sticky subject of high-temperature damages. There is an artificial timing here that compresses dynamics into statics by dealing primarily with steady states. Such an aggregate comparative-steady-state approach has proved to be a useful shortcut way for organizing thinking about eventual temperature responses to target GHG concentrations. It tells us in shorthand form what temperatures (more accurately temperature PDFs) we are eventually buying into when we set target GHG concentrations at various levels. I propose extending the same strategy a little further to at least discuss the possible present discounted damages from large temperature changes.

3   Uncertain Damages in Climate Change

From the very outset, the representation of damages from climate change presents some severe conceptual and practical problems. I follow most of the literature by postulating that damages from increased temperatures are manifested in reduced form *as if* they impair output.[7]   In my version of this just-so story, all losses from climate change will be inter-

---

[7]This interpretation could be challenged. See Weitzman (2009b).

preted as if they literally translate into a welfare-equivalent loss of consumption. There are some genuine doubts about what it means operationally to separate welfare-equivalent consumption from welfare, but here I largely follow the existing literature. As mentioned, this paper examines only the damages side, and that very simplistically. I do not try to explicitly estimate costs of achieving various GHG targets, much less attempt to determine an optimal policy by explicitly balancing the costs of achieving a given GHG target against its benefits.

Even granted that it multiplicatively diminishes consumption, no one knows how to specify a "damages function" for high temperature changes. 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 damages due to changes in agricultural productivity, changes in climate, changes in life styles, 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 the most extreme example I can imagine for making my point, suppose for the sake of argument that average global temperatures 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 mean 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. There is just too much structural uncertainty to put meaningful bounds on the unprecedented almost-unimaginable changes to planetary welfare from average global temperatures increasing by 12°C. I don't think anyone knows how to evaluate the welfare-equivalent "damages" from super-high average global temperatures, but a global temperature rise of ≈12°C has a good chance of going far beyond an absolute heat-stress limit that could extinguish many mammals on earth and impair very severely human functioning.[8]

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 time* when the RV T materializes. (This concept itself has some problematical aspects, which are ignored here, although my intention is that such intangibles as loss of the environment as we know it are somehow included in "welfare equivalent" consumption.) The most popular single formulation of a damages function in the literature is of the quadratic

---

[8]This is my own very loose translation of the study of Sherwood and Huber (2010).

form

$$\widetilde{C}_Q(T) = \frac{1}{1 + \alpha T^2},\tag{7}$$

where α is a positive parameter calibrated to give some "reasonable" values of $\widetilde{C}_Q(T)$. Most 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 (7) to conform with the damages function in the latest version of the well known DICE model of William Nordhaus (2008), where he used $\alpha{=}.002388$.[9] 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.)

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%	.2%	.1%

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

I do not find such numbers as $\widetilde{C}_Q(T)$ in Table 3 at all convincing for high temperatures. At an unimaginably high average global temperature change of $T{=}18°$C, welfare-equivalent consumption damages are "only" 44% *at that time* (when $T{=}18°$C materializes). 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 as $\widetilde{C}_Q$ in the second row of Table 3, is preordained to make extreme climate change look empirically negligible almost no matter what else is assumed.

Of course I have no objective way to determine magnitudes of high-temperature damages, but the last time that the world witnessed average temperatures very roughly $\approx8°$C or so

---

[9]Nordhaus's DICE model is perhaps the most famous integrated assessment model (IAM) in the economics of climate change. The value $\alpha = .002388$ 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.

above the present was during the Paleocene-Eocene boundary $\approx$55 mya. In this period the earth was ice free, the land was swampy, and alligators lived near the North Pole. This was also the last time that there was a *geologically rapid* increase in average global temperatures of magnitude very roughly $\approx$5°C or so above an already warm background, and it was accompanied by mass extinctions. This "PETM" event occurred geologically rapidly over a period of about 20K years, which is extremely gradual compared with current worst-case anthropogenically-induced trajectories, and it was followed by a recovery period that lasted about 150K years. It is unknown what triggered the PETM, but the leading 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 so if current emissions trends are extrapolated.[10] The major point here is that relatively rapid changes of global average temperatures $\approx$5°C above present values are extremely rare events extraordinarily far outside the scope of human experience. As for huge temperature increases like $T \approx 12°$, the planetary effects are difficult to imagine. To find a geologically instantaneous increase in temperatures of magnitude $T \approx 12°$, one would perhaps have to go back hundreds of millions of years. Others are free to calibrate any welfare-equivalent consumption loss they want in the range above $T \approx 4°$, as anybody's guess here is as good as mine. I don't think that a person needs accurate specific stories about what might happen for $T > 12°$ to imagine truly upending catastrophes undoing the planet and severely undermining the security of human civilization – at least.

I now want to "give the devil his due" by characterizing very roughly two points on a much more reactive damages function, which seems to me 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 two iconic (if arbitrary) global-average temperature changes: 6°C and 12°C. What these two iconic temperature changes might mean for the human condition and for the rest of the planet was already sketched out in the last section. 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 $\widetilde{C}_R(12°C)=1\%$.

Some IAMs and CBAs recommend a policy ramp gradualism that would approach atmospheric CO_2 levels of $\approx$700 ppm, which would arguably make GHG CO_2e levels $\approx$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$1% (depending very much on how fat-tailed is the relevant PDF). Using the proposed reactive specification of damages ($\widetilde{C}_R(6°C)=50\%$ and $\widetilde{C}_R(12°C)=1\%$), I calculated

---

[10]For more about methane clathrates, see Archer (2007) or the recent article by Shakova et al (2010).

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 (7) having the form βT^γ, so that in place of (7) we have instead a "reactive" damages function of form

$$\widetilde{C}_R(T) = \frac{1}{1 + \alpha T^2 + \beta T^\gamma}, \tag{8}$$

where (as before) α=.002388, while I calibrated β and γ so that $\widetilde{C}_R(6°C)$=50% and $\widetilde{C}_R(12°C)$=1%. The relevant parameter values are β=5.075$\times 10^{-6}$ and γ=6.754. It is readily confirmed from Table 3 that $\widetilde{C}_R(T)$ is indeed more "reactive" to higher temperature changes than $\widetilde{C}_Q(T)$. As mentioned, global mean temperatures are arbitrarily forbidden from going above $\overline{T}$=18°C, which corresponds in Table 3 to $\widetilde{C}_R(18°C)$=.1%.

4   Putting the Pieces Together

From equation (5), steady-state temperatures 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 RV with some postulated PDF, then (for any given G) T is a RV with PDF given by (6). And then, given some postulated damages function of temperature (namely $\widetilde{C}_Q(T)$ or $\widetilde{C}_R(T)$), welfare-equivalent consumption in that steady state is itself a RV. I now manufacture an artificial example 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 (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.[11] Taken together, quotations from these "famous economists" sound to 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 best estimate of most economists. My base-case future growth rate of per capita consumption is $g=2\%$ per year.[12] These base-case values imply 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 (10) with $\rho=0$ implies $g=3\%$ – probably too high. If $\eta=4$ and $g=2\%$, then (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.

For my base case I use the lognormal PDF of S (with its intermediate tail fatness), as given by (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 $\widetilde{C}_R$). After this permanent shock to the level of consumption 150 years in the future, growth continues thereafter at annual rate $g=2\%$. This is a primitive formulation, but I think it is good enough to make the point that a

---

[11] 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 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.

[12] These values of $\eta=3$ 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).

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 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 $\exp(-.06 \times 150) = \exp(-9) = .01\%$, which is a very, 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

$$\int_0^\infty U(\widehat{C}\exp(.02t))\,dt = \int_0^{150} U(\exp(.02t))\,dt + E\left[\int_{150}^\infty U(\widetilde{C}_J\exp(.02t))\,dt\right], \qquad (11)$$

where $J{=}Q$ or $J{=}R$.

Now substitute the lognormal PDF (1), (6) (for **I=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

$$\widehat{C}_J = \left(1 - e^{-6} + e^{-6}E\left[\left(\widetilde{C}_J\right)^{-2}\right]\right)^{-.5}, \qquad (12)$$

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.

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

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 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 1000 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 now to avoid significantly higher GHG concentrations a century and a half from now. No wonder, then, that optimal

IAM trajectories derived from a quadratic damages function encourage gradualist policy-ramp CO_2e levels approaching 750 ppm!

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 willingness to pay, in terms of certainty-equivalent consumption now and for all time, 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 loss of as-if-deterministic consumption accompanying a change in GHG concentrations from 550 ppm to 750 ppm by Table 4 is $(.976 - .579)/.976 = .407$ – i.e., keeping target GHG levels at 550 ppm rather than letting them rise to 750 ppm is worth spending 40.7% of present certainty-equivalent consumption at $G=550$. Or, to take another example, the welfare-equivalent change in consumption accompanying a reduction in GHG concentrations from 650 ppm to 600 ppm in Table 4 is $(.924 - .827)/.924 = .105$ – i.e., lowering target GHG levels from 650 to 600 is worth spending 10.5% of present certainty-equivalent consumption at $G=650$. Such high losses indicate a large willingness to pay, in terms of present certainty-equivalent consumption, to avoid the consequences of elevated GHG concentrations 150 years from now. Notice, though, that the willingness to pay to keep GHG concentrations below ≈ 500 ppm of CO_2e are very 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 .01%) takes care of the lower temperatures. Above ≈ 550 ppm of CO_2e, however, the danger of higher temperatures accelerates greatly the willingness to pay 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. Again, the impression is that GHG policy is most accurately viewed as an insurance policy against catastrophic 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 because it is an upper limit on optimal GHG concentrations that is approached by some optimizing IAMs and CBAs.

Generally speaking, outcomes are very dependent on how extreme tail damages and extreme tail probabilities are formulated. 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 in the paper. 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. 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 willingness to pay 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 traditional role of η. Through the 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 $\exp(-.02\times150\times4)=6.1\times10^{-6}$; for $\eta=3$ it is $\exp(-.02\times150\times3)=1.2\times10^{-4}$; and for $\eta=2$ it is $\exp(-.02\times150\times2)=2.5\times10^{-3}$. This is a very wide range for discount factors, although all of these numbers are extremely low. However, higher values of η also indicate higher relative risk aversion, which can easily 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 easily sufficient to reverse strongly the traditional role of η, because the effect of aversion to catastrophic uncertainty here easily 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 $S_{\mathbf{M}}=3$, 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 $S_{\mathbf{M}}=3$, then welfare-equivalent consumption at 750 ppm is 81% instead of 58% in Table 4. The willingness to pay 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 global warming arrives at $\tau=150$ years from now. If the time of arrival for global warming is $\tau=200$ 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 $\tau=100$ 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 mean 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.

case	CRRA η	PDF	impact yr 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.

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

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.[13]

As yet another example of possible fragility, consider the base case in Nordhaus (2008) (ρ=1.5%, η=2, g ≈2%), which from (10) corresponds to a discount rate r=5.5% per year. When Nordhaus runs instead through his DICE model "my" base case (ρ ≈0, η=3, g ≈2%), which implies r=6% per year, there is no substantive difference in outcomes. However,

---

[13]The study of Costello et al (2010) uses $\exp(-\alpha T^2)$ in place of $1/(1 + \alpha T^2)$. With α=.002388, there is hardly any difference in implied damages, even for temperature changes as high as 18°C. My tentative conclusion is that their model, which uses the form $\exp(-\alpha T^2)$, is equally non-robust to the stress-test numerical exercises of this paper, which uses the form $1/(1 + \alpha T^2)$.

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 usual $\widehat{C}_Q \approx 100\%$. Even with my reactive damages function (8), if I plug the Nordhaus base-case specification (ρ=1.5%, η=2, g=2%) into my model I get $\widehat{C}_R \approx 99\%$ for G=750 ppm of CO_2e. In other words, the willingness to pay to avoid altogether the consequences of G=750 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 ρ=1.5%, 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 values of ρ and η.

5   Discussion

I think that several themes emerge from this paper.

The paper suggests that 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 greenhouse gases, 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 temperatures 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 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 extra-ordinarily 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 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 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.

To summarize, there is an underlying generic problem with CBAs or IAMs of climate change that is not present in other, more standard, applications of CBA. In rare situations with effectively unlimited downside liability, like climate change, CBAs or IAMs can be extraordinarily sensitive to specifications of extreme tail events. Conventional CBAs or IAMs of climate change ignore this core message at their own peril.

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 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 any 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.[14]

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 welfare-equivalent

---

[14] 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.

consumption (which includes damages). I somehow doubt very much that robustness will miraculously be restored by introducing more sophisticated dynamics. The various "stress tests" of this paper do not strike me as "stressful" at all. Therefore, I suspect rather strongly that it may be difficult to dislodge altogether the verdict that a CBA of climate change is terribly sensitive to assumptions about extreme tail events – and that the primary reason for keeping GHG levels down is 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. 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

[1] Archer, David (2007). *Global Warming.* Blackwell Publishing.

[2] Arrow, Kenneth J. (1999). "Discounting, Morality, and Gaming" in Portney and Weyant (eds). *Discounting and Intergenerational Equity.* Resources for the Future.

[3] Costello, Christopher, Michael Neubert, Stephen Polasky, and Andrew Solow (2010). "Bounded uncertainty and climate change economics." *Proceedings of the National Academy of Sciences*: vol 107, no. 18, 8108-8110.

[4] Dasgupta, Partha. (2008). "Discounting climate change." *Journal of Risk and Uncertainty*, 37: 141-169.

[5] IPCC-AR4. 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 University Press, 2007 (available online at http://www.ipcc.ch).

[6] Knutti, Reto, and Gabriele C. Hegerl. (2008) "The equilibrium sensitivity of the Earth's temperature to radiation chages." *Nature Geoscience*, Vol. 1 November]: 735-743.

[7] Kriegler, Elmar, Jim W. Hall, Hermann Held, Richard Dawson, Hans Joachim Schellnhuber. "Imprecise probability asessment of tipping points in the climate system." *Proceedings of the National Academy of Sciences*, 106, 5041-5046 (31 March, 2009).

[8] Lynas, Mark (2007). *Six Degrees: Our Future on a Hotter Planet.* HarperCollins.

[9] Nordhaus, William D. (2008) *A Question of Balance.* Yale University Press.

[10] Shakhova, Natalia, Igor Semiletov, Anatoly Salyuk, Vladimir Yusupov, Denis Kosmach, Örjan Gustafsson (2010). "Extensive Methane Venting to the Atmosphere from Sediments of the East Siberian Arctic Shelf." *Science*, 327, 5970, pp. 1246 - 1250.

[11] Sherwood, Steven C., and Matthew Huber (2010). "An adaptability limit to climate change due to heat stress." *Proceedings of the National Academy of Sciences*: May 6, 2010 (www.pnas.org/cgi/doi/10.1073/pnas.0913352107).

[12] Solomon, Susan, Gian-Kasper Plattner, Reto Knutti, and Pierre Friedlingststein (2009). "Irreversible climate change due to carbon dioxide emissions." *Proceedings of the National Academy of Sciences* 106 (6): 1704-1709.

[13] Weitzman, Martin L. (2009a). "On Modeling and Interpreting the Economics of Catastrophic Climate Change." *Review of Economics and Statistics*, 91 (1) [February]: 1-19.

[14] Weitzman, Martin L. (2009b). "Additive Damages, Fat-Tailed Climate Dynamics, and Uncertain Discounting." *Economics e-journal*, 3, 2009-39 (October 22).