

NBER WORKING PAPER SERIES

STRUCTURAL UNCERTAINTY AND THE VALUE OF STATISTICAL LIFE IN
THE ECONOMICS OF CATASTROPHIC CLIMATE CHANGE

Martin Weitzman

Working Paper 13490

<http://www.nber.org/papers/w13490>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

October 2007

Department of Economics, Harvard University, Cambridge, MA 02138 (e-mail: mweitzman@harvard.edu). For helpful detailed comments on earlier drafts of this paper, but without implicating them for its remaining defects, I am grateful to Frank Ackerman, Roland Benabou, Richard Carson, Daniel Cole, Stephen DeCanio, Don Fullerton, Olle Häggström, Robert Hahn, Karl Löfgren, Michael Mastrandrea, Robert Mendelsohn, William Nordhaus, Cedric Philibert, Richard Posner, John Reilly, Richard Tol, Gary Yohe, and Richard Zeckhauser. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2007 by Martin Weitzman. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Structural Uncertainty and the Value of Statistical Life in the Economics of Catastrophic Climate Change

Martin Weitzman

NBER Working Paper No. 13490

October 2007

JEL No. Q54

ABSTRACT

Using climate change as a prototype motivating example, this paper analyzes the implications of structural uncertainty for the economics of low-probability high-impact catastrophes. The paper shows that having an uncertain multiplicative parameter, which scales or amplifies exogenous shocks and is updated by Bayesian learning, induces a critical "tail fattening" of posterior-predictive distributions. These fattened tails can have strong implications for situations (like climate change) where a catastrophe is theoretically possible because prior knowledge cannot place sufficiently narrow bounds on overall damages. The essence of the problem is the difficulty of learning extreme-impact tail behavior from finite data alone. At least potentially, the influence on cost-benefit analysis of fat-tailed uncertainty about the scale of damages -- coupled with a high value of statistical life -- can outweigh the influence of discounting or anything else.

Martin Weitzman

Department of Economics

Harvard University

Littauer 313

Cambridge, MA 02138

and NBER

mweitzman@harvard.edu

Structural Uncertainty and the Value of Statistical Life in the Economics of Catastrophic Climate Change

Martin L. Weitzman*

October 9, 2007. Comments appreciated.

Abstract

Using climate change as a prototype motivating example, this paper analyzes the implications of structural uncertainty for the economics of low-probability high-impact catastrophes. The paper shows that having an uncertain multiplicative parameter, which scales or amplifies exogenous shocks and is updated by Bayesian learning, induces a critical “tail fattening” of posterior-predictive distributions. These fattened tails can have strong implications for situations (like climate change) where a catastrophe is theoretically possible because prior knowledge cannot place sufficiently narrow bounds on overall damages. The essence of the problem is the difficulty of learning extreme-impact tail behavior from finite data alone. At least potentially, the influence on cost-benefit analysis of fat-tailed uncertainty about the scale of damages – coupled with a high value of statistical life – can outweigh the influence of discounting or anything else.

1 Introduction

This paper is about the effects of structural uncertainty on the expected utility analysis of system-wide potential catastrophes. The key unknown structural parameter in this paper is a critical multiplier that amplifies (or scales) an uncertain exogenous shock or impulse to the system. Very roughly – at a high level of abstraction and without trying to push an imperfect

*Department of Economics, Harvard University, Cambridge, MA 02138 (e-mail: mweitzman@harvard.edu). For helpful detailed comments on earlier drafts of this paper, but without implicating them for its remaining defects, I am grateful to Frank Ackerman, Roland Benabou, Richard Carson, Daniel Cole, Stephen DeCanio, Don Fullerton, Olle Häggström, Robert Hahn, Karl Löfgren, Michael Mastrandrea, Robert Mendelsohn, William Nordhaus, Cedric Philibert, Richard Posner, John Reilly, Richard Tol, Gary Yohe, and Richard Zeckhauser.

analogy too far – the role of this uncertain multiplicative amplifier or scale parameter can be illustrated by the role of uncertain “climate sensitivity” in discussions of global warming. Let $\Delta \ln CO_2$ be sustained relative change in concentrations of atmospheric carbon dioxide while ΔT is equilibrium temperature response. Climate sensitivity S is an amplifying or scaling multiplier for converting $\Delta \ln CO_2$ into ΔT by the (reasonably accurate) linear approximation $\Delta T \approx (S/\ln 2) \times \Delta \ln CO_2$. The International Panel on Climate Change in its IPCC4 (2007) Executive Summary states: “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.” (For IPCC4, “likely” is $P > 66\%$, while “very unlikely” is $P < 10\%$.)

In this paper I am mostly concerned with the 17% of those S “values substantially higher than 4.5°C” which “cannot be excluded.” Eighteen recent studies of climate sensitivity with 18 probability density functions (PDFs) of S lie behind the above-quoted IPCC4 summary statement. From Figure 1 in Box 10.2 of IPCC4 (2007), it is apparent that the upper tails of these 18 PDFs tend to be long and fat. For neatness of its graphical display, Box 10.2 arbitrarily truncates all PDFs at 10°C. But tiny positive tail probabilities in many of the studies stretch well beyond 10°C because the science makes it difficult to constrain absolutely the high end. The upper 5% probability level averaged over all 18 studies is $\bar{S}=6.2^\circ\text{C}$, which I take as approximately meaning that $P[S > 6^\circ\text{C}] \approx 5\%$. For the purposes of this paper I assume $P[S > 8^\circ\text{C}] \approx 2\%$, which is roughly consistent with the averaged upper tails in Box 10.2. Such numbers are exceedingly crude ballpark estimates of what amounts to huge climate impacts with tiny probabilities, but the subject matter of this paper concerns just such kind of situations and my example here does not depend on precise numbers.

Global average warming of 8°C masks tremendous local and seasonal variation, which can be expected to produce temperature increases *much* greater than this at particular times in particular places. Societies and ecosystems in a world whose mean temperature will have changed in the course of one or two (or even three) centuries by $\Delta T > 8^\circ\text{C}$ (for U.S. readers: $8^\circ\text{C} = 14.4^\circ\text{F}$) are located in *terra incognita*, since such high average temperatures have not existed for at least tens of millions of years. So much global warming so quickly would be unprecedented perhaps even on a time scale of hundreds of millions of years. There exist some truly frightening examples of possible consequences of such magnitude of mean temperature increases: sudden disintegration of the Greenland or West Antarctic ice sheets with dramatic raising of sea level by perhaps fifteen meters or so, shutdowns or even re-

versals of the warming component of large-scale oceanic circulation systems like the Gulf Stream, major disruptions of large-scale weather patterns like monsoons, highly consequential worldwide changes in regional freshwater availability, amplified-positive-feedback greenhouse warming due to heat-induced (i.e., endogenous, and therefore *in addition to* climate sensitivity as conventionally defined) rapid releases of the immense amounts of greenhouse gases currently sequestered in arctic permafrost or (more remotely) in the even-much-vaster deposits of offshore methane hydrates – and so forth and so on. It is difficult to imagine what $\Delta T > 8^\circ\text{C}$ might mean for life on earth, but such geologically-instantaneous mean temperature changes are significantly larger than what separates us now from recent past ice ages and at a minimum would trigger massive species extinctions and monumental ecosystem changes. Furthermore, stabilizing anthropogenically-injected GHG (greenhouse gas) stocks at anything like twice pre-industrial-revolution levels looks now like an extraordinarily ambitious goal. Given current trends in GHG emissions, we will attain such a doubling within about 30-40 years and will then go well beyond that amount unless relatively drastic measures are taken starting soon. Projecting current trends in business-as-usual GHG emissions, a *tripling* of GHG concentrations would be attained relative to pre-industrial-revolution levels within about a century and a steady-state-equilibrium global-mean-temperature response might in this case involve for the tail something like $P[\Delta T > 10^\circ\text{C}] \approx 1\%$ (for U.S. readers: $10^\circ\text{C}=18^\circ\text{F}$).

All of the above horrifying examples of climate-change mega-disasters are incontrovertibly possible. They were purposely selected to be especially lurid in order to drive home a valid point. The tiny probabilities of nightmare impacts of climate change are all such crude ballpark estimates that there is a tendency in the literature to dismiss them on the “scientific” grounds that they are much too highly speculative to be taken seriously because they are statistically indistinguishable from zero. By contrast with the conventional wisdom of *not* taking seriously extreme-temperature-change tiny probabilities *because* such crude estimates are highly speculative and statistically insignificant, the purpose of this paper is to prove that the exact opposite logic is in fact true by giving a rigorous sense in which, other things being equal, the more speculative and fuzzy are the tiny tail probabilities of high-impact extreme events, the less ignorable and the more serious is the situation for an agent whose welfare is measured by present discounted expected utility. The seriousness of such situations comes from combining strict relative risk aversion with a multiplicative factor having effectively open-ended uncertainty that amplifies impulses into bad impacts with long fat tails.

Oversimplifying enormously, how warm the climate ultimately gets is approximately a product of two factors – the amount of GHG concentrations and a critical climate-sensitivity scaling multiplier. Both of these factors are uncertain, but the scaling multiplier is much

more open-ended on the high side with a much longer upper tail. This critical scale parameter reflecting large *scientific* uncertainty is then used as a multiplier for converting aggregated GHG emissions – an input mostly reflecting *economic* uncertainty – into eventual temperature changes. The paper will show that a generalization of this form of interaction can be re-packaged and analyzed as an aggregative macroeconomic model with essentially the same reduced form (structural uncertainty about some unknown open-ended scale-multiplying parameter amplifying an uncertain economic input). This form of interaction (coupled with finite data, under conditions of relative risk aversion) can have very strong consequences for cost-benefit analysis (CBA) when catastrophes are theoretically possible, because it can easily drive applications of expected utility (EU) theory much more than anything else.

When fed back into an economic analysis, the great open-ended uncertainty about eventual mean planetary temperature change cascades into yet-greater yet-more-open-ended uncertainty about eventual changes in utility or welfare. Not only is it very difficult to estimate tail probabilities of high- ΔT outcomes, but translating this via ambiguous local and seasonal weather consequences combined with unknown adaptation capabilities into aggregated utility-equivalent units for people living a century or two from now introduces enormous further fuzziness. When such future utility includes the existence values placed by people at that time on wild species, *in situ* conservation, natural or historical habitats, and the availability of outdoor activities more generally, the already-huge welfare uncertainties are yet further compounded. Even if climate sensitivity were bounded by some big number, the value of what might be called “welfare sensitivity” is effectively bounded only by some *very* big number representing something like the value of statistical civilization as we know it or maybe even the value of statistical life on earth as we know it.

Without further belaboring the point, the overall utility effects of global warming that might accompany a 2% chance of $S > 8^\circ\text{C}$ are sufficiently open-ended and fuzzy that it seems fair to say offhand there might be a non-negligible probability of a catastrophe. In his book *Catastrophe: Risk and Response*,¹ Richard A. Posner defines the word catastrophe “... to designate an event that is believed to have a very low probability of materializing but that if it does materialize will produce a harm so great and sudden as to seem discontinuous with the flow of events that preceded it.” Posner adds: “The low probability of such disasters – frequently the *unknown* probability, as in the case of bioterrorism and abrupt global warming – is among the things that baffle efforts at responding rationally to them.” In this paper I address what a rational economic response in the form of expected present discounted utility theory might offer in the way of guidance for thinking coherently about the economics of

¹Posner (2004). See also the insightful review by Parson (2007). Sunstein (2007) covers some similar themes from a somewhat different perspective.

highly-uncertain catastrophes with tiny but highly-unknown probabilities.

As was pointed out, the prime example of an unsure structural parameter that best illustrates this paper’s thesis is a standard-deviation-like scale parameter whose mode of interaction parallels the interaction of a climate-sensitivity amplifying multiplier. Suppose the true value of this generic scaling parameter is unknown because of limited past experience, a situation that can be modeled *as if* inferences must be made inductively from a finite number of data observations. At a high level of abstraction, each data point might be interpreted as representing an outcome from a particular scientific or economic study. (Some of the scientific studies of climate sensitivity might actually be examples of this because they rely on noisy temperature responses from past natural experiments of changed radiative forcings of unsure magnitude.) The paper shows that having an uncertain scale parameter in such a setup can add a significant tail-fattening effect to posterior-predictive expectations, even when Bayesian learning takes place with arbitrarily large (but finite) amounts of data. 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 probability distribution. It is inherently difficult to learn extreme bad-tail probabilities from finite samples alone because, by definition, we don’t get many data-point observations of such catastrophes. 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 is an event, the more unsure is our estimate of its probability of occurrence. In this spirit (from being constructed out of inductive knowledge), the empirical studies of climate sensitivity are perhaps preordained to find the fat-tailed power-law-like PDFs which they seem, approximately, to find in practice.

This paper suggests that standard approaches to modeling the economics of climate change (even those that purport to treat risk by Monte Carlo simulations) very likely fail to account adequately for the implications of large uncertain impacts with small probabilities. The overarching general message is that from inductive experience alone one cannot acquire sufficiently accurate information about the probabilities of tail disasters to prevent the expected marginal utility of an extra sure unit of consumption from becoming unbounded for any utility function having strict relative risk aversion. To close the model by bounding expected marginal utility below $+\infty$ (or expected utility above $-\infty$), this paper relies on a concept akin to the “value of statistical life” (VSL) – except that here it represents something more like the rate of substitution between consumption and the mortality risk of a catastrophic extinction of the natural world or civilization as we know these concepts.

With this particular way of closing the model, subsequent EU-based CBA will then depend critically upon an exogenously-imposed VSL-like parameter, which is a generalization of the value of a statistical human life and is presumably *very* big. Practically, a high VSL-like parameter means for situations with potentially unlimited downside exposure (like climate change) that a Monte Carlo simulation must go *very* deep into the extreme-negative-impact fat tail to merit credibility as an accurate and fair CBA. In this sense (by forcing there to be such utter dependence upon a concept like the value of a statistical life), structural or deep uncertainty is potentially much more of a driving force than discounting or risk for cost-benefit applications of EU theory to open-ended situations with potentially unlimited exposure. For such situations where there do not exist prior limits on damages (like climate change from greenhouse warming), expected present discounted utility analysis of costs and benefits is likely to be dominated by considerations and concepts related more to catastrophe insurance than to the consumption-smoothing consequences of long-term discounting at one or another particular interest rate.

2 A Disturbing Example

Let C be consumption. Consider a representative agent with a standard CRRA (constant relative risk aversion) utility function of the form

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

for $C > 0$, which implies that

$$U'(C) = C^{-\eta}. \tag{2}$$

Technically speaking, the model of this paper requires only that the coefficient of relative risk aversion η is positive. But the main application here involves a situation where the value of statistical life is extremely large, which, it will turn out, is essentially only consistent with $\eta > 1$. Although the condition $\eta > 1$ can be *derived* from the model (and later it will be), analytically it is easier just to assume it in the first place.

For analytical crispness, the model of this paper has only two periods – the present and the future. Applied to climate change, the future would be one or two centuries hence. Instead of working directly with C , in this paper it is more convenient to work with (and think in terms of) $\ln C$. If present consumption is normalized to unity, then the growth of consumption between the two periods is

$$Y \equiv \ln C, \tag{3}$$

where in this model Y is a random variable capturing *all* uncertainty that influences future values of $\ln C$. For the purposes of this paper, Y includes not just economic growth narrowly defined, but also the consumption-equivalent damages of adverse climate change. Actually, this paper is mostly concerned with the extraordinarily-small probability of an extraordinarily-large negative realization of Y that might accompany extreme climate change. (Note: here the “bad” tail of Y is its *lower* tail.)

The “stochastic discount factor” or “pricing kernel” is an expression of the form

$$M(C) = \beta \frac{U'(C)}{U'(1)} \quad (4)$$

for time-preference parameter β ($0 < \beta \leq 1$). When (2) holds, then from (4) the amount of present consumption that the agent would be willing to give up in the present period to obtain one extra sure unit of consumption in the future period is

$$E[M] = \beta E[\exp(-\eta Y)], \quad (5)$$

which is a kind of shadow price for discounting future costs and benefits in project analysis.

Throughout this paper I use the price of a future sure unit of consumption as the single most useful overall indicator of the present cost of future uncertainty. Other like indicators – such as welfare-equivalent deterministic consumption – give similar results, but the required analysis in terms of mean-preserving spreads is slightly more elaborate and slightly less intuitive. Focusing on the behavior of $E[M]$ is understood in this paper, therefore, as being a metaphor for understanding what drives the results of cost-benefit or welfare analysis more generally in situations of potentially unlimited exposure to catastrophic impacts.

Using standard notation, let lower-case y denote a realization of the upper-case random variable Y . If Y has PDF (probability density function) $f(y)$, then the price of future consumption (5) can be written as

$$E[M] = \beta \int_{-\infty}^{\infty} e^{-\eta y} f(y) dy, \quad (6)$$

which means that $E[M]$ is “essentially” the Laplace transform or moment-generating function of $f(y)$. This is helpful because the properties of the expected stochastic discount factor are the same as the properties of the moment-generating function of a probability distribution, about which a great deal is already understood. For example, if $Y \sim N(\mu, s^2)$ then plugging the usual formula for the expectation of a lognormal random variable into (6) gives the

familiar expression

$$E[M] = \exp\left(-\delta - \eta\mu + \frac{1}{2}\eta^2 s^2\right), \quad (7)$$

where $\delta = -\ln \beta$ is the instantaneous rate of pure time preference. Expression (7) shows up in innumerable asset-pricing Euler-equation applications as the expected value of the stochastic discount factor or pricing kernel when consumption is lognormally distributed. Equation (7) is also the basis of the well-known generalized-Ramsey formula for the riskfree interest rate

$$r^f = \delta + \eta\mu - \frac{1}{2}\eta^2 s^2, \quad (8)$$

which (in its deterministic form, for the special case $s = 0$) plays a key role in recent debates about what social interest rate to use for intergenerational cost-benefit discounting of policies to mitigate GHG emissions. This intergenerational-discounting debate has mainly revolved around choosing “ethical” values of the rate of pure time preference δ , but this paper will demonstrate that, for any $\eta > 0$, the effect of δ in formula (8) is theoretically overshadowed by the effect of an uncertain scaling parameter s .

Although not phrased in this way, the existing literature already contains an example that can be interpreted as showing a sense in which EU-maximizing agents are significantly more averse to structural “uncertainty” (meaning here a situation where the structure of the data generating process is unknown and must be estimated statistically) than they are to pure “risk” (the structure of the data generating process is known, but future stochastic realizations are unknown – as in (8) where $Y \sim N(\mu, s^2)$ with both parameters known). The example used here to convey this basic idea is a relatively simple specification consisting of the workhorse isoelastic or CRRA utility function (1) along with familiar probability distributions: lognormal, Student- t , gamma.² One may then ask whether the insight that structural “uncertainty” has potentially more impact on EU analysis than pure “risk” is due here to the particular quirks of this relatively simple example or, alternatively, it represents a generic insight of broader scope. The answer, given in the next section, is that the result is generic (or at least much broader than the example). The relatively simple formulation of this section will then help to motivate the subsequent development of a more general theory of catastrophic change involving structural uncertainty about the true value of the relevant

²An example with these particular functional forms leading to existence problems from indefinite expected-utility integrals blowing up was first articulated by Geweke (2001). It subsequently was independently rediscovered in a context of discounting by Weitzman (2007a), who further developed its meaning and implications for asset pricing in a nonstationary setting. Even earlier, in a very different context, Schwarz (1999) introduced a similar point to the main theme of this paper (and a generalization of Geweke’s example) involving tail-thickening of posterior-predictive distributions into power-law tails under general conditions. Mathematically, this entire set of issues is closely related to the classical generalizations of the central limit theorem by Gnedenko and Kolmogorov (1949) to cover random variables having heavy-tailed distributions.

scale parameter.

Throughout this paper, the structural scale parameter controlling the tail spread of a probability distribution is the most critical unknown. For convenience in the super-simple example of this section, it is assumed that $Y \sim N(\mu, s^2)$ where the mean μ *is known* but the scale parameter s *is unknown*. In an extremely loose sense this unknown structural scale parameter is a highly-stylized abstraction of the effect that is embodied in an uncertain climate-sensitivity amplifying multiplier, which was discussed previously in the introduction. With this rough analogy in this super-simple example, $Y \leftrightarrow A - B\Delta T$, where A and B are positive constants.

The structural uncertainty concerning s is modeled here *as if* this scale parameter is a subjectively-distributed random variable (denoted S) whose PDF must be inferred by inductive reasoning from n observed data points. At a very high level of abstraction, these data points might be interpreted as outcomes from various economic-scientific studies very roughly akin to the eighteen climate-sensitivity studies discussed in the introduction. Let $\mathbf{y} = (y_1, \dots, y_n)$ be a sample of n i.i.d. random draws from the data-generating process of the normal distribution whose PDF is

$$h(y | s) = \frac{1}{\sqrt{2\pi}s} \exp\left(-\frac{(y - \mu)^2}{2s^2}\right). \quad (9)$$

The sample variance is

$$\nu_n \equiv \frac{1}{n} \sum_{j=1}^n (y_j - \mu)^2 \quad (10)$$

and the likelihood function for the random variable S here is

$$L(s; \mathbf{y}) \propto \frac{1}{s^n} \exp\left(-\frac{n\nu_n}{2s^2}\right). \quad (11)$$

In a Bayesian framework, deriving the agent’s posterior distribution of S requires that some prior distribution be imposed on S . The choice of a “noninformative” prior in the most *general* case represents a thorny issue within Bayesian statistics. However, for the *particular* case of a one-dimensional scale parameter, which is the case here, practically all definitions of a “noninformative” prior give the same PDF. This traditional reference prior forms the Bayesian mirror image of classical linear-normal regression analysis. This prior (called Jeffreys prior, which is explained in any textbook on Bayesian statistics) is a diffuse or uniform distribution of $\ln S$ on $(0, \infty)$, meaning that the (improper) prior PDF expressed in terms of s is

$$p_0(s) \propto \frac{1}{s}. \quad (12)$$

The “precision,” θ , is commonly defined to be the reciprocal of the variance, so that

$$\theta \equiv \frac{1}{s^2}. \quad (13)$$

The posterior probability density is proportional to the product of the prior $p_0(s)$ times the likelihood $L(s; \mathbf{y})$. When the change of variables (13) is plugged into the product of (12) times (11), the posterior becomes conveniently expressed in terms of θ by the gamma distribution

$$\Gamma(\theta) \propto \theta^{a-1} \exp(-b\theta), \quad (14)$$

where

$$a = \frac{n}{2} \quad (15)$$

and

$$b = \frac{n\nu_n}{2}. \quad (16)$$

The mean of the gamma distribution (14) is a/b while its variance is a/b^2 , so that from (15) and (16),

$$E[\theta | \mathbf{y}] = \frac{1}{\nu_n} \quad (17)$$

while

$$V[\theta | \mathbf{y}] = \left(\frac{2}{\nu_n}\right) \frac{1}{n}. \quad (18)$$

After integrating out the precision from the conditional-normal distribution (9), the unconditional or marginal posterior-predictive PDF of the growth rate Y is

$$f(y) \propto \int_0^\infty \sqrt{\theta} \exp(-\theta(y - \mu)^2/2) \Gamma(\theta) d\theta. \quad (19)$$

Straightforward brute-force integration of (19) (for (14), (15), (16)) shows that $f(y)$ is the Student- t distribution with n degrees of freedom:

$$f(y) \propto \left(1 + \frac{(y - \mu)^2}{n\nu_n}\right)^{-(n+1)/2}. \quad (20)$$

(Any Bayesian textbook shows that a normal with gamma precision becomes a Student- t). Note that, asymptotically, the limiting tail behavior of (20) is a fat-tailed power-law PDF

whose exponent is $n + 1$. In the next section of the paper, this same kind of power-law-tail result will be shown to hold much more generally than the particular example of this section.

When the posterior-predictive distribution of Y is (20) (from s being unknown), then (6) becomes

$$E[M] = +\infty, \tag{21}$$

because the moment-generating function of a Student- t distribution is infinite. What accounts technically for the economically-stunning counterintuitiveness of the finding (21) is a form of pointwise but nonuniform convergence. When $n \rightarrow \infty$ in (20), $f(y)$ becomes the familiar normal form $\exp(-(y - \mu)^2/2\nu_\infty^2)$, which then, as $y \rightarrow -\infty$, approaches zero faster than $\exp(-\eta y)$ approaches infinity, thereby leading to the well-known finite formula (7) for $E[M]$. Given any fixed n , on the other hand, as $y \rightarrow -\infty$ expression (20) tends to zero only as fast as the power-law polynomial $(-y)^{-n-1}$, so that now in formula (6) it is the exponential term $\exp(-\eta y)$ that dominates asymptotically, thereby causing $E[M] \rightarrow +\infty$.

Something quite extraordinary seems to be happening here, which is crying out for further elucidation! Thousands of applications of EU theory in thousands of articles and books are based on formulas like (7) or (8). Yet when it is acknowledged that s is unknown (with a standard noninformative reference prior) and its value in formula (7) or (8) must instead be inferred as if from a data sample that can be *arbitrarily large* (but finite), expected marginal utility explodes in (21). The question then naturally arises: what is EU theory trying to tell us when its conclusions for a host of important applications – in CBA, asset pricing, and many other fields of economics – seem so sensitive merely to the recognition that conditioned on finite realized data the distribution implied by the normal is the Student- t ?

I want to emphasize as emphatically as I can at this relatively early stage of the paper that the problem (both here *and throughout the rest of the paper*) is *not* a mathematically illegitimate use of the symbol ∞ in formula (21), which incorrectly seems to offer an easy way out of the dilemma that $E[M] \rightarrow +\infty$ by somehow discrediting this application of EU theory on the narrow grounds that infinities are not allowed in a legitimate theory of choice under uncertainty. It is easy to put arbitrary bounds on utility functions, or to truncate probability distributions arbitrarily, or to introduce *ad hoc* priors that arbitrarily cut off or otherwise severely dampen high values of S or low values of C . Introducing any of these changes formally closes the model in the sense of replacing the symbol ∞ by an arbitrarily-large but finite number. Indeed, the model of this paper will be closed later in just such a fashion by placing a lower bound on consumption of the form $C \geq D$, where the lower bound $D(\lambda) > 0$ is defined indirectly by a “value of statistical life” parameter λ . However, removing the infinity symbol in this or any other way does not eliminate (or even marginalize) the underlying problem because it then comes back to haunt in the form of an arbitrarily large

expected stochastic discount factor, whose exact value depends hypersensitively upon obscure bounds, truncations, severely-dampened or cut-off prior PDFs, or whatever other tricks have been used to banish the ∞ symbol. One can easily remove the ∞ in formula (21), but one cannot so easily remove the underlying economic problem that expected stochastic discount factors – which lie at the heart of cost-benefit, asset-pricing, and many other important applications of EU theory – can become arbitrarily large just by making what seem like unobjectionable statistical inferences about limiting tail behavior. The take-away message here is that reasonable attempts to constrict the length or the fatness of the “bad” tail (or to modify the utility function) still can leave us with uncomfortably big numbers whose exact value depends non-robustly upon artificial constraints or parameter settings that we really do not understand.

What is happening in this example is a particular instance of a general idea. To repeat: *the* core underlying problem is the difficulty of learning limiting tail behavior inductively from finite data. Seemingly thin-tailed probability distributions (like here the normal), which are actually only thin-tailed *conditional on known structural parameters* of the model, become tail-fattened (like here the Student- t) after integrating out the uncertainty. This core issue cannot be eliminated in any clean way, and of necessity it must influence any utility function that is sensitive to low values of consumption. Utility isoelasticity *per se* is inessential to the reasoning here (although it makes the argument easier to understand), because the expected stochastic discount factor $E[M]$ is $+\infty$ in this setup for *any* relatively-risk-averse utility function satisfying the curvature requirement: $\inf_{C>0} [-CU''(C)/U'(C)] > 0$.

The Student- t “child” posterior-predictive density from a large number of observations looks almost exactly like its bell-shaped normal “parent” except that the probabilities are somewhat more stretched out, making the tails appear relatively fatter at the expense of a slightly-flatter center. Intuitively, a normal density “becomes” a Student- t from a tail-fattening spreading-apart of probabilities caused by the variance of the normal having itself a (inverted gamma) probability distribution. It is then no surprise from EU theory that people are more averse qualitatively to a relatively fat-tailed Student- t posterior-predictive child distribution than they are to the relatively thin-tailed normal parent which begets it. A perhaps more surprising consequence of EU theory is the *quantitative strength* of this endogenously-derived aversion to the effects of unknown tail-structure. The story behind this quantitative strength is that fattened-posterior bad tails represent structural or deep uncertainty about the possibility of rare high-impact disasters that – using colorful language here – “scare” any agent having a utility function with strictly positive relative risk aversion. One obvious technical way for the model to contain this scary “Student- t explosion” (or, more generally, “power-law-tail explosion”) is to exclude it *a priori* by imposing some kind

or another of an absolute lower bound $D > 0$ on allowed values of $C \geq D$.

The next issue to be investigated is the extent to which this particular example generalizes. It turns out that there is a rigorous sense in which this scary fattened-posterior-tail effect holds for (essentially) *any* probability distribution characterized by having an uncertain scale parameter. When indeterminateness is compounded by probability distributions themselves having probability distributions, then posterior-predictive tails must inevitably become fattened, with potentially strong consequences for some important applications of EU theory – including the economic analysis of climate change.

3 The General Model

In order to focus sharply on structural parameter uncertainty, the model of this section is patterned closely as a generalization of last section's example and is deliberately sparse. To create families of probability distributions that are simultaneously fairly general and analytically tractable, the following generating mechanism is employed. Suppose Z represents a random variable normalized to have mean zero and variance one. Let $\phi(z)$ be *any* piecewise-continuous PDF with

$$\int_{-\infty}^{\infty} z \phi(z) dz = 0 \tag{22}$$

and with

$$\int_{-\infty}^{\infty} z^2 \phi(z) dz = 1. \tag{23}$$

It should be noted that the PDF $\phi(z)$ is allowed to be extremely general. For example, the distribution of Z might have finite support (like the uniform distribution, which signifies that unbounded catastrophes will be absolutely excluded *conditional* on the value of the finite lower support being known), or it might have unbounded range (like the normal, which allows unbounded catastrophes to occur but assigns them a thin bad tail *conditional* on the variance being known). The *only* restrictions placed on $\phi(z)$ are the extremely weak regularity or technical conditions that $\phi(0) > 0$ with $\phi(z)$ being continuous in an open neighborhood of $z = 0$, and that

$$\int_{-\infty}^{\infty} \exp(-\alpha z) \phi(z) dz < \infty \tag{24}$$

for all $\alpha > 0$, which is automatically satisfied if Z has finite lower support.

With $s > 0$ given, make the change of random variable of the linear form $y = sz + \mu$, which implies that the conditional PDF of y is

$$h(y | s) = \frac{1}{s} \phi\left(\frac{y - \mu}{s}\right), \quad (25)$$

where s and μ are structural parameters having the interpretations

$$\mu = E[Y | s] = \int_{-\infty}^{\infty} y h(y | s) dy \quad (26)$$

and

$$s^2 = V[Y | s] = \int_{-\infty}^{\infty} (y - \mu)^2 h(y | s) dy. \quad (27)$$

For this paper, what matters most is structural uncertainty about the scale parameter controlling the tail spread of a probability distribution, which is the most critical unknown in this setup. This scale parameter s may be conceptualized extremely loosely as a highly-stylized abstract generalization of a climate-sensitivity amplifying or scaling multiplier. (In this crude analogy, $Z \leftrightarrow \Delta \ln CO_2 / \ln 2$, $SZ \leftrightarrow \Delta T$, $Y \leftrightarrow A - B\Delta T$.) Without significant loss of generality, it is assumed for ease of exposition that in (25) the mean μ *is known*, while the standard-deviation *scale parameter* s *is unknown*. The case where μ and s are *both* unknown involves more intricate notation but otherwise gives essentially identical results.

The point of departure here is that the conditional PDF of growth rates $h(y | s)$ given by (25) is known to the agent and, while the true value of s is unknown, the situation is *as if* some finite number of i.i.d. observations are available on which to base an estimate of s by some process of inductive reasoning. Suppose that the agent has observed the random sample $\mathbf{y} = (y_1, \dots, y_n)$ of growth-rate data realizations from n independent draws of the distribution $h(y | s)$ defined by (25) for some unknown fixed value of s . An example relevant to this paper is where the sample space represents the outcomes of various economic-scientific studies and the data $\mathbf{y} = (y_1, \dots, y_n)$ are interpreted at a very high level of abstraction as the findings of n such studies, so that n here is then a measure of the degree of inductive knowledge of the situation.

From (25) the relevant likelihood function of s is

$$L(s; \mathbf{y}) \propto \prod_{j=1}^n h(y_j | s). \quad (28)$$

The prior PDF of S is taken to be a generalization of (12) of the form

$$p_0(s) \propto s^{-k} \quad (29)$$

for any number k . As k can be chosen to be *arbitrarily large*, the non-dogmatic prior distribution (29) can be made to place *arbitrarily small* prior probability weight on big values of s . It should be appreciated that *any invariant prior must be of the form (29)*. Invariance (discussed in the Bayesian-statistical literature) is considered desirable as a description of a “noninformative” reference prior that favors no particular value of the scaling parameter s over any other. For such a noninformative reference prior, it seems reasonable to impose a condition of scale invariance that might be justified by the following kind of reasoning. If the action taken in a decision problem should not depend upon the unit of measurement, then a plausible principle of scale-parameter invariance might ask of a complete-ignorance prior that

$$p_0(s) \propto p_0(\alpha s), \quad (30)$$

and the only way that (30) can then hold for all $\alpha > 0$ and all $s > 0$ is when the (necessarily improper) PDF is of the form (29).

The posterior probability density $p_n(s | y)$ is proportional to the product of the prior $p_0(s)$ from (29) multiplied by the likelihood $L(s; \mathbf{y})$ from (28), which here yields

$$p_n(s | \mathbf{y}) \propto p_0(s) \prod_{j=1}^n h(y_j | s). \quad (31)$$

Integrating out the agent’s uncertainty about s described by the probability density (31), the unconditional or marginal posterior-predictive density of the growth-rate random variable Y is

$$f(y) = \int_0^{\infty} h(y | s) p_n(s | \mathbf{y}) ds, \quad (32)$$

and (6) then becomes

$$E[M] = \beta \int_{-\infty}^{\infty} e^{-\eta y} f(y) dy. \quad (33)$$

4 The Key Role of a “VSL-like Parameter”

To jump ahead of the story just a bit, the general model of Section 3 has essentially the same unsettling property as the disturbing example of Section 2 – namely that $E[M]$ is unbounded.

Technically, for the analysis to proceed further some mathematical mechanism is required to close the model by bounding $E[M]$. A variety of bounding mechanisms are possible, with the broad general conclusions of the model not being tied to any one particular mechanism. In this paper I close the model by placing an *ad hoc* positive lower bound on consumption, which is denoted D (for “death”). This lower bound D is not *completely* arbitrary, however, because it can be related conceptually to a kind of “value of statistical life” (VSL) parameter. This has the advantage of tying conclusions to a familiar economic concept whose ballpark estimates can at least convey some very crude quantitative implications for the economics of climate change. In this empirical sense the glass is half full. However, the glass is half empty in the empirical sense that an accurate CBA of climate change can end up being distressingly dependent on some very large VSL-like coefficient about whose size we are highly unsure.

The critical coefficient that is behind the lower bound on consumption is called the *VSL-like parameter* and is denoted λ . This “VSL-like parameter” λ is intended to be akin to the already-somewhat-vague concept of the value of a human statistical life, only in the context here it represents the yet-far-fuzzier concept of something more like the value of statistical life on earth as we know it, or perhaps the value of statistical civilization as we know it. In this paper I am going to take λ to be some very big number that indirectly controls the convergence of the integral defining $E[M]$ by implicitly generating a lower bound $D(\lambda) > 0$ on consumption. An empirical first approximation of λ (normalized per capita) might be given by conventional estimates of the value of a statistical human life, which may be too small for the purposes at hand but will at least give some crude empirical idea of what is implied numerically.

The parameter λ that is being used here to truncate the extent of catastrophic damages is akin to the “fear of ruin” coefficient introduced by Aumann and Kurz (1977) to characterize an individual’s “attitude toward risking his fortune” in binary lotteries. Foncel and Treich (2005) later analyzed this fear-of-ruin coefficient and showed that it is the same thing analytically as VSL. The particular utility function I use here is essentially identical (but with a different purpose in a different context) to a specification used recently by Hall and Jones (2007), which, it was argued by them, explains broadly an array of stylized facts about health spending while being roughly consistent with empirical VSL estimates.

The basic idea in terms of an application to my model here is that a society trading off a decreased probability of its own catastrophic demise against the cost of lowering the probability of that catastrophe is facing a decision problem conceptually analogous to how people might make private trade-offs between decreased consumption as against a lower probability of their own personally-catastrophic end – which they do all the time. In this spirit, suppose for the sake of developing the argument that the analysis is allowed to proceed

as if the treatment of the most catastrophic conceivable impact of climate change is very roughly analogous to the simplest possible economic model of the behavior of an individual who is trading off increased consumption against a slightly increased probability of death.

Suppose D is some disastrously low value of consumption, conceptualized as representing some kind of an analogue of a starvation level, below which level the individual dies. Let the utility associated with death be normalized at zero. The utility function $U(C; D)$ is chosen to be of the analytically convenient form

$$U(C; D) = \frac{C^{1-\eta} - D^{1-\eta}}{1-\eta} \quad (34)$$

for $C > D$, and

$$U(C; D) = 0 \quad (35)$$

for $0 \leq C \leq D$.

Without loss of generality, current consumption is normalized as it was before at $C = 1$. For simplicity, suppose the agent begins with something close to a zero probability of death in the current period. Let $A(q)$ be the amount of extra consumption the individual requires to exactly compensate for $P[C \leq D] = q$. In free translation, q is the probability of death. From EU theory, $A(q)$ satisfies the equation

$$(1 - q)U(1 + A(q); D) = U(1; D). \quad (36)$$

Differentiating (36) with respect to q and evaluating at $q=0$ gives

$$-U(1; D) + U'(1; D)A'(0) = 0. \quad (37)$$

The “value of statistical life” (or, in the terminology here, the “VSL-like parameter”) is defined as the rate of substitution between consumption and mortality risk, here being

$$\lambda \equiv A'(0). \quad (38)$$

Equations (37), (38) can be inverted to give the implied lower bound on consumption D as an implicit function of the VSL-like parameter λ . Proceeding this way for the isoelastic utility function (1) yields, after manipulation, the equation

$$D(\lambda) = [1 + (\eta - 1)\lambda]^{-1/(\eta-1)}. \quad (39)$$

In this paper I am interested in analyzing what happens for very large values of λ . In order for equation (39) to make sense in such situations by ensuring $D(\lambda) > 0$ requires here that

$$1 + (\eta - 1)\lambda > 0, \tag{40}$$

which can “essentially” (up to limiting measure zero) hold for arbitrarily large values of λ only if

$$\eta > 1. \tag{41}$$

In other words, for what it is worth the theory here is telling us that the large empirical values of statistical life (relative to consumption), which seem to characterize our world, are “essentially” compatible only with the strict relative risk aversion implied by (41).

Very rough ballpark estimates of the per-capita value of a statistical human life might be of the order of magnitude of a hundred times per-capita consumption.³ From a wide variety of empirical studies in disparate contexts, a plausible value of the coefficient of relative risk aversion might be two.⁴ For $\eta=2$, the value $\lambda \approx 100$ plugged into formula (39) gives $D(100) \approx .01$. An interpretation of λ as a parameter representing the per-capita value of statistical civilization or the value of statistical life on earth (as we currently know or understand these concepts) may very well involve higher values of λ than ≈ 100 and lower values of D than $\approx .01$. In any event, I note here for further reference that a Monte Carlo simulation assessing the EU impacts of losing up to 99% of consumption in the far-distant dark reaches of a bad fat tail is very different from any simulations now being done with any existing empirical model or study of the economics of climate change.

Let $E[M \mid \lambda]$ represent the expected value of the stochastic discount factor for $M(C)$ given by formula (4) when $C > D(\lambda)$ and for $M(C)=0$ when $0 \leq C \leq D(\lambda)$. The next section explores what happens to $E[M \mid \lambda]$ for extremely large values of λ .

5 The Dismal Theorem

The following “Dismal Theorem” (hereafter sometimes abbreviated as “DT”) shows that $E[M] \rightarrow +\infty$ under quite general circumstances when there is structural uncertainty concerning the unknown scaling parameter s – *and* when λ might be *very* big.

³For this particular application of using a VSL-like parameter to analyze the extent of the worst imaginable climate-change catastrophe, I think that the *most* one might hope for is accuracy to within about an order of magnitude – anything more being false precision. Even the empirical estimates for the value of a much-better-defined statistical human life have a disturbingly wide range, but $\lambda \approx 100$ is roughly consistent with the meta-analysis in Bellavance, Dione, and Lebeau (2007) or the survey of Viscusi and Aldi (2003).

⁴This is the point estimate selected by Hall and Jones (2007) in a conceptually-similar model and defended by them with references to a wide range of studies on page 61.

Theorem 1 For *any given* n and k ,

$$\lim_{\lambda \rightarrow \infty} E[M \mid \lambda] = +\infty. \quad (42)$$

Proof. Combining the interpretation of $D(\lambda)$ with the appropriate modifications of (33), (32) – and tracing the link of equations all the way back to (25) – implies that

$$E[M \mid \lambda] \propto \int_0^\infty \frac{1}{s^{k+n+1}} \prod_{j=1}^n \phi\left(\frac{y_j - \mu}{s}\right) \left[\int_{\ln D(\lambda)}^\infty e^{-\eta y} \phi\left(\frac{y - \mu}{s}\right) dy \right] ds. \quad (43)$$

Make the change of variable $z = (y - \mu)/s$, use the fact from (39) that $D(\infty) = 0$, and reverse the order of integration to rewrite (43) as

$$\lim_{\lambda \rightarrow \infty} E[M \mid \lambda] \propto \int_{-\infty}^\infty \phi(z) \left[\int_0^\infty e^{-\eta z s} \frac{1}{s^{k+n}} \prod_{j=1}^n \phi\left(\frac{y_j - \mu}{s}\right) ds \right] dz. \quad (44)$$

Pick any value of z' for which simultaneously $z' < 0$ and $\phi(z) > 0$ in an open neighborhood of z' . Then note that

$$\lim_{s \rightarrow \infty} \left\{ e^{-\eta z' s} \frac{1}{s^{k+n}} \right\} = +\infty, \quad (45)$$

implying that expression (44) also approaches $+\infty$ as $\lambda \rightarrow \infty$, which concludes this basic sketch of a proof. ■

The underlying logic behind the strong result of Theorem 1 is described by the limiting behavior of (45) for large values of s . Given any values of n and k , the *probability* of a disaster declines *polynomially* in the scale s of the disaster from (45), while the marginal-utility *impact* of a disaster increases *exponentially* in the scale s of the disaster. It is intuitive, and can readily be proved, that the tail of the random variable Y essentially behaves like the tail of the random variable S . Therefore, *irrespective of the original parent distribution*, the effect of an uncertain scale parameter fattens the tail of the posterior-predictive child distribution so that it behaves asymptotically like a power-law distribution whose power coefficient from (45) is $n+k$. In this sense, power-law tails need not be postulated, because they are essentially unavoidable in posterior-predictive distributions. No matter the number of observations, the race to the bottom of the bad tail between a polynomially-contracting power-law probability times an exponentially-expanding marginal utility impact is won in the limit every time by the marginal utility impact, so long as there is positive relative risk

aversion.⁵

In the model of this paper, people are averse to two types of indeterminacy – what might be called the “risk” of not knowing y given s , and what might be called the “uncertainty” of not knowing s . The Dismal Theorem can be interpreted as providing a rigorous sense in which aversion to structural or deep “uncertainty” (in the form of not knowing the scale parameter s) is potentially far greater than aversion to the pure “risk” of not knowing the realized value of the growth-rate random variable y (but when s is known). However, it is also true that the interpretation and application of Theorem 1 is sensitive to a subtle but important behind-the-scene tug of war between pointwise-but-nonuniform limiting behavior in λ and pointwise-but-nonuniform limiting behavior in n .

To see more clearly how the issue of “risk” vs. “uncertainty” plays out in determining $E[M]$ under pointwise-but-nonuniform convergence, suppose that, unbeknownst to the agent, the true value of s is s^* . Since the prior $p_0(s)$ defined by (29) assigns positive probability to an open interval around s^* , the imposed specification has sufficient regularity for large-sample likelihood dominance to cause strong (i.e., almost sure) convergence of the posterior distribution (31) of S to its true data-generating value $s = s^*$. This in turn means that the posterior-predictive distribution of growth rates (32) converges strongly to its true data-generating distribution $h(y | s^*)$ and $E[M]$ converges strongly to its true value:

$$n \rightarrow \infty \implies E[M | \lambda] \xrightarrow{a.s.} \beta \int_{-\infty}^{\infty} e^{-\eta y} \frac{1}{s^*} \phi\left(\frac{y - \mu}{s^*}\right) dy. \quad (46)$$

Condition (46) signifies that for *any given* $\lambda < \infty$ (which via (39) puts a positive lower bound $D(\lambda)$ on C , and thereby a finite upper bound on M), in the limit as full structural knowledge is approached (because $n \rightarrow \infty$) pure “risk” is more important than structural “uncertainty.” What is happening here is that as the strength of inductive knowledge n is increasing in the form of more and more data observations piling up, it is becoming increasingly apparent that the probability of C being anywhere remotely as low as $D(\lambda)$ is ignorable – even after taking into account the possible EU impacts of disastrously-low utilities. A conventional pure-risk-like application of thin-tail EU theory essentially corresponds to a situation where there is enough inductive-plus-prior knowledge to identify the relevant structure because $n+k$ is reasonably large relative to the VSL-like parameter λ – and relative

⁵As stated here, DT depends upon a prior of the polynomial form (29), but this is not much of a limitation because k can be *any number*. To undo the infinite limit in (42) requires a prior that approaches zero faster than *any polynomial in* $1/s$ (as $s \rightarrow \infty$). In this case the limit in (42) is a finite number, but its (potentially arbitrarily large) value will depend critically upon the strong *a priori* knowledge embodied in the presumed-known parameters of such a prior – and the prior-sensitivity message that such a formulation ends up delivering is very similar to the message delivered by the model of this paper.

to the much-less-controversial parameters β and η .

Concerning the conventional parameters β and η , we have at least some very rough idea of what kinds of numbers might be empirically relevant. (For example, reasonable point estimates might involve values of β consistent with $\approx 99\%$ per year and η consistent with ≈ 2 .) In complete contrast, any discussion about climate change concerning the empirically-relevant value of the nonconventional VSL-like parameter λ (relative to $n+k$) belongs to a much more abstract realm of discourse. It is therefore understandable to want climate-change CBA to be restricted to dealing with worst-conceivable damages of, say for example, $\approx 25\%$ of consumption (equivalent to picking $D \approx 3/4$) – with anything worse than this being disregarded (as being “too speculative” or “not based on hard science”) by chopping off the rest of the bad tail, discarding it, and then forgetting about it. This is actually the *de facto* strategy employed by most of those relatively few existing CBAs of climate change that even bother to concern themselves at all with uncertainty regarding high-impact damages. Alas, to be confident in the validity of such a cutoff strategy in a situation where we are grossly unsure about λ or D (relative to $n+k$) effectively requires *uniform* convergence of $E[M]$ for *all* (conceivable) values of λ or D . Otherwise, for any given level of inductive-plus-prior knowledge $n+k$, a skeptical critic could always come back and ask how robust is the CBA to the highly-unsure truncation value of $D(\lambda)$. Similar robustness questions apply to any *a priori* presumption or imposition of thin-tailed PDFs.

In the spirit of the above discussion, it is critical here to note that with (46) the a.s. convergence of $E[M | \lambda]$ to its true value is *pointwise but not uniform in n* . No matter how much data-evidence n exists – or even can be *imagined* to exist – DT says that $E[M | \lambda]$ is always exceedingly sensitive to very large values of λ . In this sense, structural “uncertainty” always has the potential to trump pure “risk” for situations of potentially-unlimited downside exposure when no plausible bound $D(\lambda) > 0$ can confidently be imposed by prior knowledge. Here “risk” means that the data generating process is known exactly (only the outcome is random), while “uncertainty” means that (as well as the outcome being random) the parameters of the data generating process are unknown and must be estimated statistically.

This paper has stressed repeatedly the basic theme that no finite sample can accurately assess relative probabilities or magnitudes of the most extreme disasters lurking in the dark-distant tails of distributions. The dominant statistical-economic truth behind DT is that the expected utility of a Bayesian agent with strict relative risk aversion may be driven to an arbitrarily large extent by this seemingly unavoidable limitation of learning-inference *unless* potential exposure is somehow limited by prior information. The Dismal Theorem can therefore be interpreted as implying a spirit in which it may be unnecessary to append to the theory of decision making under uncertainty an *ad hoc* extra postulate of “ambiguity

aversion.” At least for situations where there is fundamental uncertainty about catastrophes coexisting with a fear of ruin, EU theory itself already tells us precisely how the “ambiguity” of structural-parameter uncertainty can be especially important and why people may be much more averse to it than to pure objective-frequency “risk.”

The Dismal Theorem has both a general point and a particular application to the economics of climate change. The general point is that Theorem 1 embodies a very strong form of a “generalized precautionary principle” for situations of potentially unlimited exposure. From experience alone one cannot acquire sufficiently accurate information about the probabilities of disasters in the bad tail to make the expected marginal utility of an extra sure unit of consumption become independent of the VSL-like parameter – thereby potentially allowing this VSL-like-parameter aspect to dominate cost-benefit applications of EU theory. To drive home the main theme of this paper yet again, the underlying general problem that DT is illustrating concerns a fundamental limitation on the ability of empirical learning or inductive knowledge to shed light on extreme events. Even in a stationary world, it is not possible to learn enough about the frequency of tail events from finite samples alone to make expected stochastic discount factors be independent of artificially-imposed bounds on the extent of possibly-ruinous disasters.

The part of the distribution of possible outcomes that can most readily be learned (from inductive information that comes in a form as if conveyed by data) concerns the relatively-more-likely outcomes in the body of the distribution. From previous experience, past observations, plausible interpolations or extrapolations, and perhaps even the law of large numbers, there may be at least some modicum of confidence in being able to construct a reasonable picture of the middle regions of the PDF. As we move towards probabilities in the periphery of the distribution, however, we are increasingly moving into the unknown territory of subjective uncertainty where our probability estimate of the probability distributions themselves becomes increasingly diffuse because the frequencies of rare events in the tails cannot be pinned down by previous experiences or past observations. Climate change generally and climate sensitivity specifically are prototype examples of this general principle, because we are trying to extrapolate inductive knowledge far outside the range of limited past experience from natural forcing experiments. Since the unknown scale parameter S (whose uncertainty drives the economic analysis) is an abstract generalized version of an amplifying multiplier whose role is very crudely analogous to the role of open-ended climate sensitivity, it is then perhaps no accident or surprise that the empirical PDF of climate sensitivity (which is, after all, based on inductive information from data and studies) seemingly has a thick power-law-type tail.

The degree to which the kind of “generalized precautionary principle” embodied in the Dismal Theorem is relevant for a particular application must be decided on a case-by-case “rule of reason” basis. It depends generally upon the extent to which prior λ -knowledge (and prior k -knowledge) combine with inductive-posterior n -knowledge in a particular case to fatten or to thin the bad tail. In the particular application to the economics of climate change, with so obviously limited data and limited experience about the catastrophic reach of climate extremes, to ignore or suppress the significance of rare tail disasters is to ignore or suppress what economic-statistical decision theory is telling us here loudly and clearly is potentially the most important part of the analysis.

Global climate change unfolds over a time scale of centuries and, through the power of compound interest, a standard CBA of what to do now to mitigate GHGs is hugely sensitive to the discount rate that is postulated. This has produced some sharp disagreements among economists about what is an “ethical” value of the rate of pure time preference δ (and the CRRA coefficient η) to use for intergenerational discounting in the deterministic version ($s = 0$) of the Ramsey equation (8) that forms the analytical backbone for most studies of the economics of climate change.⁶ For the model of this paper, which is based on structural uncertainty, arguments about what values of δ to use in equation (7) or (8) translate into arguments about what values of β to use in the model’s structural-uncertainty generalization of the Ramsey equation, which is the expectation formula (33). (A zero rate of pure time preference $\delta = 0$ in (8) corresponds to $\beta = 1$ in (33).) In this connection, Theorem 1 seems to be saying that no matter what values of β or η are selected, so long as $\eta > 0$ and $\beta > 0$ (equivalent to $\delta < \infty$), any big- λ CBA of GHG-mitigation policy should be presumed (until shown otherwise empirically) to be affected by deep structural uncertainty.

Expected utility theory is telling us here analytically that the debate about discounting may be secondary to a debate about the open-ended catastrophic reach of climate disasters. While it is always fair game to challenge the assumptions of a model, when theory provides a generic result (like “free trade is Pareto optimal” or “steady growth eventually outstrips one-time change”) the burden of proof is commonly taken as residing with whoever wants to over-rule the theorem in a particular application. The take-away message here is that the burden of proof in the economics of climate change is presumptively upon whoever wants to model or conceptualize the expected present discounted utility of feasible trajectories under

⁶While this contentious intergenerational-discounting issue has long existed (see, e.g., the various essays in Portney and Weyant (1999)), it has been elevated to recent prominence by publication of the controversial *Stern Review of the Economics of Climate Change* (2007). The *Review* argues for a base case of preference-parameter values $\delta \approx 0$ and $\eta \approx 1$, on which its strong conclusions depend analytically. Alternative views of intergenerational discounting are provided in, e.g., Dasgupta (2007), Nordhaus (2007), and Weitzman (2007b). The last of these also contains a heuristic exposition of the contents of this paper, as well as giving *Stern* some credit for emphasizing the great uncertainties associated with climate change.

greenhouse warming *without* considering that structural uncertainty might matter more than discounting or pure risk *per se*. Such a body-of-the-distribution modeler should be prepared to explain why the fattened bad tail of the posterior-predictive distribution is *not* empirically relevant and does *not* play a significant role in the analysis.

6 Concluding Discussion

A common reaction to the conundrum for CBA implied by the Dismal Theorem is to acknowledge its mathematical logic but to wonder how it is to be used constructively for deciding what to do in practice. After all, horror stories about theoretically-possible speculative disasters can be told for many situations without this aspect necessarily paralyzing decision-making (or freezing the status quo by effectively blocking all progress whenever there exists a theoretical possibility of severe downside risks).

The Dismal Theorem says that in the limit as the VSL-like parameter becomes very large, agents will pay virtually any price to eliminate deep uncertainty – a vexing idea that is difficult to wrap one’s mind around. Is this an incompleteness result showing that the familiar ordinary form of CBA based on EU theory with known thin tails is somehow nullified by a severe restriction on what inductive information can teach us about basic structural uncertainty? Phrased differently, is DT an economics version of an impossibility theorem which signifies that there are fat-tailed situations where economic analysis is up against a strong constraint on the ability of *any* quantitative analysis to inform us *without* committing to a VSL-like parameter and an empirical CBA framework that is based on *some* explicit numerical estimates of the miniscule probabilities of all levels of catastrophic impacts up to absolute disaster? And if ordinary thin-tailed CBA is thereby constrained, then what are we supposed to use in its place? Even if it were true that DT represents a valid economic-statistical precautionary principle which, at least theoretically, might dominate decision making, would not putting into practice this “generalized precautionary principle” freeze all progress if taken too literally? (Should we have foregone the industrial revolution because of the GHGs it generated?) With such a severe barrier to what we are able to say in a situation of deep structural uncertainty (when it is combined with practically unlimited potential for downside exposure), what should we do when *some* kind of a policy response to a complex chaos-prone dynamic system like climate change is required?

I simply do not know the full answers to the above wide range of legitimate questions that DT raises. I don’t think anyone does. But I also don’t think that such questions can be deflected aside in good conscience. When uncertainty is considered at all in climate change, the artificial practice of using thin-tailed PDFs – especially the widespread practice

of imposing *de minimis* low-probability-threshold cutoffs that casually dictate what part of the high-impact bad tail is to be truncated and discarded from CBA – seems (to me) arbitrary and problematic.⁷

Any interpretation or application of the Dismal Theorem is rendered exceedingly tricky by the unusual (for economics) nonuniform convergence of $E[M]$ or $E[U]$ in all of its *other* parameters relative to the key VSL-like parameter λ . This nonuniform convergence enables $E[M]$ or $E[U]$ to explode (for any other given parameter values) as $\lambda \rightarrow \infty$. One might try to argue that the values of $E[M]$ or $E[U]$ are ultimately an empirical matter to be decided empirically (by analytical formulas or simulation results), with relevant parameter values of λ , n , k , δ , η , μ and so forth being taken together as an empirically-plausible ensemble. The idea that the values of $E[M]$ or $E[U]$ should depend on empirically-reasonable values of λ and the other parameters is, of course, right on some level – and it sounds reassuring. Yet, as a practical matter, the fact that $E[M]$ and $E[U]$ are so sensitive to large values of λ (or small values of D , about which we can have little confidence in our own *a priori* knowledge) casts a very long shadow over any empirical CBA of a situation to which DT might apply. In ordinary run-of-the-mill limited-exposure thin-tailed situations, there is at least the underlying theoretical reassurance that finite-cutoff-based CBA might (at least in principle) be an arbitrarily-close *approximation* to something that is accurate and objective. In fat-tailed unlimited-exposure DT situations, by contrast, there is no such theoretical assurance underpinning the arbitrary cutoffs, which is ultimately due to the lack of uniform convergence of $E[M]$ or $E[U]$ with respect to λ or D .

One does not want to abandon lightly the ideal that CBA should bring independent empirical discipline to any application by being based upon empirically-reasonable parameter values. Even when the Dismal Theorem might apply, CBA based upon empirically-reasonable parameter values (including λ) might reveal useful information. At the same time, one does not want to be obtuse by insisting that DT *per se* makes no practical difference for CBA because the VSL-like coefficient λ is in any event just another parameter whose value is to be determined empirically and then simply plugged into the analysis along with some guesses about “the damages function” for catastrophes (combined with speculative extreme-tail probabilities). So some sort of a tricky balance is required between being overawed by DT into abandoning CBA altogether and being underawed by DT into insisting that it is just another empirical issue to be sorted out by business-as-usual CBA.

There exist perhaps half a dozen or so serious “nightmare scenarios” of environmental disasters comparable in conceivable worst-case impact to catastrophic climate change. These

⁷Adler (2007) sketches out in some detail the many ways in which *de minimis* low-probability-threshold cutoffs are arbitrary and problematic in more-ordinary regulatory settings.

might include: biotechnology, nanotechnology, asteroids, strangelets, pandemics, runaway computer systems, nuclear proliferation.⁸ It may well be that each of these possibilities of environmental catastrophe deserves its own CBA application of DT. Even if this were true, however, it would not lessen the need to reckon with the strong potential implications of DT for CBA in the particular case of climate change.

Perhaps it is no more than just an intuition, but for what it is worth I somehow don't feel that the handful of other conceivable environmental catastrophes are as critical as climate change. Let me illustrate some of what I have in mind with the specific example of the widespread cultivation of crops based upon bioengineered genetically-modified organisms (GMOs). At first glance, the two situations might appear casually similar. In both cases, there is deep unease about artificial tinkering with the natural environment, which can generate frightening tales of a world turned upside down and a planet that has been ruined by human activity. Suppose for the sake of argument that in the case of GMOs the overarching fear of disaster concerns the possibility that widespread cultivation of so-called "Frankenfood" could somehow allow bioengineered genes to escape into the wild and wreak havoc on delicate ecosystems and native populations (including, perhaps, humans), which have been fine-tuned by millions of years of natural selection. At the end of the day I think that the potential for environmental disaster with Frankenfood is somewhat different from the potential for environmental disaster with climate change – along the lines of the following loose reasoning.

Theorem 1 and the subsequent discussion of nonuniform convergence in λ and n imply that deep uncertainty about the unknown scale of a disaster has the potential to dominate EU cost-benefit calculations when the amount of prior-plus-inductive knowledge $k+n$ is small relative to the VSL-like parameter λ of such disasters. I think that in the case of Frankenfoods interfering with wild organisms that have evolved by natural selection, there is at least *some* basic underlying principle that plausibly dampens the extent of catastrophic jumping of artificial DNA from cultivars to landraces. After all, nature herself has already tried endless combinations of mutated DNA and genes over countless millions of years, and what has evolved in the fierce battle for survival is only an infinitesimal subset of the very fittest permutations. In this regard there exists at least some inkling of a prior argument making it fundamentally implausible that Frankenfood artificially selected for traits that humans find desirable will compete with or genetically alter the wild types that nature has selected via Darwinian survival of the fittest. Wild types have already experienced innumerable small-step genetic mutations perhaps comparable to large-step human-induced artificial modifications – and these natural small-step modifications have already been shown

⁸Many of these are discussed in Posner (2004), Sunstein (2007), and Parson (2007).

not to have survival value in the wild. I think that analogous arguments may also apply for invasive “superweeds,” which (at least so far) represent a minor cultivation problem and have not displayed any ability to displace landraces. Besides all this, and importantly in this context, safeguards in the form of so-called “terminator genes” can be inserted into the DNA of GMOs, which directly prevents GMO genes from reproducing themselves.

Contrast the above discussion about the plausible limits of disaster (and the limits of a precautionary principle) in the case of genetic engineering with the lack of any such prior limitations on the magnitude of disasters possible from climate change. The climate-change “experiment,” whose eventual outcome we are trying to infer now, concerns the planet’s response to a geologically-instantaneous exogenous injection of GHGs. Such a “planetary experiment” of an exogenous injection of *this much GHGs this fast* seems unprecedented in Earth’s history stretching back perhaps even hundreds of millions of years. Can anyone honestly say now, from very limited information or experience, what are reasonable upper bounds on the eventual global warming or climate change that we are currently trying to infer will be the outcome of such a first-ever planetary experiment? What we *do* know about climate science and extreme tail probabilities is that planet Earth hovers in an unstable climate equilibrium⁹, chaotic dynamics cannot be ruled out, and all eighteen current studies of climate sensitivity cited by IPCC4 taken together are estimating on average that $P[S > 6^\circ\text{C}] \approx 5\%$. To my mind this open-ended aspect makes GHG-induced global climate change vastly more worrisome than global cultivation of Frankenfood.

I think this example perhaps suggests that it may (barely) be possible to make a few meaningful distinctions among the handful of situations where the Dismal Theorem might conceivably apply seriously. While my discussion here is hardly conclusive and we surely cannot rule out a biotech disaster, I would say on the basis of this line of argument that it seems to me *very very* unlikely, whereas a climate disaster seems to me “only” *very* unlikely. In the language of this paper, synthetic biology seems to me more like a high- $(k+n)$ situation *relative to* climate change, which by comparison feels to me more like a low- $(k+n)$ situation. It is true that a climate-change catastrophe develops more slowly than some other potential catastrophes and there is perhaps somewhat more chance for learning and mid-course corrections relative to, say, biotechnology. So the possibility of “learning by doing” may be a more distinctive feature of climate change disasters than other disasters and should be part of an optimal climate-change policy. This being said, the climate response to GHGs has tremendous built-in inertial lags and I don’t think there is a smoking gun in the biotechnology scenario – or in most other catastrophe scenarios – quite like having average expert assessment in climate change be that $P[S > 8^\circ\text{C}] \approx 2\%$.

⁹On the nature of this unstable climate equilibrium, see Hansen et al (2007).

I have tried to make an example contrasting climate change with synthetic biology and I have tried to argue that the latter feels less dangerous than the former. However, this is ultimately a subjective assessment intended only to illustrate a possible thought process. The applicability of DT to each catastrophe scenario must be decided on a case by case basis. To repeat, the fact that DT might also apply to some other environmental catastrophes, like biotechnology, is *not* a valid reason for excluding it from applying to climate change.

Many situations have natural limited-exposure-like bounds on the possible damages that might materialize, for which the theory of this paper may be less relevant or perhaps even have no relevance. When 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 tail fatness is not such a paramount concern. But a few 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. This paper shows that the EU analysis of those relatively few deep-uncertainty situations with potentially catastrophic reach – like climate change – can give very different conclusions than what might emerge from a typical CBA of a more ordinary limited-exposure situation. The conclusions can even be very different from what would come out of a *conventional* Monte Carlo simulation of a model whose core structure resembles the model of this paper but with a discrete-grid approximation or with a finite number of runs. The dismal outcome embodied in Theorem 1 of this paper can be approached by a Monte Carlo simulation only as a double limit where the grid size and the number of runs both go to infinity.

A so-called “Integrated Assessment Model” (hereafter “IAM”) for climate change is a multiple-equation computer-simulated model that combines dynamic economics with geophysical and geochemical circulation dynamics for purposes of analyzing the economic impacts of global climate change. An IAM is essentially a model of economic growth with a controllable externality of endogenous greenhouse warming. IAMs have proven themselves indispensable for understanding the economics of climate change – if for no other reason than being able to describe outcomes from a complicated interplay of the long lags and big inertias that are involved. Most existing economic analyses of climate change treat central forecasts of damages as if they were certain, or, if they address uncertainty at all, use thin-tailed PDFs including, especially, truncation of PDFs at an arbitrary level. When viewed in the light of the Dismal Theorem, therefore, most conventional climate-change IAMs may have limitations that have not yet been sufficiently recognized or acknowledged.

Samplings based upon conventional Monte Carlo simulations of the economics of climate

change with any existing IAM may give a very misleading picture of the EU consequences of alternative GHG-mitigation policies.¹⁰ The core underlying problem is that while it might be true in *expectations* that utility-equivalent damages of climate change are enormous, when chasing a fat tail this will not be true for the overwhelming bulk of Monte Carlo *realizations*. To see this most clearly in a crisp thought experiment, imagine what would happen to the simple stripped-down model of this paper in the hands of a Monte Carlo IAM simulator.

A finite grid may not reveal the true expected stochastic discount factor or true expected discounted utility in simulations of this model (even in the limit of an infinite number of runs) because the most extreme negative impacts in the fattened tails will have been truncated and evaluated at but a single point representing an artificially-imposed lower bound on the set of all possible values of consumption-equivalent outcomes from all conceivable negative impacts. Such arbitrarily-imposed *de minimis* threshold-cutoff truncations are typically justified (when anyone bothers to justify them at all) on the thin-tailed logic that probabilities of extremely rare events are statistically insignificantly different from zero – and hence can be ignored. This logic might conceivably suffice for known thin tails, but the conclusion is highly incorrect for the rare and unusual class of fat-tailed potentially-high-impact economic problems to which climate change seemingly belongs. Back-of-the-envelope calculations cited earlier in this paper appear to indicate that the validity of a Monte Carlo simulation of the economics of climate change requires seriously probing into the implications of disastrous temperatures and catastrophic impacts in incremental steps that might conceivably cause up to a 99% (or maybe even more) decline of welfare-equivalent consumption before the modeler is allowed to cut off the rest of the bad fat tail and discard it. This paper says that any climate-change IAM that does *not* go out on a limb by explicitly stating and including in a Monte Carlo simulation the ultra-miniscule but fat-tailed probabilities of ultra-catastrophic impacts (down to about one percent or so of current welfare-equivalent consumption) is in possible violation of best-practice economic analysis because (by ignoring the extreme tails) it could constitute a serious misapplication of EU theory. The policy relevance of such a model might then remain under a dark cloud until this fat-tail issue is addressed seriously and resolved empirically.

An additional consideration is that a finite sample of Monte Carlo simulations may not reveal true expected utility in this model (even in the limit of an infinite grid) because the limited sample may not be able to go deep enough into the fattened tails where the most extreme damages are. Nor will typical sensitivity analysis of Monte Carlo simulations

¹⁰Tol (2003) showed the empirical relevance of this issue in some actual IAM simulations. I am grateful to Richard Carson for suggesting the inclusion of an explicit discussion of why a Monte Carlo simulation may fail to account fully for the implications of uncertain large impacts with small probabilities.

necessarily penetrate sufficiently far into the fattened-tail region to represent accurately the EU consequences of disastrous damages. For any IAM (which presumably has a core structure resembling the model of this paper), special precautions are required to ensure that Monte Carlo simulations represent accurately the low-utility impacts of fat-tailed PDFs.

Instead of the existing IAM emphasis on estimating or simulating the economic impacts of what are effectively the more plausible climate-change scenarios, to at least compensate partially for finite-sample bias the model of this paper calls for a dramatic *oversampling* (relative to probability of occurrence) of those stratified climate-change scenarios associated with the most adverse imaginable economic impacts in the bad fat tail. With limited sampling resources for the big IAMs, Monte Carlo analysis could be used much more creatively – not to defend a specific policy result, but to experiment seriously with what happens to the outcomes with fat-tailed uncertainty in the limit as the grid size and number of runs simultaneously increase. Of course this emphasis on sampling climate-change scenarios in proportion to marginal-utility-weighted probabilities of occurrence forces us to estimate subjective probabilities down to extraordinarily tiny levels and also to put utility weights on disasters with damage impacts up to perhaps being welfare-equivalent to losing 99% of consumption – but that is the price we must be willing to pay for having a genuine economic assessment of potentially-catastrophic climate change.

In situations of potentially unlimited damage exposure – like climate change – a reframing of the focus of economic analysis might be appropriate, with more emphasis on even a slightly better treatment of the worst-case tail extremes (and what might be done about them, at what cost) relative to refining the calibration of merely-likely outcomes or arguing over discount rates. A clear implication of this paper is that greater research effort is relatively ineffectual when it is targeted at describing and estimating the central tendencies of what we already know fairly well about the economics of climate change in the more-plausible scenarios. A much more fruitful goal of research in my opinion is to aim at understanding even slightly better the deep uncertainty (which potentially permeates the economic analysis) concerning the *less* plausible scenarios located in the bad fat tail. I also believe that an important complementary research goal is to comprehend much better *all* of the options for dealing with high-impact climate-change extremes, which should include undertaking well-funded detailed studies and experiments about the feasibility and cost-effectiveness of geoengineering options to slim down the bad fat tail quickly in case of emergency.¹¹

¹¹It is a secret well kept within the climate-change community of scholars that (with the unfortunately-limited information we currently possess) geoengineering via injection of stratospheric sulfate aerosol precursors looks at first glance now like an incredibly cheap and effective way to slim down the bad fat tail quickly. For more on the economics of geoengineering, see Barrett (2007). In my opinion there is an acute – even desperate – need for much more extensive research on (and experimentation with) various geoengineering

When analyzing the economics of climate change, perhaps it might be possible to reason somewhat by analogy with insurance for extreme events and to compare, say, a homeowner's cost of buying fire insurance or a young adult's cost of buying life insurance with the cost to the world economy of buying an insurance policy going some way towards mitigating the extreme high-temperature possibilities. On a U.S. national level, rough comparisons could perhaps be made with the potentially-huge payoffs, small probabilities, and significant costs involved in countering terrorism, building anti-ballistic missile shields, or eliminating hostile dictatorships that might harbor weapons of mass destruction. A crude natural metric for calibrating cost estimates of climate-change environmental-insurance policies might be that the U.S. already spends approximately 3% of national income on the cost of a clean environment.¹² All of this having been said, the bind we find ourselves in now on climate change starts from a high- λ , low- k situation to begin with, and is characterized by extremely slow convergence in n of inductive knowledge towards resolving the deep uncertainties – relative to the lags and irreversibilities from *not* acting before structure is more fully identified.

The point of all of this is that economic analysis is not completely helpless in the presence of deep structural uncertainty and potentially unlimited exposure. The analysis is more frustrating and more subjective – and it looks a lot less conclusive – because it requires some form of speculation (masquerading as an “assessment”) about the extreme bad-fat-tail probabilities and utilities. Compared with the thin-tailed case, CBA of fat-tailed potential catastrophes is inclined to favor paying a *lot* more attention to learning *now* how fat the bad tail might be and – if the tail is discovered to be too heavy for comfort after the learning process – is a *lot* more prone to investing big-budget real money on mitigation measures to slim it down. This slimming-down of overweight tails is likely to be a perennial theme in the economic analysis of catastrophes. The key economic issues here are: what is the cost of such a tail-slimming weight-loss program and how much of the bad fat does it remove from the overweight tail?

The above attempts at doling out constructive suggestions notwithstanding, it is painfully apparent that the Dismal Theorem makes economic analysis trickier and more open-ended in the presence of deep structural uncertainty – and there is no sense pretending otherwise. The economics of fat-tailed catastrophes raises difficult conceptual issues which cause the analysis to appear less scientifically conclusive and to look more contentiously subjective than what comes out of an empirical CBA of more-usual thin-tailed situations. But if that is the way things are with fat tails, then that is the way things are, and it is a fact to be lived

options for dealing with potential runaway climate change.

¹²U.S. Environmental Protection Agency (1990), executive summary projections for 2000 updated and extrapolated by me to 2007.

with rather than a fact to be evaded just because it looks less scientifically objective in cost-benefit applications. A CBA of a situation with known thin tails, even including whatever elements of subjective arbitrariness it might otherwise have, can at least in principle make statements of the form: “if the tails are cut off here, then EU theory will still be an accurate approximation of what is important.” Such accuracy-of-approximation cutoff statements, alas, are far from being out there automatically for fat-tailed CBA. Perhaps in the end the economist can help most by *not* presenting a cost-benefit estimate for a fat-tailed situation with potentially unlimited exposure as if it is accurate and objective – and not even presenting the analysis as if it is an *approximation* to something that is accurate and objective – but instead by stressing more the fact that such an estimate might conceivably be arbitrarily inaccurate depending upon what is subjectively assumed about the fatness of the tails or where they have been cut off. This is unsatisfying and not what economists are used to doing, but in rare situations like climate change where the Dismal Theorem applies we may be deluding ourselves and others with misplaced concreteness if we think that we are able to deliver anything much more precise than this with even the biggest and most-detailed climate-change IAMs as currently constructed and deployed.

The contribution of this paper is to phrase exactly and to prove rigorously a basic theoretical principle that holds under strict relative risk aversion and potentially unlimited exposure. In principle, what might be called the catastrophe-insurance aspect of such a fat-tailed unlimited-exposure situation, which can never be fully learned away, dominates the social-discounting aspect, the pure-risk aspect, or the consumption-smoothing aspect. Even if this principle in and of itself does not provide an easy answer to questions about how much catastrophe insurance to buy (or even an easy answer in practical terms to the question of what exactly *is* catastrophe insurance buying for climate change or other applications), I believe it still might provide a useful way of framing the economic analysis of catastrophes.

References

- [1] **Adler, Matthew D.** “Why De Minimis?” AEI-Brookings Joint Center Related Publication 07-17, dated June 2007.
- [2] **Aumann, Robert J., and Mordecai Kurz.** “Power and Taxes.” *Econometrica*, 1977, 199, pp. 1137-1161.
- [3] **Barrett, Scott.** “The Incredible Economics of Geoengineering.” Johns Hopkins mimeo dated 18 March 2007.

- [4] **Bellavance, Francois, Georges Dionne and Martin Lebeau.** “The Value of a Statistical Life: A Meta-Analysis with a Mixed Effects Regression Model.” HEC Montreal working paper 06-12, dated 7 January 2007.
- [5] **Dasgupta, Partha.** “Commentary: the Stern Review’s Economics of Climate Change.” *National Institute Economic Review*, 2007, 199, pp. 4-7.
- [6] **Foncel, Jerome, and Nicolas Treich.** “Fear of Ruin.” *The Journal of Risk and Uncertainty*, 2005, 31, pp. 289-300.
- [7] **Geweke, John.** “A note on some limitations of CRRA utility.” *Economics Letters*, 2001, 71, pp. 341-345.
- [8] **Gnedenko, Boris V., and and Andrei R. Kolmogorov.** *Limit Distributions for Sums of Independent Random Variables*. In Russian, 1949. English Translation by K. L. Chung, Cambridge: Addison-Wesley, 1954.
- [9] **Hansen, James et al.** “Climate change and trace gases.” *Phil. Trans. R. Soc. A*, 2007, 365, pp. 1925-1954.
- [10] **Hall, Robert E. and Charles I. Jones.** “The Value of Life and the Rise in Health Spending.” *Quarterly Journal of Economics*, 2007, 122, pp. 39-72.
- [11] **IPCC4.** *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>).
- [12] **Nordhaus, William D.** “The Stern Review on the Economics of Climate Change.” *Journal of Economic Literature*, 45 (3) (September 2007).
- [13] **Parson, Edward A.** “The Big One: A Review of Richard Posner’s *Catastrophe: Risk and Response*.” *Journal of Economic Literature*, 2007, XLV (March), pp. 147-164.
- [14] **Portney, Paul R. and Weyant, John P, eds.** *Discounting and Intergenerational Equity*. Washington, DC: Resources for the Future, 1999.
- [15] **Posner, Richard A.** *Catastrophe: Risk and Response*. Oxford University Press, 2004.
- [16] **Sunstein, Cass R.** *Worst-Case Scenarios*. Harvard University Press, 2007.
- [17] **Schwarz, Michael.** “Decision Making Under Extreme Uncertainty.” Stanford University Graduate School of Business: Ph.D. Dissertation, 1999.

- [18] **Stern, Nicholas et al.** *The Economics of Climate Change*. Cambridge University Press, 2007.
- [19] **Tol, Richard S. J.** “Is the Uncertainty about Climate Change Too Large for Expected Cost-Benefit Analysis.” *Climatic Change*, 2003, 56, pp. 265-289.
- [20] **Viscusi, W. Kip and Joseph E. Aldy.** “The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World.” *Journal of Risk and Uncertainty*, 2003, 5, pp. 5-76.
- [21] **Weitzman, Martin L.** “Subjective Expectations and Asset-Return Puzzles.” *American Economic Review*, 97 (4) (September 2007).
- [22] **Weitzman, Martin L.** “The Stern Review of the Economics of Climate Change.” *Journal of Economic Literature*, 45 (3) (September 2007).
- [23] **U.S. Environmental Protection Agency.** *Environmental Investments: The Cost of a Clean Environment*. Washington, DC: U.S. Government Printing Office, 1990.