8 Some dynamic economic consequences of the climate-sensitivity inference dilemma

Martin L. Weitzman

1 INTRODUCTION

Equilibrium climate sensitivity is a key parameter that serves as a very useful macro-indicator of the eventual aggregate response of temperature change to the aggregate level of greenhouse gases (GHGs). Decades of scientific research have failed to constrain the upper range of climate sensitivity, which means that alarmingly high distant-future temperature responses are not excluded. This chapter highlights a generic statistical-inference mechanism that makes it difficult to thin down to zero inherently fat-tailed probability estimates of rare extreme outcomes.

Let $\Delta \ln CO_2$ be sustained relative change in concentrations of atmospheric carbon dioxide while ΔT is equilibrium mean global surface temperature response. Equilibrium climate sensitivity (here denoted λ) is a benchmark amplifying or scaling multiplier for converting $\Delta \ln CO_2$ into ΔT by the (reasonably accurate) linear approximation $\Delta T \approx (\lambda / \ln 2) \times \Delta \ln CO_2$. As the Intergovernmental Panel on Climate Change in its IPCC-AR4 (2007) Executive Summary phrases it:

> 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 (italics in original).

In this chapter I am mostly concerned with the roughly 15 per cent of those λ 'values substantially higher than 4.5° C' which 'cannot be excluded'. A grand total of 22 peer-reviewed studies of climate sensitivity published recently in reputable scientific journals and encompassing a wide variety of methodologies (along with 22 imputed probability density functions (PDFs) of λ) lie indirectly behind the above-quoted IPCC-AR4 (2007) summary statement; they are of that document in Table 9.3 and Box 10.2 listed. It might be argued that these 22 studies are of uneven reliability and their complicatedly-related PDFs cannot easily be combined, but for the simplistic purposes of this illustrative example I do not perform any kind of formal Bayesian model-averaging or meta-analysis (or even engage in informal cherry picking). Instead I just naively assume that all 22 studies have equal credibility and, for my purposes here, that their PDFs can be simplistically aggregated. The upper 5 per cent probability level averaged over all 22 climate-sensitivity studies cited in IPCC-AR4 (2007) is 7° C, while the median is 6.4° C,[1] which I take as signifying approximately that $P[\lambda > 7° C] \approx 5\%$. Looking at the upper tails of these 22 PDFs, one might roughly presume from a simplistically-aggregated PDF of these 22 studies that $P[\lambda > 10° C] \approx 1\%$. Even if my numbers are somewhat off, it still seems apparent that the upper tails of the PDFs of λ appear to be long and fat. A fat-tailed PDF assigns a relatively much higher probability to rare events in the extreme tails than does a thin-tailed PDF.[2] Although both limiting probabilities are infinitesimal, the ratio of a thick-tailed probability divided by a thin-tailed probability approaches infinity in the limit.

A critical question is this. Why, after decades of extensive research, do these upper tails of climate sensitivity PDFs seem so intractably long and fat? The climate science literature appears to have coalesced around an answer along the lines that there might be some physical basis for believing that $1 / \lambda$ is approximately normally distributed, which would make λ itself have something like a fat-tailed Cauchy-Lorentz PDF. As a story about feedback processes and measurement errors, some of this logic can make sense and it provides some much-needed insight into the physical nature of an issue that is crucially important for understanding a key driver of global warming. However, as a story about statistical inference, this reasoning is partial and incomplete in the sense that it appears to rely on a very specific data generating process (DGP) that is not fully rigorously specified – nor does this story pinpoint formally what exactly is the relevant problem here of prediction under uncertainty. In this chapter, I suggest that the core logic behind a fat-tailed PDF of climate sensitivity perhaps transcends the underlying physics and may even be more generic than somewhat partial reasoning about the properties of ratios of random variables in a particular DGP. I argue that the relevant posterior-predictive PDF of virtually *any* high-impact low-probability rare event has a deeply built-in tendency to be fat tailed – almost irrespective of the underlying DGP. When these fat tails matter because catastrophic damages have essentially unlimited liability – as with climate change – this aspect is capable of driving the economic analysis.

The essence of the climate sensitivity dilemma highlighted in this chapter is the difficulty of learning extreme-impact tail behavior from finite data alone. Loosely speaking, the driving mechanism is that the operation of taking 'expectations of expectations' or 'probability distributions of probability distributions' spreads apart and fattens the tails of the reduced-form compounded posterior-predictive PDF. It is inherently difficult to determine extreme bad-tail probabilities from finite samples alone because, by definition, we don't get many data-point observations of infrequent events. Therefore, rare disasters located in the stretched-out fattened tails of such posterior-predictive distributions must inherently contain an irreducibly large component of deep structural uncertainty. The underlying sampling-theory principle is that the rarer an event, the more unsure is our estimate of its probability of occurrence – and the larger the sample size required to rule it out for practical purposes. In this spirit (from being constructed out of inductive knowledge), the empirical studies of climate sensitivity are perhaps pre-ordained to find the fat-tailed power-law-like PDFs which they seem, approximately, to find in practice.

Climate sensitivity is not nearly the same thing as temperature change. In previous work,[3] I tried to fudge the distinction by arguing that the shapes of both PDFs are very roughly similar because a doubling of anthropogenically injected CO_2-equivalent greenhouse gases (GHGs) relative to pre-industrial-revolution levels is essentially unavoidable within about the next 40–50 years and the GHGs are very likely to remain well above this level for at least the subsequent century or so after first attaining it. But such a discrete two-period formulation suppresses the continuous-time dimension by ignoring the fact that higher ΔT values are (continuously) correlated with later times of arrival. This chapter addresses more centrally the relationship between dynamic ΔT trajectories and climate sensitivity. I show that a previous two-period result, that fat-tailed climate sensitivity can have strong economic implications, survives being recast into a more complete dynamic specification, even though (other things being equal) the higher the temperature realization, the later this temperature realization is expected to arrive. When fat climate-sensitivity tails are combined with very uncertain high-temperature damage, this aspect can dominate the discounting aspect in calculations of expected present discounted utility – even at empirically plausible real-world interest rates and even when taking full account of the important continuous correlation that, conditional upon its realization, the higher the temperature, the later is its expected time of arrival.

A central theme of this line of research is that with finite data, it is practically inevitable that rare extreme events will have fat tails. These fat tails, which stubbornly resist the accumulation of finite data, reflect back at us our own prior ignorance concerning how to model or represent or quantify rare extreme events. My message is that we must learn to live with the idea that the answers to cost–benefit analyses of what to do about climate change may very well depend – at least to some degree – upon subjective judgments about how bad it might get, with what probabilities, in the most extreme situations.

2 A DYNAMIC AGGREGATIVE MODEL OF GLOBAL WARMING

This section compresses into a single differential equation what is arguably the simplest meaningful dynamic model of the physical process of global warming. Of course this particular one-equation model cannot possibly capture the full complexity of climate change. However, I think the highly aggregated approach taken here is realistic enough to serve as a springboard for meaningful discussions of some basic climate change issues which, for the purposes of this chapter, are actually clarified when tightly framed in such stark simplicity.

Perhaps the single most useful concept for understanding the process of climate change is that of radiative forcing. GHGs such as CO_2 are prime examples of this, but there are many others, such as solar intensity, aerosols and particulates. (The radiative forcing from CO_2 happens to be proportional to the logarithm of its atmospheric concentration, but this is not true in general for all forcings.) A key property of radiative forcings is that the various components and subcomponents can be aggregated simply by adding them all up because they combine additively.

Suppose for simplicity that throughout times $t < 0$, the planetary climate system has been in a state of long-run equilibrium at a constant temperature with constant radiative forcing. Imagine that starting at time $t = 0$, and continuing throughout times $t \geq 0$ a sustained perturbation (relative to times $t < 0$) of radiative forcing of constant magnitude ΔR_f has been additionally imposed. (Whether this constant additional radiative forcing ΔR_f is itself exogenous or endogenous is irrelevant in this context because only the reduced-form total forcing matters here.) If the earth were a black body planet with no atmosphere and no further feedbacks, the long-run temperature response would be $\Delta T \rightarrow \lambda_0 \Delta R_f$, where λ_0 is the feedback-free equilibrium-climate-sensitivity constant defined by the fundamental physics of a black-body reference system described by the Stefan-Boltzman law. Even in richer, more realistic situations with feedbacks and complicated dynamics, other things being equal, it is not a bad assumption that at any time t the temperature moves with an instantaneous velocity approximately proportional to $\lambda_0 \Delta R_f(t) - \Delta T(t)$ – that is, the (linearized) basic equation of temperature motion is

$$\dot{T} = \frac{1}{h}[\lambda_0 \Delta R_f(t) - \Delta T(t)]. \tag{1}$$

The positive coefficient h in (1) represents the thermal inertia of the system, in this application primarily standing for the overall planetary capacity of slab-like oceans to take up heat.

The full temperature dynamics of this idealized planetary system can then most simply be described as follows. At time $t \geq 0$, suppose that a system previously in long-run equilibrium is now subjected to an exogenously imposed radiative forcing of $\Delta F(t)$. Let the total change in radiative forcing at time $t \geq 0$ be denoted $\Delta R_f(t)$. If the endogenously induced radiative forcing at time $t \geq 0$ is denoted $\Delta R_i(t)$, then

$$\Delta R_f(t) = \Delta F(t) + \Delta R_i(t). \tag{2}$$

In a simple linear feedback system applied to the problem at hand, the temperature change $\Delta T(t)$ causes a comparatively fast-acting (relative to (1)) endogenous feedback response on induced radiative forcing $\Delta R_i(t)$ according to the formula

$$\Delta R_i(t) = \frac{f}{\lambda_0} \Delta T(t), \tag{3}$$

where the (linear) feedback factor f is a basic parameter of the system. The relevant feedback factors in climate change involve cloud formation, water vapor, albedo, among others. A key property of linear feedback factors is that (as with radiative forcing) the various components and subcomponents can be aggregated simply by adding them all up because they combine additively.

Plugging (3) and (2) into (1) yields after simplification the fundamental differential equation

$$\dot{T} = \frac{1}{h}[\lambda_0 \Delta F(t) - (1 - f) \Delta T(t)], \tag{4}$$

whose solution is

$$\Delta T(t) = \frac{\lambda_0}{h} \int_0^t \Delta F(s) \exp\left(-\left(\frac{1 - f}{h} \right) (t - s) \right) ds. \tag{5}$$

The oversimplifications of physical reality that have gone into the one-equation temperature change trajectory (5) are too numerous and too tedious to recount here. There is only one major defence of this ultra-macro approach: it seems fair to say that (5) captures the dynamic interplay of forces along a global-warming path better than any alternative single-differential-equation formula. If we want a sharply focused formulation of the big moving picture of a dynamic global-warming trajectory in terms of its primary contributing ingredients, then we are pretty much stuck with (5).

In what follows, it will be analytically convenient to work with the special case where exogenously imposed radiative forcing is constant, so that for all times, $t \geq 0$,

$$\Delta F(t) = \bar{F}, \tag{6}$$

which simplifies (5) into

$$\Delta T(t) = \frac{\lambda_0}{1-f}\bar{F}\left[1 - \exp\left(-\left(\frac{1-f}{h}\right)t\right)\right]. \tag{7}$$

For notational neatness, assume that all units are expressed in terms of a doubling of CO_2. The equilibrium climate sensitivity is then defined as

$$\lambda \equiv \lim_{t \to \infty} \frac{\Delta T(t)}{\bar{F}}, \tag{8}$$

and it is readily apparent from applying (8) to (7) that

$$\lambda = \frac{\lambda_0}{1-f}, \tag{9}$$

which is one of the most basic relationships of climate change.

Even accepting the enormous oversimplifications of reality that go into an equation like (5) (or (7)), there remain massive uncertainties concerning the appropriate values of the structural parameters. The critical feedback parameter f (and hence, by (9), climate sensitivity λ) is perhaps the biggest uncertainty in the system. While this chapter concentrates on this particular uncertainty, it should be appreciated that the relevant values of h and of past forcings $\{\Delta F(s)\}$ are also very uncertain. Just glancing at equation (5) is highly suggestive of why it is so difficult in practice to infer f (or λ) directly from data. The record of past natural forcing experiments is extremely noisy and such components as aerosol concentrations are notoriously difficult to identify. Furthermore, it is readily shown that the first-order response of a system like (5) to a change in forcings does

not involve long-run parameters like f (or λ) at all, but more centrally concerns the overall ability of the oceans to take up heat as embodied in the thermal inertia coefficient h, which itself is not very well known in this aggregative context. It is statistically very difficult to distinguish between a high-f low-h world and a low-f high-h world. To be able to deduce f (or λ) at all precisely would require a long and fairly accurate time series of past natural forcings along with a decent knowledge of the relevant thermal inertia – none of which is readily available. A more detailed look at how most scientists frame and view the difficulty of inferring climate sensitivity is examined in the next section.

3 WHY IS CLIMATE SENSITIVITY SO UNPREDICTABLE?

The title of this section is taken from the title of an influential recent *Science* article by Roe and Baker (2007), which highlights nicely the core dilemma here. Their explanation overlaps in its logic and spirit with a long series of preceding scientific insights that were similar in tone but were less formally articulated. The starting point for this genre of explanations is the observation that forcings ΔF and feedback factors f are both linearly additive in individual subcomponents, while neither temperature responses ΔT nor climate sensitivity λ display this linear additivity property.

What bothers scientists most about the climate-sensitivity issue is that, even after some three decades of intensive research, almost no progress has been made on providing a meaningful upper bound for climate sensitivity. I would argue that this concern is somewhat misstated and perhaps even misdirected. It is not the absence of an absolute upper bound on climate sensitivity per se that is disturbing or, for that matter, even mysterious. The absence of an upper bound on λ just means that the right tail of the corresponding climate-sensitivity PDF is very long and stretched out because very high values cannot absolutely be ruled out. However, it is not the length of the right tail PDF that is disturbing for policy implications, but rather its thickness. A great many catastrophic possibilities in our world have long tails, but we do not worry about them because we may have some reason to believe that these long tails are thin with probability and their asymptotic PDFs converge rapidly towards zero. The real problem with estimated climate sensitivity PDFs is that the right tail is too 'fat' (or 'thick' or 'heavy') with probability to allow us to feel comfortable with our current state of knowledge. In this sense, the scientific concern about the lack of an absolute upper bound on climate sensitivity is somewhat misconstrued because the real issue is not that the right upper tail is too long (which only means that high values are empirically or theoretically conceivable) but that it is too fat (which means that high values are possible with uncomfortably large probabilities). Although the scientists themselves do not make a distinction between tails that are too long and tails that are too fat, in this section I restate their explanations in terms of the more substantive and more genuinely disturbing issue of why the upper tail of the PDF of climate sensitivity is so fat with probability (as opposed to being so stretched out or long or unbounded).

Inferences about climate sensitivity come from two broad classes or categories of studies. The first group is computer simulations of large-scale climate models with randomized parameters. The second group is noisy observations of past natural experiments via proxies that essentially mimic $\Delta T/\Delta F$. I begin with the first category.

Roe and Baker (2007) come at the climate-sensitivity estimation problem from the perspective of the first category of perturbed-physics simulations of large ensembles of computer models. In computer-simulated numerical modeling of climate, there are hundreds of parameters governing all manner of minute details, such as fall speed of raindrops, how reflectivity changes as snow ages, exchange of turbulent fluxes in the boundary layer, evapotranspiration through plant roots, and so forth. In most cases, these uncertain parameters represent not-directly-observable 'effective' coefficients that are stochastically perturbed in the simulations – and the physical meaning of what they actually represent can be quite unclear. The major feedback parameters that climate scientists typically analyze (water vapor/lapse rate, clouds, surface albedo and so forth) are some very complicated functions of obscure combinations of model parameters. The climate system has complex, non-linear, and even chaotic features.

Despite these non-linear complications and the overall messiness of climate dynamics, Roe and Baker (2007) argue that in practice, feedbacks still combine more-or-less additively. If

$$f = \sum_{j=1}^{m} f_j \tag{10}$$

and if each feedback sub-factor f_j is distributed more-or-less independently of the other feedback sub-factors, then if m is large enough and each f_j is small enough, by the central limit theorem, the overall feedback factor f is distributed approximately normally. The argument closes by noting that climate sensitivity λ defined by (9) is then basically distributed as one over a normal PDF, which is essentially a skewed Cauchy-Lorentz-like distribution which has a long upper tail. As I have indicated, I think the real issue here is that this PDF has a fat upper tail (not that it is long). In my version of this story, the upper tail of the Cauchy-Lorentz-like PDF of λ is fat because it behaves asymptotically like a power-law distribution $\propto 1/\lambda^2$. (To make the story airtight, one must set aside issues about values of f greater than one, or dividing by zero, or artificial truncations, which, alas, are far from being trivial details because how they are resolved can substantially alter the logic of the argument and its conclusions. Furthermore, the independence assumption is suspicious because of likely negative correlations among the f_j, analogous to temperature-constrained f being negatively correlated with h – more on this later.)

Turning to the second category of empirical measurements, a less formal version of what seems generically to be a very similar story to the above Roe-Baker version (of why it is difficult to obtain an upper bound on climate sensitivity) has been present in the scientific literature for some time now. This long-present story concerns noisy observations of past natural experiments by proxies that essentially mimic ΔT and ΔF, from which $1/\lambda$ is essentially estimated as $\Delta F/\Delta T$. For concreteness, I use the recent formulation in an influential survey article by Allen et al. (2006) entitled *Observational Constraints on Climate Sensitivity*. The mechanism behind this story is analogous to the underlying mechanism of the Roe-Baker story except that here ΔF plays the role of f and ΔT plays the role of λ. Again, the key point of departure is that even though the climate system has complex, non-linear, and even chaotic features, in practice observed changes in radiative forcings still combine more or less additively. If

$$\Delta F = \sum_{j=1}^{m} \Delta F_j \tag{11}$$

and if each radiative sub-forcing ΔF_j is distributed more-or-less independently of the other radiative sub-forcings, then if m is large enough and each ΔF_j is small enough, by the central limit theorem, the overall radiative forcing ΔF is distributed approximately normally.

If one writes

$$\Delta F = \frac{1}{\lambda}\Delta T \tag{12}$$

then, it is further asserted by Allen et al., that the dominant uncertainties in empirical observations are on the left-hand side of equation (12) because empirical uncertainty in measured or inferred ΔT is typically much smaller than empirical uncertainty in measured or inferred forcings. This line of reasoning – normal measurement errors on noisy observations of ΔF – strongly suggests estimating the coefficient $1/\lambda$ in (12) by regressing observations of ΔF as the dependent variable on observations of ΔT

as the independent variable. (There is nothing inherently wrong with this approach so long as one keeps in mind that the causality we actually believe in, which is also what we need for prediction purposes, goes in the reverse direction: $\Delta F \rightarrow \Delta T$.) In the case of regressing ΔF on ΔT in (12), if the variance of the error term were known, then under standard Bayesian assumptions, the posterior-predictive distribution of $1/\lambda$ would be normal. Once again, the argument is concluded by noting that climate sensitivity λ is then distributed as one over a normal PDF; this, again setting aside issues about negative values or dividing by zero or artificial truncations (that, unfortunately for this argument, are actually substantial), is essentially a skewed Cauchy-Lorentz-like distribution with a long upper tail. And once more the real issue is that this Cauchy-Lorentz-like distribution has a fat upper tail (not that it is long per se), which comes about because this inverted-normal PDF tail behaves asymptotically like a power-law distribution $\propto 1/\lambda^2$.

In the previous paragraph, I have done my best to represent accurately what I think is the prevailing scientific wisdom about why there are observational constraints causing the PDF of climate sensitivity to have a fat upper tail. Alas, I fear this physical reasoning may be somewhat incomplete as stated and perhaps is not even fully rigorous. In what follows, I try to give a more careful rendition of the core inference-prediction problem. Interpreted carefully, the above-stated idea that there are normally-distributed measurement errors on noisy observations of ΔF translates (12) into a statement about the conditional PDF of ΔF given ΔT, where both random variables are drawn from some as-yet-unspecified DGP whose joint realizations are bivariate observations of $(\Delta F, \Delta T)$. The formally correct linear-normal translation here is that given any realized value of ΔT, the conditional PDF of ΔF is

$$\Delta F \mid \Delta T \sim \mathcal{N}(a + b\Delta T, \nu). \tag{13}$$

However, what we are really interested in for climate-change prediction is not (13), but rather the PDF of $\Delta T \mid \Delta F$. In other words, what we really want to know is what happens to ΔT for a given postulated ΔF. It cannot be stressed strongly enough that there is simply no way to infer the PDF of $\Delta T \mid \Delta F$ from a knowledge of the PDF of $\Delta F \mid \Delta T$ alone. Essentially, one is required additionally to postulate some knowledge of the bivariate DGP that is jointly generating the observed values of $(\Delta F, \Delta T)$. To illustrate what is involved in this kind of inference-prediction problem as applied to climate change, pretend for the sake of analytical simplicity that the PDF of empirical observations of $(\Delta F, \Delta T)$ is bivariate normal. This is undoubtedly not literally true, but the bivariate normal assumption will

serve here to highlight crisply the critical inference-prediction problems involved in explaining why climate sensitivity is so unpredictable – here taken as meaning that the climate sensitivity PDF has a thick upper tail.

Suppose, then, that the complete DGP (data generating process) of what has been measured in past noisy observations is postulated to be bivariate normal in $(\Delta F, \Delta T)$, which (using standard notation) is equivalent to having the key properties:

$$\Delta F \sim \mathcal{N}(\mu_F, \sigma_F^2), \tag{14}$$

$$\Delta T \sim \mathcal{N}(\mu_T, \sigma_T^2), \tag{15}$$

$$\Delta F \mid \Delta T \sim \mathcal{N}\left(\mu_F + \rho\frac{\sigma_F}{\sigma_T}[\Delta T - \mu_T], \sigma_F^2(1 - \rho^2)\right), \tag{16}$$

$$\Delta T \mid \Delta F \sim \mathcal{N}\left(\mu_T + \rho\frac{\sigma_T}{\sigma_F}[\Delta F - \mu_F], \sigma_T^2(1 - \rho^2)\right), \tag{17}$$

$$\rho = \frac{E[(\Delta T - \mu_T)(\Delta F - \mu_F)]}{\sigma_T \sigma_F}. \tag{18}$$

Notice that (16) has the same form as (13), where $a \equiv \mu_F - \rho\sigma_F\mu_T/\sigma_T$, $b \equiv \rho\sigma_F/\sigma_T$, $\nu \equiv \sigma_T^2(1 - \rho^2)$. However, the incomplete logic involved in trying to infer the tail properties of the PDF of climate sensitivity as anything like the reciprocal of the PDF of b from (13) (which is like trying to infer the tail properties of $\Delta T \mid \Delta F$ from the PDF of $\Delta F \mid \Delta T$ alone) becomes apparent once the entire DGP has been carefully specified – as, for example, in (14)–(18) above.

Once granted that the observed data is being generated by independent draws from the bivariate normal distribution in equations (14)–(18), there is no question but that the critical equation we are interested in for predicting the temperature-change response to a given change in forcing is (17). The bivariate normal system in equations (14)–(18) is telling us the overall DGP for the noisy observations we are measuring in the past data from natural forcing experiments. We are not really interested in this full noise-generating DGP (14)–(18) as an end in itself. For example, we are not really interested in identifying all five parameters in the five-equation system (14)–(18). We are only interested in predicting as accurately as possible what will be the noise-free true future temperature response to a hypothetically known noise-free true change in radiative forcing. Given the model assumptions, this best predictor will come only from regression estimation of (17), irrespective of the rest of the bivariate-normal DGP system (14)–(18).

If we want to conceptualize some 'true' value of climate sensitivity λ that is hidden from us by noisy disturbances via errors of measurement or observation (and most scientists seem to prefer to think this way), a natural interpretation of how to estimate λ is as follows. Suppose λ were known. In the absence of any errors of measurement or observation, the true relationship is that a known 'true' change in constant radiative forcing ΔF^* would induce (for each parametrically fixed value of ΔF^*) a known 'true' equilibrium temperature response ΔT^* according to the linear-proportional (but not affine) formula

$$\Delta T^* \mid \Delta F^* = \lambda \Delta F^*. \tag{19}$$

The 'true' value of λ in equation (19) is the Holy Grail of climate sensitivity we seek, since knowing it would allow us to predict the equilibrium temperature response to various radiative forcings corresponding to various GHG scenarios. The only way that the standard scientific linear description (19) can be made compatible with (17) is by imposing on (17) the a priori known (on the basis of scientific first principles) additional constraint that the affine term is zero:

$$\mu_T - \rho\frac{\sigma_T}{\sigma_F}\mu_F = 0, \tag{20}$$

in which case (17) becomes transformed into

$$\Delta T \mid \Delta F \sim \mathcal{N}(\lambda \Delta F,\ V), \tag{21}$$

where simple algebra then shows that $\lambda = \mu_T/\mu_F$ and $V = \sigma_F^2(1 - \rho^2)$, neither of which is directly observable.

If we are allowed to imagine that the noisy observations come in the form of n independent realizations from the bivariate normal DGP described by equations (14)–(18), then the DGP for observation i $(i = 1, 2, \ldots, n)$ of (21) can be written in the familiar linear-normal regression form

$$\Delta T_i = \lambda \Delta F_i + \varepsilon_i, \tag{22}$$

where each 'error term' ε_i is independent identically distributed $\mathcal{N}(0, V)$. Interestingly, fat tails on the posterior-predictive distribution of λ also emerge from this more complete description of the DGP via the following tail-fattening mechanism. Let $\hat{\lambda}$ be the least-squares estimator

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} \Delta T_i \Delta F_i}{\sum_{i=1}^{n} \Delta F_i^2}, \tag{23}$$

where the sample variance is

$$\hat{V} = \frac{1}{n}\sum_{i=1}^{n} (\Delta T_i - \hat{\lambda}\Delta F_i)^2. \tag{24}$$

Under standard Bayesian reference prior assumptions, the posterior-predictive distribution of λ is Student-t with $n - 1$ degrees of freedom:

$$\phi(\lambda) \propto \left(1 + \frac{(\lambda - \hat{\lambda})^2}{\hat{V}}\right)^{-(n+1)/2}. \tag{25}$$

The Student-t PDF (35) is fat-tailed for all $n < \infty$, displaying the asymptotic behavior of a power law in λ with exponent $n + 1$. The fatness of the tails is directly proportional to $\hat{V}$ and inversely proportional to n, so that the empirical real-world fact that the upper tail of climate sensitivity is actually very fat traces back to a relative scarcity of independent observations (n is small) combined with very noisy observations ($\hat{V}$ is large).

The above reasoning is just one specific example of a generic argument that when you are trying to infer the value of a parameter way outside the range of usual experience or data, you end up with a thick-tailed posterior-predictive distribution. This thick tail reflects back the underlying thickness of the standard non-informative reference family of priors. The posterior-predictive distribution is more thin in proportion to the number of observations, but it is still technically thick for any given n. I have already explained this mechanism in some detail in earlier work and don't elaborate further here.[4] What I think all of this shows is that a careful restatement of the Allen et al. (2006) argument still produces fat tails, although by a somewhat different mechanism than the original one they had in mind.

The Roe and Baker tail-fattening mechanism may appear to be along slightly different lines, but I think it is ultimately more similar than different. Something suspicious is happening there when $f \approx 1$ because we don't really believe it is easily physically possible to have $f > 1$ on the grounds that this would result in limitless runaway warming. So probably there is some prior information restricting $f = \Sigma f_j$ from being greater than one. In other words, if a bunch of f_j values are big, then a bunch of other f_j values should be small in order to compensate and keep f from being greater than one. This is somewhat analogous to the argument that observed temperature constrains f and h to be negatively correlated. It means the f_j are not independently distributed and the appeal to the central limit theorem is suspect. But the very complicated macro interactions restricting $f = \Sigma f_j < 1$ are not well modeled by the micro assumption that the f_j are independently distributed. The Roe and Baker story may be a good mechanical description of where the long thick tail is coming from in the simulations that human beings perform on computers, but the physics behind the independence of micro f_j within the simulations is questionable because it is inconsistent with the macro-physics that $f < 1$.

I think that the Roe and Baker tail-fattening mechanism and my tail-fattening mechanism explained above are more similar than different in the following sense. In both situations, we don't know how to represent things far outside the range of average experience. In both cases, the thick tails are coming from prior assumptions built into the modeling process rather than from hard science that justifies these assumptions. In both cases, the fat tails reflect more these prior assumptions than any actual empirical knowledge of overall system behavior in that extreme region where f is close to one and λ is very big.

I would emphasize that none of this dependence on subjective prior modeling assumptions makes the problem any less real. To make economic decisions today, we must work with the fat posterior-predictive distributions, which are all coming essentially from lack of prior knowledge about extreme values, reinforced by lack of empirical experience with extreme values. Throughout the rest of the chapter I just assume that the PDF of climate sensitivity λ has a thick upper tail – for whatever reason – and examine the consequences for the economics of climate change. Thus, in what follows, the PDF of λ, which is denoted $\varphi(\lambda)$, is presumed to have an upper tail that declines to zero less rapidly than exponentially (for example, polynomially, as with a power-law PDF).

4 ECONOMIC DYNAMICS OF FAT-TAILED TEMPERATURES

From (5) it is immediately apparent that, other things being equal, at any given time, higher values of f (and hence of λ) imply higher values of ΔT. However, it is also true that higher values of λ (or of f) make the system take a longer time for ΔT to reach its long-run equilibrium value. To examine this issue more closely, suppose in what follows that the planet has been subject to a sustained doubling of CO_2. With the convention being followed here that all units are expressed in terms of a doubling of

CO_2, this means that $\bar{F} = 1$. Substituting $\bar{F} = 1$ and (9) into (7), the time trajectory of temperature change (expressed in terms of λ) becomes

$$\Delta T_\lambda(t) = \lambda\left[1 - \exp\left(-\frac{\lambda_0 t}{h\lambda}\right)\right]. \tag{26}$$

Equation (26) indicates that $\Delta T \rightarrow \lambda$. The question I now examine is length of the time it takes for ΔT to attain the fraction α of λ, where $0 < \alpha < 1$. Call this time $t_\alpha(\lambda)$. From (26), $t_\alpha(\lambda)$ must obey the equation

$$\alpha\lambda = \lambda\left[1 - \exp\left(-\frac{\lambda_0 t_\alpha(\lambda)}{h\lambda}\right)\right], \tag{27}$$

which can be rewritten as

$$t_\alpha(\lambda) = \psi(\alpha)\lambda, \tag{28}$$

where

$$\psi(\alpha) = \frac{h}{\lambda_0}(-\ln(1-\alpha)). \tag{29}$$

From (28), it is apparent that $t_\alpha(\lambda) \propto \lambda$. Thus, for any given α, the larger the climate sensitivity λ, the longer it takes for the system to attain the temperature change $\Delta T = \alpha\lambda$. Nevertheless, if the PDF of λ is fat-tailed at the upper end, then eventually there is a positive probability of ΔT becoming unboundedly high. The formal statement of this possibility is that for any value of T', however large, there exists some time t' such that $t > t'$ implies $P[\Delta T(t) > T'] > 0$. The question I now seek to address is what this possibility of very high temperatures arriving at very distant times does to economic welfare analysis. The answer depends (among many other things) on what is assumed about damage at very high temperatures.

Most existing Integrated Assessment Models (IAMs) treat high-temperature damage by an extremely casual extrapolation of whatever specification is arbitrarily assumed to be the low-temperature 'damage function'. The high-temperature 'damage function' extrapolated from the low-temperature 'damage function' is remarkably sensitive to assumed them because an extraordinarily wide variety of them can be made to fit virtually identically the low-temperature damage that has been assumed by the modeler. In the IAM literature, most damage functions reduce welfare-equivalent consumption by the quadratic-polynomial multiplier $1/[1 + \gamma(\Delta T)^2]$ with γ calibrated to some postulated loss for

$\Delta T \approx 2-3°$ C. There was never any more compelling rationale in the first place for this particular loss function than the comfort that economists feel from having worked with it before. In other words, the quadratic-polynomial specification is being used to assess climate-change damages for no better reason than casual familiarity with this particular form from other cost-of-adjustment dynamic economic models, where it had been used primarily for its analytical simplicity.

I would argue strongly on a priori grounds that if, for some unfathomable reason, climate-change economists want dependence of damages to be a function of $(\Delta T)^2$, then a far better choice of functional form at high temperatures for a welfare-equivalent quadratic-based consumption-loss multiplier is the exponential form $\exp(-\gamma(\Delta T)^2)$. Why? Look at the specification choice abstractly. What might be called the 'attenuating pressure' on welfare (denoted A) is arriving here as the arbitrarily imposed quadratic form $A(\Delta T) = (\Delta T)^2$, around which some further structure is imposed to convert it into utility units. With isoelastic utility, the exponential specification is equivalent to $dU/U \propto dA$, while for high A the polynomial specification is equivalent to $dU/U \propto dA/A$. For me it is obvious that, between the two, the former is much superior to the latter. Why should the impact of dA on dU/U be artificially and unaccountably diluted through dividing dA by high values of A in the latter case? The same argument applies to any polynomial in ΔT. Of course I cannot prove that my favored choice here is the more reasonable of the two functional forms for high ΔT (although I truly believe that it is), but no one can disprove it either – and *this* is the point.

The value of γ required for calibrating welfare-equivalent consumption at $\Delta T = 3°$ C to be, say, 98 per cent of consumption at $\Delta T = 0°$ C is so miniscule that both the polynomial-quadratic multiplier $1/[1 + \gamma(\Delta T)^2]$ and the exponential-quadratic multiplier $\exp(-\gamma(\Delta T)^2)$ give virtually identical outcomes for relatively small values of $\Delta T \leq 5°$ C, but at ever higher temperatures, they gradually yet ever-increasingly diverge. With a fat-tailed PDF of λ and a very long time horizon, there can be a big difference between these two functional forms in the implied willingness to pay (WTP) to avoid or reduce uncertainty in ΔT. I next calculate the WTP to avoid uncertain ΔT when the consumption-loss welfare-equivalent quadratic-based multiplier is of the exponential form $\exp(-\gamma(\Delta T)^2)$. In what follows I use a utility function of the constant elasticity form

$$U(C) = \frac{C^{1-\eta}}{1-\eta}, \tag{30}$$

where the coefficient of relative risk aversion is $\eta > 1$ and $C(0)$ is normalized to unity.

Suppose the economy grows at a given rate $g > 0$. The rate of pure time preference is $\delta > 0$. Suppose there is some arbitrarily-imposed time horizon H. The random variable of climate sensitivity λ has a thick upper tail in its PDF $\varphi(\lambda)$. The base case thought experiment here is a doubling of CO_2 beginning at time zero. I now ask: what is the WTP – in terms of a constant fraction of consumption foregone at all times t between 0 and H – to avoid altogether this temperature uncertainty? The answer as a function of the time horizon is denoted here as $WTP(H)$.

Making use of (28), (29) and (30), $WTP(H)$ must satisfy the condition:

$$\frac{\int_0^H ((1 - WTP(H))\exp(gt))^{1-\eta}\exp(-\delta t)\,dt}{1-\eta}$$

$$= \frac{\int_0^1 \left[\int_0^{H/\psi(\alpha)} (\exp(gt_\alpha(\lambda) - \gamma\alpha^2\lambda^2))^{1-\eta}\exp(-\delta t_\alpha(\lambda))\varphi(\lambda)\,d\lambda\right] d\alpha}{1-\eta} \tag{31}$$

As $H \to \infty$, it is not difficult to show that with a fat tail in $\varphi(\lambda)$, the integral on the right-hand side of (31) approaches $-\infty$ because the term in $\exp(-\gamma\alpha^2\lambda^2)$ dominates everything else. This fact in turn implies that

$$\lim_{H\to\infty} WTP(H) = 1. \tag{32}$$

When the consumption-reducing welfare-equivalent damage multiplier has the exponential form $\exp(-\gamma(\Delta T)^2)$, then as the horizon $H \to \infty$, the above result (32) implies at the limit that the WTP to avoid (or even reduce) fat-tailed uncertainty in ΔT approaches 100 per cent of consumption. This does not mean, of course, that we should be spending 100 per cent of consumption to eliminate the climate-change problem. But this example does highlight the remarkable ability of miniscule refinements of the damage function (when combined with fat tails) to dominate climate-change cost–benefit analysis – and the remarkable fragility of policy advice coming out of conventional thin-tailed IAMs with polynomial damage.

I think this example shows that a previous two-period result that fat-tailed climate sensitivity can have strong economic implications survives being recast as a more complete dynamic specification, even though (other things being equal) the higher the temperature realization, the later this temperature realization is expected to arrive. When fat climate-sensitivity tails are combined with very uncertain high-temperature damage, this aspect can dominate the discounting aspect in calculations of expected present discounted utility – even at empirically plausible real-world interest rates and even when taking full account of the important continuous correlation that, conditional upon its realization, the higher the temperature, the later its expected time of arrival.

The model I have used throughout this chapter for the sake of analytical sharpness is so incredibly oversimplified that it can legitimately be criticized on an enormous number of counts as being grossly unrealistic. The temperature dynamics is primitive, climate change involves much more than an instantaneous doubling of atmospheric CO_2, the utility function may be wrong (especially for low consumption), results depend on the postulated exponential-quadratic damage function, policy is much richer than a double-or-nothing CO_2 decision, there is some possibility of learning and adaptive mitigation (although the inertial commitment of GHGs already in the pipeline is distressingly long), technological change is ignored, and so forth and so on. Nevertheless, I believe that a fair conclusion from this example is that any economic analysis of climate change that does *not* include an explicit treatment of rare climate catastrophes (no matter how far off in the future they may occur) is problematic and its policy conclusions are under a dark cloud until this fat-tailed disaster aspect is modeled explicitly and addressed seriously.

5 CONCLUSION

This chapter has two main goals. The first goal is an attempt to place the physical-science measurement-based discussion of why climate sensitivity is so unpredictable into a broader context of statistical inference, prediction and decision making. Here the chapter makes two basic points: (1) it is not the fact that it is difficult to place an upper bound on climate sensitivity that is worrisome, but rather the fact that the upper tail of its PDF is fat with probability; and (2) the fatness of the upper tail of the PDF of climate sensitivity comes primarily from being generically built into any situation where we are trying to estimate the probabilities of rare outlier events from limited data based on incomplete structural knowledge.

The second goal of the chapter is to show that previous findings from a two-period discrete-time formulation survive the introduction of continuous time and more realistic dynamics that take explicit account of the fact that higher temperatures arrive later. The overarching message of the line of research leading to this chapter continues to be that, at least potentially, the influence on cost–benefit economic analysis of fat-tailed structural uncertainty about climate change, coupled with great unsureness about high-temperature damage, can outweigh the influence of discounting or anything else. My message is that we must learn to live with the idea that the answers to cost–benefit analyses of what to do about climate change may very well depend – at least to some degree – upon subjective judgments about how bad it might get, with what probabilities, in the most extreme situations.

ACKNOWLEDEMENTS

For helpful detailed comments on earlier drafts of this paper, but without implicating them for its remaining defects, I am grateful to Harvey Lam and Richard Tol.

I am also indebted to Gerard Roe for generously helping to guide me through some technical scientific aspects in e-mails, and private conversations, and via his insightful paper Roe (2007). He should not, however, be blamed for my interpretations or the formulation of this chapter.

NOTES

1. Details of this calculation are available from the author upon request. Eleven of the studies in Table 9.3 overlap with the studies portrayed in Box 10.2. Four of these overlapping studies conflict on the numbers given for the over-5 per cent level. For three of these differences, I chose the Table 9.3 values on the grounds that all of the Box 10.2 values had been modified from the original studies to make them have zero probability mass above 10° C. (The fact that all PDFs in Box 10.2 have been normalized to zero probability above 10° C biases my over-5 per cent averages here towards the low side.) With the fourth conflict (Gregory et al., 2002a), I substituted 8.2° C from Box 10.2 for the ∞ in Table 9.3 (which arises only because the method of the study itself does not impose any meaningful upper-bound constraint). The only other modification was to average the three reported volcanic-forcing values of Wigley *et al.* (2005a) in Table 9.3 into one study with the single over-5 per cent value of 6.4° C.
2. As I use the term in this paper, a PDF has a 'fat' (or 'thick' or 'heavy') tail when its moment generating function (MGF) is infinite – that is, the tail probability approaches zero more slowly than exponentially. The standard example of a fat-tailed PDF is a power-law-family distribution, although, for example, a lognormal PDF is also fat-tailed, as is an inverted-normal or inverted-gamma. By this definition, a PDF whose MGF is finite has a 'thin' tail. A normal or a gamma are examples of thin-tailed PDFs, as is *any* PDF having finite supports.
3. Weitzman (2008).
4. See Weitzman (2008), where the fat-tailed properties of this example are extended to a much broader family of distributions than the normal and Student-*t*.

REFERENCES

Allen, M., N. Andronova, B. Booth, S. Dessai, D. Frame, C. Forest, J. Gregory, G. Hegerl, R. Knutti, C. Piani, D. Sexton and D. Stainforth (2006). 'Observational Constraints on Climate Sensitivity'. Chapter 9 in *Avoiding Dangerous Climate Change*, ed. H.J. Schellnhuber, Cambridge University Press, pp. 281-289.

Allen, Myles R. and David J. Frame (2007). 'Call Off the Quest' *Science* (October 26), 318, pp. 582–583.

Gregory, J.M., R.J. Stouffer, S.C.B. Raper P.A. Stott and N.A. Rayner (2002). 'An Observationally Based Estimate of the Climate Sensitivity'. Journal of Climate, 15, pp. 3117–3121.

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 University Press, 2007 (available online at http://www.ipcc-wg1.unibe.ch/publications/wg1-ar4/wg1-ar4.html, accessed 16 May 2010).

Roe, Gerard H (2007). 'Feedbacks, timescales, and seeing red'. Mimeo, September 2.

Roe, Gerard H. and Marcia B. Baker (2007). 'Why is Climate Sensitivity So Unpredictable?' *Science* (October 26), 318, pp. 629–632.

Weitzman, Martin L (2008). *Climate Change*. Working Paper, February 8 (available online at www.economics.harvard.edu/files/faculty/61_Modeling.pdf, accessed 16 May 2010).

Wigley, T.M.L., C.M Ammann, B.D. Santer and S.C.B. Raper (2005). 'Effects of Climate Sensitivity on the Response to Volcanic Forcing'. Journal of Geophysical Research, 110, D09107, doi: 10.1029/2004JD005557.