
ECONOMIC POLICY IN THE FACE OF SEVERE TAIL EVENTS

WILLIAM D. NORDHAUS

Yale University

Abstract

From time to time, something occurs which is outside the range of normal expectations. We will call these “tail events” in the sense that they are way out of the tail of a probability distribution. I consider the question of the implications of tail events for economic policy and climate-change economics. This issue has been analyzed by Martin Weitzman who proposed a Dismal Theorem. The general idea is that, under limited conditions concerning the structure of uncertainty and risk aversion, society has an indefinitely large expected loss from high-consequence, low-probability events. Under such conditions, standard economic tools such as cost-benefit analysis cannot be applied. The present study is intended to put the Dismal Theorem in context and examine the range of its relevance, with an application to catastrophic climate change. I conclude that tail events are sometimes of extreme importance, and we must be extremely careful to include them in situations of deep uncertainty. However, we conclude that no loaded gun of strong tail dominance has been uncovered to date.

1. Climate Policy in the Age of Tail Events

In an earlier era, climate policy seemed to be a straightforward exercise in weighing costs and benefits in light of the economic costs of reducing emissions and the lower economic damages from reduced concentrations of greenhouse gases (GHGs). Most analyses in this framework called for modest near-term reductions in emissions, followed by increasingly tight controls

William D. Nordhaus, Sterling Professor of Economics, Yale University, 28 Hillhouse Avenue, New Haven, CT 06510, United States (william.nordhaus@yale.edu).

The author is grateful for comments on these issues from William Brainard, Gary Yohe, Richard Tol, Martin Weitzman, three anonymous referees, and the editors. The author has no conflict of interest.

Received October 1, 2010; Accepted July 14, 2011.

© 2012 Wiley Periodicals, Inc.

Journal of Public Economic Theory, 14 (2), 2012, pp. 197–219.

in the decades to come. In this view, the science and policy could improve our estimates as there was no great urgency in terms of either the damages or sharp discontinuities.

Three developments in the natural and social sciences have upended the earlier stance on climate policy. The first was the argument in *The Stern Review* that societies must take a longer-term view than was customary. While debates about the appropriate discount rate have been prominent in the analyses of energy and climate change for decades, emphasizing this point led to a reappraisal of the priority for major near-term emissions reductions.

A second development was the change in the view of climate change from one of gradual and smooth changes to one with potentially abrupt climate change and sharp and irreversible tipping points. This view emphasized the importance of potential thresholds or catastrophic impacts. Important examples that have been discussed are reversal of the Atlantic thermohaline circulation, disintegration of the Greenland and West Antarctic Ice Sheets, shifts in monsoonal patterns, and die-off of the Amazon rain forests. Economic analysis of tipping points and discontinuities in climate change is in its infancy.

The third change has been the increasing concern with the potential of “tail events” in the impacts of climate change. This paper deals with the third of these issues. I focus in particular on the implications of the combination of outcomes that are potentially catastrophic in nature and have “fat tails.” These events might be the result of crossing poorly understood thresholds, so the second development overlaps with this third one. That is, it is not just “low probability, high consequence” situations that are of concern, but we must also consider situations in which the low probabilities do not decline sufficiently rapidly to offset the high consequences. The combination of situations is one in which our standard analyses need to be modified or may even break down. In the extreme case, the combination of fat tails and unlimited exposure implies that the expected loss from certain risks such as climate change is infinite.

The plan of the present study is as follows. We begin in the first section with a discussion of tail events along with a definition of tail dominance. We provide an overview of the basic analysis underlying Weitzman’s Dismal Theorem, and then provide a heuristic example in which the basic structure is easily seen. We put the analysis in the context of policy decisions and note that the usual version of the Dismal Theorem actually contains no policy decisions.

The subsequent section sketches the critical assumptions underlying strong tail dominance as exemplified by the Dismal Theorem. It shows that two central assumptions are unboundedness of both the uncertain variable and the marginal utility of consumption as consumption approaches zero.

The final section explores in greater depth the question of tail dominance in the important case of climate change. This example is motivated both because of its importance in current policy discussions and because it

is the policy example that Weitzman explores in his analysis. We discuss the probability distribution of a critical uncertain variable, the temperature sensitivity coefficient (TSC) that Weitzman analyzes. We also explore the distribution of consumption declines using the analog of consumption “disasters” in economic history. Using two different databases, we find that consumption disasters appear not to have a tail that is sufficiently fat to trigger the Dismal Theorem.

We conclude that a loaded gun of strong tail dominance has not been discovered to date. At the same time, the results of the Dismal Theorem are sufficiently powerful to serve as a reminder that we must constantly be alert to this possibility. Perhaps climate change is not the Dismal event, but there is a sufficiently large number of other high-consequence, low-probability events that we need to pay careful attention to tail events.

1.1. Tail Events and the Dismal Theorem

A tail event is an outcome which, from the perspective of the frequency of historical events or perhaps only from intuition, should happen only once in a thousand or million or centillion years. Momentous tail events were the first detonation of an atomic weapon over Hiroshima in 1945, the sharp rise in oil prices in 1973, the 23% fall in stock prices in October 19, 1987, the destruction of the World Trade Towers in 2001, and the collapse of the world financial system in 2007–2008.

Distributions with fat or heavy tails have a substantial history in statistics.¹ Statisticians have known for almost a century that events with fat-tailed or heavy tailed distributions may behave in an unintuitive way. The Pareto-Lévy distributions are stable distributions (have central limit theorems) and behavior that in the limit has the shape of the Pareto distribution in the tail, but the Pareto distribution is not a stable distribution. The Pareto distribution in the tail is prominent in discussions because, unlike the Pareto-Lévy class, it has an analytical representation for its probability distribution function. There is now a substantial literature on the statistics of fat tails, and it is increasingly diffusing into other areas such as finance, physics, economics, and politics. Much of the empirical work has focused on the extremes of the distributions and has used the Pareto distribution.

For the Pareto distribution, the probability density function of an event X is $f(X) = k X^{-(\beta+1)}$, where $\beta > 0$ is the Pareto “shape parameter,” which reflects the importance of tail events; X is the variable of interest; and k is

¹ There is a vast literature on the subject today. While most current work focuses on the Pareto distribution, which is used in this paper, statisticians have investigated “Pareto-Lévy distributions.” These are a family of stable distributions of which the normal is one and indeed is the only one with a finite variance. See Paul (1923), Lévy (1924), and Feller (1950). The early works in economics in this area by Mandelbrot (1963) and Samuelson (1967) are notable.

a constant that ensures that the sum of probabilities is 1. If β is very small, then the tail is very fat and the variable has a highly dispersed distribution. Depending upon the value of β , the distribution can have infinite variance, or even infinite mean, as in the case of the Cauchy distribution seen below.²

Relatively little work has examined the implications of fat tails for economic modeling and policy. In a recent series of papers, Martin Weitzman (2009) has proposed a dramatically different conclusion from standard analysis in what he has called the Dismal Theorem. This result holds that under certain circumstances the expected value of utility or marginal utility does not exist and we therefore cannot apply standard economic analysis. The circumstances that Weitzman proposes are ones with fat tails and with strong risk aversion.

He summarizes the basic point as well as its application to climate change as follows: “The burden of proof in climate-change cost-benefit analysis (CBA) is presumptively upon whoever calculates expected discounted utilities without considering that structural uncertainty might matter more than discounting or pure risk. Such a middle-of-the-distribution modeler should be prepared to explain why the bad fat tail of the posterior-predictive PDF is not empirically relevant and does not play a very significant—perhaps even decisive—role in climate-change CBA.”³

The presence of consequential tail events is potentially of great importance for both economic modeling and for economic analyses of climate change. The purpose of this note is to put tail events and the Dismal Theorem in context and analyze the range of their applicability. I conclude that Weitzman raises important issues about the selection of distributions in the analysis of decision-making under uncertainty. However, the assumptions underlying the theorems are very restrictive, so the broad claim to have reversed the burden of proof on the use of expected utility analysis needs to be qualified.

1.2. Some Preliminaries on Tail Dominance

The Dismal Theorem concerns how tail events can dominate our analysis. We can describe the point in a qualitative way in terms of what can be described as “tail dominance.” We can classify problems into three classes:

² There is no generally accepted definition of the term fat tails, also sometimes called heavy tails. One set of definitions divides distributions into three classes. A thin-tailed distribution has a finite domain (such as the uniform), a medium-tailed distribution has exponentially declining tails (such as the normal), and a fat-tailed distribution has power-law tails (such as the Pareto distribution). See Schuster (1984). Weitzman proposes a new definition, that a fat-tailed distribution is one whose moment generating function is infinite. As we will see below, this is also the condition for the Dismal Theorem. We will also see that within a class of distributions the condition will depend on incidental parameters such as the degrees of freedom.

³ These quotations are from Weitzman (2009).

- (1) *Tail irrelevance.* Certain problems are ones in which the distribution of the random variable makes no (or little) difference to the policy or the outcome. For example, some problems are ones in which certainty equivalence applies. In these, clearly the tails are irrelevant.
- (2) *Weak tail dominance.* A second class of cases is one where the outcomes or policies are affected by the tails of the distributions but where outcomes or policies converge as we examine the tails more deeply. For example, it might be the case that the best decision on climate-change policy changes markedly as we consider uncertainties, but the answer converges quickly to a best policy as we continue to look at increasingly unlikely events.
- (3) *Strong tail dominance.* A final class of cases is one where there the outcome does not converge as we continue to look at more and more extreme events. In other words, the optimal policy or the outcome does not exist (say because it is unbounded).

1.2.1. Earthquakes and the dismal theorem

The example of earthquakes is a good application of fat tails and the Dismal Theorem. Seismologists have determined that the size of earthquakes tends to follow a power-law distribution, this known as the Gutenberg–Richter Law (see Gutenberg and Richter 1954; Pisarenko and Sornette 2003). Most estimates indicate that the Pareto coefficient on the energy of earthquakes has a parameter of about $\beta = 1$.

Figure 1 shows an illustration using data on the largest recorded earthquakes from the U.S. Geological Survey (2011), but this is a typical result (see Christensen, Danon, Scanlon, and Bak 2002 for another example). The variable E is the energy of the earthquake. This is calculated as the $1.5 \times$ the moment magnitude. For those familiar with the Richter scale, the moment magnitude is based on a refined version of that scale and has the same units. A difference in magnitude of 1 in the Richter or moment-magnitude scale, as between a magnitude 9 and a magnitude 8, converts into a difference in the energy or destructive power of a factor of $10^{1.5} = 31.6$. Research indicates that the trough to crest amplitude of a tsunami for an ocean earthquake is proportional to the square of the magnitude of the earthquake (see Abe 1981).

Now consider how the fat-tailed nature of earthquakes might work in the context of the Dismal Theorem. Suppose that we are designing a structure (perhaps an airport, or a clam shack, or maybe a nuclear power plant) and need to consider what “size” earthquakes to plan for. This would require estimates of damages, costs, and consider risk aversion as well. We cannot consider every possible outcome, so what might we use as a rule of thumb? One rule of thumb might be to plan for the largest earthquake or tsunami that had been experienced in the region. If the historical record is 200 years, then that might to be a reasonable idea. It would consider risks that occur

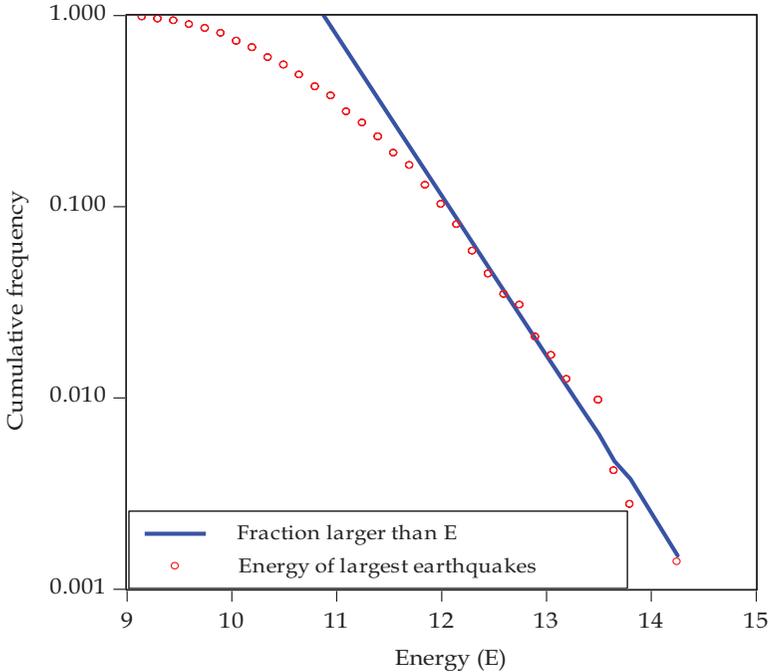


Figure 1: The power law distribution for earthquakes.

Notes: The horizontal axis shows the logarithm to the base 10 of the earthquake energy, while the vertical axis shows the fraction of earthquakes that were at least as large as the magnitude of the earthquake. The slope of the curve is the Pareto slope and in this case is estimated to be $\beta = 1$ at the upper tail.

Source: Estimates by the author based on historical data on the largest recorded earthquakes from United States Geological Survey at http://earthquake.usgs.gov/earthquakes/world/historical_mag_big.php

with a frequency of no more than $1/2\%$ per year, but ignore the larger risks as too remote to be worth examining.

Now suppose that an analysis reveals that this policy leads to certain rules concerning the size of seawalls, or backup systems for nuclear power plants. Perhaps the structure would be placed at least 5 m above mean high tide because that was the largest tsunami ever measured in the region. We should ask, in the spirit of the Dismal Theorem, if the rule based on historical experience is a robust one. Or do we run the risk of catastrophic events if we use this rule of thumb?

If the event magnitudes and damages are normally distributed, this approach might be sensible. It would say that we run only a $1/2\%$ annual chance of getting a larger event. But, more important, it says that when rare events occurs, they are likely to be *only slightly larger* than the historical experience.

How much larger are the tail events for different distributions. Let's call the design event a "200-year event," meaning that we calculate that events larger than this would occur only once every 200 years. Then, if the distribution is normal, we can calculate that a tail event (one that is larger than the 200-year event) would on average be about 12% larger than the design event.

The answer is completely different for fat-tailed distributions; here, outliers can be huge. If the distribution is a Pareto with a coefficient of $\beta = 1.5$ (often found for stock prices), then the average outlier is around 200% larger than the 200-year event. For the earthquake distribution with $\beta = 1$, the average outlier is actually infinite, but in finite samples the outlier is typically 1000% or more larger than the 200-year event.

This point was vividly seen in the case of the March 2011 Japanese earthquake/tsunami. There is some uncertainty about its size, but current estimates suggest a magnitude of 9.0. At this size, there have been only four earthquakes in recorded history that exceeded it. The largest prior Japanese earthquake was recorded at 8.5. This implies that the energy of the 2011 earthquake was six times more powerful than the largest Japanese earthquake ever measured. This is the power law with a vengeance.

We can visualize the point about outliers with the example of human heights and tsunamis. Female heights are close to a normal distribution. If we take the largest from random sample of 200 women, on average the tallest women would be 71.7 inches tall (this being a 2.7-sigma event). Suppose that we see an unusually tall women, one from the upper tail of the distribution. On average, for a normal distribution, such a person would be 72.7 inches tall. In a population of 1 million women, the expected value of the tallest would be around 79 inches. So if we see an unusually tall woman, perhaps $6\frac{1}{2}$ feet tall, we would be surprised, but this would not change our view of homo sapiens. By contrast, the March 2011 Japanese tsunami was like a 30-foot person striding down the street.⁴

With this information in hand, let's return to our question of designing for rare events. One approach would be to repeat the analysis using longer and longer odds. We might move from the 200-year event to the 400-year event or the 1000-year event. In other words, we would look at larger and larger potential earthquakes (all diminishing in probability according to the power law). The question is whether our optimal design is robust to the changing odds and likely damages. We might find that little changes as we go further out the tail. The additional risk-corrected damages are small compared to the additional costs. This might be our results in analyzing the impact of earthquakes on clam shacks.

But perhaps, as we move further out the tail, we might determine that we need to change the design in a major way, say in the case of a nuclear

⁴ All the calculations about outliers were calculated using replicated samples of 1 million observations using EViews 7.0.

power plant or a national treasure. As we move to the 200-year event ($1/2\%$ annual probability), perhaps we change our seawalls or containment vessels. Perhaps we move our plant to a completely different location. If further out the tail, perhaps we decide we should not build it at all. Under the logic of the Dismal Theorem, for this case, there is in reality no optimal policy because the policy continues to change as we move further out the tail.

So the earthquake/tsunami example shows the logic of the Dismal Theorem. In certain conditions, the combination of risk aversion and fat tails leads to a never-ending chain of changing optimal decisions. As we increase the point in the tail where we cut off our calculations, the best policy continues to change. There is no optimal policy.

1.3. Analytics of Strong Tail Dominance and the Dismal Theorem

An early example of the implications of strong tail dominance was in John Geweke (2001). Geweke was concerned about the use of constant relative risk aversion (CRRA) utility in the context of Bayesian learning in economic-growth models. Recall that a CRRA utility function is of the form $U(c) = c^{1-\alpha}/(1-\alpha)$, for $\alpha \neq 1$, where c is a measure of consumption and α is the elasticity of the marginal utility of consumption [$U(c) = \ln(c)$, when $\alpha = 1$]. Weitzman usually takes the elasticity to be $\alpha > 1$, and I will follow that assumption in this discussion. A central assumption in both Geweke's and Weitzman's analysis is that consumption has a structural uncertainty that is lognormally distributed,

$$\ln(c) \propto \bar{c} + \varepsilon, \quad (1)$$

where ε is $N(\mu, \sigma^2)$, with mean μ and standard deviation σ .

Geweke provided a number of examples in which expected utility would exist (be finite) or would be unbounded depending upon the value of α and the probability distribution of consumption. For example, if consumption is lognormally distributed with known mean and variance, then expected utility exists (is finite) for all α . A degenerate case comes when consumption is lognormally distributed with unknown mean and variance, and when the parameters of the distribution are derived from Bayesian updating. In this case, the distribution of the parameters is a normal-gamma distribution and expected utility is unbounded (negative infinity) for $\alpha \neq 1$. This example is of particular interest because the sampling distribution for the standard deviation of a normal distribution is a t-distribution, which is in the gamma family. The existence of expected utility is "fragile" with respect to changes in the distributions of random variables or changes in prior information. Fragile in this context denotes that, for constant relative risk aversion (CRRA) utility functions, the expected utility exists with some distributions but not for others.

1.4. An Alternative Version of the Dismal Theorem

We can get similar results to the Geweke–Weitzman analysis by using an alternative and simplified version of the Dismal Theorem. We retain the basic analysis, remove the Bayesian structure, and change the distribution assumption for consumption. The parameter conditions for divergence are slightly different between Weitzman and the present approach, but the general insights are the same.

For the utility function, we retain the CRRA utility function, assuming always that $\alpha > 1$. A high value of α signifies high risk-aversion or inequality-aversion. For the probability distribution, we work with the Pareto distribution (i.e., assume the distribution is Paretian in the lower tail). We assume that $1/c$ has a Pareto distribution; i.e., we look at tail events for declines in consumption. For small c , this implies that $f(c) \propto c^{\beta+1}$, $\beta > 0$. Note in this context that a low value of β signifies a distribution with a fatter tail.

Define the *conditional utility* at consumption level c (which denotes the probability times utility) as $V(c) = f(c) U(c)$. For this specification, $V(c) = f(c) U(c) \propto -c^{\beta+1} c^{1-\alpha} = -c^{\beta+2-\alpha}$. Expected utility [the integral of $V(c)$] exists if the conditional utility is bounded as c tends to zero. The expected utility over the interval between zero and some positive level of consumption, \bar{c} , converges to a finite number as $c \rightarrow 0$ if and only if $\beta + 3 - \alpha > 0$. (An alternative approach would be to consider approach to some catastrophic minimum consumption level, but that raises no new issues.)

Weitzman works with *conditional marginal utility*. The reason is that he is interested in the “pricing kernel” or stochastic discount factor, which indicates how much consumption society would give up today to obtain a sure unit of consumption in the future. The conditional marginal utility is defined as $CMU(c) = f(c) U'(c)$. Expected marginal utility [the integral of $CMU(c)$] is bounded as $c \rightarrow 0$ if and only if $\beta + 2 - \alpha > 0$.

As a first example, assume that the distribution is modestly thick (with $\beta = 2$) and risk aversion is moderate, with $\alpha = 2$. In this case, the conditional utility is $V(c) \propto -c^{2+2-2} = -c^2$. A minimal amount of calculation will show that this combination of parameters leads to bounded expected utility and bounded expected marginal utility. On the other hand, assume that the tails are very fat, as with $\beta = 1$, and that there is a strong risk aversion, with $\alpha = 4.5$. In this case, the conditional utility is $V(c) \propto -c^{1+2-4.5} = -c^{-1.5}$. For this case, both expected utility and expected marginal utility are unbounded.

The intuition behind these results is straightforward: The Dismal Theorem holds if the distribution is not only fat tailed but *very fat-tailed* (meaning that β is small), or if the utility function shows not only risk aversion but *very high risk aversion* (meaning that α is large).

While this example simplifies the logic of the argument behind the Dismal Theorem, it shows some important points. It shows that fat tails *per se* are not sufficient to lead to unbounded expected utility or expected marginal

utility. Moreover, the question of boundedness depends upon both the parameters of the utility function as well as the parameters of the probability distribution function.

1.5. The Role of Policy in Tail Events

Implicit in the Dismal Theorem is that the results of policies may have unbounded utility or expected utility. Recall Weitzman's argument that cost-benefit analyses "coming out of a thin-tail-based model remains under a very dark cloud" until the tail issues are resolved. In fact, there is no CBA in the analysis underlying the Dismal Theorem. Indeed, there are no policies. This is clearly an important extension because the purpose of CBA and decision analysis is to steer away from the disasters that lurk in the tail events. How might we extend the analysis to the questions of policy?

1.5.1. Policy as a binary variable

One of the major concerns among scientists is that there are potentially disastrous effects of continuing business as usual in the face of global warming (see, e.g., Lenton *et al.* 2008). Suppose we have discovered some disasters. We need to introduce policy.

Assume that policy (say concerning climate change in this example) is represented by a policy variable Z . An ineffective policy is business as usual, so set the policy variable at zero ($Z = 0$). A very effective policy will prevent climate change, so set the policy variable at one ($Z = 1$). Using this convention, we can rewrite Weitzman's model as a variant of Equation (1) by adding the policy variable:

$$\ln(c) \propto \bar{c}(Z) + \mu \times (1 - Z). \quad (2)$$

This equation contains a deterministic term, $\bar{c}(Z)$, and a multiplicative random term, $\mu \times (1 - Z)$. As in Weitzman's climate-change analysis, μ is the critical uncertain parameter, which is a generalized temperature sensitivity coefficient (TSC). The uncertainty is captured in μ , so the ε term is not needed. Equation (2) states that the uncertainty is given by the distribution of μ , but the uncertainty can be removed by a very effective policy. If policy is very effective, then $Z = 1$ and $\mu \times (1 - Z) = 0$. If policy is ineffective, then $Z \neq 1$ and $\mu \times (1 - Z) \neq 0$. If the underlying distribution of μ is normal, the estimated policy multiplier (call it $\hat{\mu}$) has a t-distribution, which is fat-tailed in Weitzman's framework. This implies that the expected utility for the CRRA utility function is unbounded (negative infinity).

1.5.2. Policy as a continuous variable

In most cases, and definitely for climate-change, policy is a continuous and even multi-dimensional variable. This genuinely complicates the analysis. We

can write the expected value of policy as follows:

$$V(Z) = \int_{\mu} f [T(\mu; Z)] U [c(Z, T), T(\mu; Z)] d\mu. \quad (3)$$

In this formulation, f is the distribution of climatic outcomes given policy and the uncertain parameter, while U is utility, which is a function of consumption and climatic outcomes. The cost-benefit optimum comes where (3) is maximized with respect to the policy Z . The maximization comes when $V'(Z^*) = 0$, and the optimal policy is Z^* .

The important point is that there is no particular relationship between tail dominance for the expected utility in $V(Z)$ and tail dominance in the optimal policy equation, $V'(Z^*) = 0$. It might well be the case that the policy equation has weak tail dominance even though the expected value has strong tail dominance.

For example, suppose that asteroid collisions display strong tail dominance but no policy exists that can prevent the very worst outcomes. Here utility is tail dominant but policy is tail irrelevant. If policy enters into (3) in a separable fashion, then the uncertainties about the TSC would cancel out in the CBA.

An interesting application of policy would be geoengineering by solar-radiation management to prevent potentially catastrophic climate change. Suppose that it turns out that we have a very bad draw for the TSC, and this triggers several major further feedbacks with potentially catastrophic results. This might lead to the case where the Dismal Theorem would apply without policy. However, suppose that we consider a policy of injecting of sulfate particles into the stratosphere to reflect or absorb solar radiation. This might trim the tail of the distribution before the catastrophic events occurred. There might be some unpleasant side effects, similar to cancer therapy that cures the patient but leaves damage in its wake. But this would be a case of strong tail dominance without policy converted into weak-tail dominance with policy.

The major conclusion here is that extending the Dismal Theorem to consider the application of policy requires a different structure and can have different results regarding the importance of the tails.

2. Some Key features of Tail Dominance

2.1. Key Features of Strong Tail Dominance

The Dismal Theorem of strong tail dominance depends upon some special assumptions. Using the simplified analysis presented above, first, it is necessary that the value of the utility function tends to minus infinity (or to plus infinity for marginal utility) as consumption tends to zero. This first condition holds for all CRRA utility functions with $\alpha > 1$, but not for all utility functions with risk aversion. Second, it is necessary that the (posterior)

probability distribution of consumption has fat tails. The fat tails for the distribution of consumption means that the probability associated with low values of consumption declines less rapidly than the marginal utility of consumption increases. We discuss these questions in turn.

2.2. Utility with Near-Zero Consumption

We first discuss some problems that arise with CRRA for near-zero consumption. The CRRA functions that Weitzman analyzes (with $\alpha > 1$) assume that zero consumption has utility of minus negative infinity (and unbounded positive marginal utility) as consumption goes to zero. This has the unattractive and unrealistic feature that societies would pay unlimited amounts to prevent an infinitesimal probability of zero consumption. For example, assume that there is a very, very tiny probability that a killer asteroid might hit Earth, and further assume that we can deflect that asteroid for an expenditure of \$10 trillion. The CRRA utility function implies that we would spend the \$10 trillion *no matter how small was the probability*. Even if the probability were 10^{-10} , 10^{-20} , or even $10^{-1,000,000}$, we would spend a large fraction of world income to avoid these infinitesimally small outcomes (short of going extinct to prevent extinction).

A more reasonable approach would be to assume that near-zero consumption is extremely but not infinitely undesirable. This is analogous in the health literature to assuming that the value of preventing an individual's statistical death is finite. To be realistic, societies do tolerate a tiny probability of zero consumption if preventing zero consumption is ruinously expensive.

2.3. Fat Tails and the Distribution of Parameters

The second crucial condition for the Dismal Theorem is that the probability distribution of consumption has "fat tails" as consumption approaches zero. Recalling Equation (2) above, Weitzman derives this by assuming a very specific functional form for the distribution of consumption. The condition is that the structural distribution of consumption is lognormal, the uncertain policy multiplier is normally distributed, and knowledge about the distribution of the policy multiplier is attained through sampling or Bayesian learning.⁵

However, the results are not robust to minor changes in specifications. For example, a finite upper limit might be placed on the uncertain parameter, perhaps, in Weitzman's example of the TSC, from fundamental physics.

⁵ Weitzman's analysis contains a discussion of a Bayesian approach. He relies on the application of a "non-dogmatic prior distribution" in the form of a generalized power law, $p(\mu) \propto \mu^{-(1+\beta)}$ [using the notation of Equation (2)] with a limiting argument as $\beta \rightarrow \infty$. I believe that the results can be obtained using an improper infinite uniform prior, which provides the same intuition as the classical discussion in the text.

Alternatively, the underlying distribution of the uncertain parameter might, with sampling, lead to a distribution in which the parameter has thin tails. There is little reason to think that the specific distribution used in the analysis is the correct one.

The statistical approach in Equation (2) proceeds in the absence of any prior information. This is not the way that most natural or social scientists derive their subjective probability distributions for important questions such as those regarding climate change, tax policy, or toxicology.⁶ In doing statistical estimates of the radius of the universe, physicists might require that the parameter be non-negative. In the case of the temperature sensitivity, most of current knowledge comes from the application of physical principles, and until recently, *none* of scientists' judgments on the TSC came from sampling of historical data. In general, subjective distributions on scientific parameters are derived from time series, expert opinion, statistical analyses, theory, and similar sources. There would seem little reason to force this complex process into the straightjacket of the model in Equation (2).

3. Empirical Issues in the Application to Climate Change

The Dismal Theorem of strong tail dominance is a useful reminder that analysts should think carefully about the distribution functions of parameters when undertaking an analysis of uncertainties. In particular, the counterintuitive nature of fat-tailed distributions, where “23-sigma” events such as the stock market fall in October 1987 can happen in historical time, needs to be part of any serious analysis of risk. The events in financial markets during 2007–2009 are useful reminders of that important and oft-neglected point. The question addressed here is, how strong is the evidence for strong tail dominance as a general rule and in particular as applied to climate change.

The array of uncertainties and the vast possible distributions from thin to medium to fat make a general discussion of the full set of issues beyond the scope of this paper, and in fact well beyond the frontier of research in empirical studies in climate change. I will focus this discussion on a single uncertain parameter, the TSC, to illustrate how difficult and indeed treacherous the terrain is. But it must be emphasized that the potential areas of uncertainty are much broader than this single example. Several studies such as Hope (2006) and Nordhaus (2007) have looked at a larger array of uncertainties, including those surrounding the carbon cycle, population growth, productivity, energy technologies, and the like. But it must be noted that, when we are facing time horizons of 100 years or more in the future, it is

⁶An excellent example of how different studies can be synthesized is the National Academy of Sciences report on the toxicological effects of methylmercury (NRC, 2000). This report uses both expert assessment and meta-analysis to determine safe exposure levels.

very difficult to get a foothold on the means of the distributions, and estimates of the dispersion are even slippier. It would be fantastic to believe that we have any grip at all on higher moments or Pareto parameters for any of these and even more so for all of them.

What can we do in the face of this vast uncertainty about uncertainty itself? Mainly, just trudge on and try to find our way, knowing that eventually the uncertainties about what the years 2020, 2050, and 2100 hold will be revealed when we get there.

3.1. Estimates of the TSC

The central example in Weitzman's exposition of the Dismal Theorem is the example of the TSC. To begin with, he assumes that the TSC enters in a multiplicative way as shown in Equation (2). For our purpose, we can rewrite Equation (2) as

$$\ln(c_{2200}) = \ln(\bar{c}_{2200}) + TSC \times f(Z). \quad (4)$$

This equation might relate the log of consumption two hundred years in the future (which is the date that Weitzman identifies) to a base value and the product of the *TSC* and $f(Z)$. I interpret Z as a climate-change policy variable in which $f(Z) = 0$ when effective climate change policies are taken (perhaps zero net carbon emissions over the next two centuries), and $f(Z) = 1$ for a business-as-usual case of rapid growth in carbon emissions over the next two centuries. Weitzman does not introduce an explicit policy variable such as Z , but it is implicit in the analysis and discussion of policy and models.

3.1.1. Weitzman's estimates of the TSC

The central empirical component of Weitzman's analysis is that the posterior distribution of *TSC* is extremely dispersed. I quote Weitzman's analysis of this issue at length:⁷

In this paper, I am mostly concerned with the roughly 15% of those TSC_1 "values substantially higher than 4.5 °C" which "cannot be excluded" by the IPCC Fourth Assessment's Summary. A grand total of twenty-two peer-reviewed studies of climate sensitivity published recently in reputable scientific journals and encompassing a wide variety of methodologies (along with 22 imputed PDFs of TSC_1) lie indirectly behind the above-quoted IPCC-AR4 (2007) summary statement. These 22 recent scientific studies cited by IPCC-AR4 are compiled in Table 9.3 and Box 10.2. It might be argued that these 22 studies are of uneven reliability and their complicatedly-related PDFs cannot easily be combined, but for the simplistic purposes

⁷ Weitzman (2009), pp. 5, 7. Note that I have for convenience of exposition changed Weitzman's S_1 and S_2 to TSC_1 and TSC_2 to conform to the notation used here.

of this illustrative example I do not perform any kind of formal Bayesian model-averaging or meta-analysis (or even engage in informal cherry picking). Instead I just naively assume that all 22 studies have equal credibility and for my purposes here their PDFs can be simplistically aggregated. The upper 5% probability level averaged over all 22 climate-sensitivity studies cited in IPCC-AR4 (2007) is 7°C while the median is 6.4°C , which I take as signifying approximately that $P[\text{TSC}_1 > 7^{\circ}\text{C}] \approx 5\%$. Glancing at Table 9.3 and Box 10.2 of IPCC-AR4, it is apparent that the upper tails of these 22 PDFs tend to be sufficiently long and fat that one is allowed from a simplistically-aggregated PDF of these 22 studies the rough approximation $P[\text{TSC}_1 > 10^{\circ}\text{C}] \approx 1\%$.

Instead of TSC_1 , which stands for climate sensitivity narrowly defined, I work throughout the rest of this paper with TSC_2 , which (abusing scientific terminology somewhat here) stands for a more abstract “generalized climate-sensitivity-like scaling parameter” that includes heat-induced feedbacks on the forcing from the above-mentioned releases of naturally-sequestered GHGs, increased respiration of soil microbes, climate-stressed forests, and other weakening of natural carbon sinks. Without further ado I just assume for purposes of this simplistic example that $P[\text{TSC}_2 > 10^{\circ}\text{C}] \approx 5\%$ and $P[\text{TSC}_2 > 20^{\circ}\text{C}] \approx 1\%$, implying that anthropogenic doubling of $\text{CO}_2\text{-e}$ eventually causes $P[\Delta T > 10^{\circ}\text{C}] \approx 5\%$ and $P[\Delta T > 20^{\circ}\text{C}] \approx 1\%$, which I take as my base-case tail estimates in what follows.

Many people would agree that a 5% chance of a 10°C change, or a 1% chance of a 20°C change, would be a dangerous prospect for human societies. However, the procedures used to derive these numbers are flawed.

Weitzman’s estimates are in the spirit of a meta-analysis of existing statistical studies of the TSC.⁸ The problem with his procedure is the following. If we have studies with any statistical independence, then we would *never* take the average of the 95th or the 99th percentile as the appropriate estimate of those percentiles of the underlying distribution. Those numbers might be reasonable estimates of the 95th or the 99th percentile of the next study, but they are not good estimates of the percentiles of the underlying distribution. Even if the underlying “data” were generated in a fashion that would be appropriate for a meta-analysis (which they are not), an appropriate analysis would be to estimate the distribution that generated the “data” and calculate the percentiles from the underlying distribution.⁹ The Weitzman procedure

⁸ There is a large literature on meta-analysis and similar techniques. See Cooper, Hedges, and Valentine (2009) for a review of the field.

⁹ One key to the problem with Weitzman’s approach is the treatment of Gregory *et al.* (2002). That study reports a 95th percentile of ∞ , which is probably because of low power at the upper end. If this were included, then under Weitzman’s procedure, the 95th percentile would also be ∞ .

would be correct only if the TSC estimates are based on exactly the same data, so that the distributions are identical. This is clearly not the case, as an examination of the sources, methods, and the distributions makes clear.

An example will illustrate the difficulty with the procedure. Suppose we want to estimate the 95th percentile of the estimated mean for a random normal variable, Y , for which we have 10,010 independent observations, where $Y \approx N(0, 1)$. It just happens that the observations come in two groups, with group A containing the first 10 observations and group B containing the next 10,000 observations. The expected 95th percentile of the estimated mean for the first group is 0.699, while the expected 95th percentile for the second group is 0.01955. Under the Weitzman procedure, we would average these to get an overall standard deviation of 0.359. The correct answer is to combine the two samples, which yields an expected 95th percentile of the estimated mean of 0.01955. This illustrates why Weitzman's procedure will generally overestimate the 95th percentile of the distribution.

The issue of the appropriate distribution of the TSC is indeed a difficult one. There are many studies of the issue (again, IPCC 2007 discusses several of them). These are a useful source to provide guidance on the central estimates and the range or dispersion of estimates of the TSC. However, without a clear understanding of what is generating the data in the historical periods or the different climate models, it seems dubious that we can at present estimate the shape of the tails of the TSC distribution function.

3.2. The Distribution of Consumption Declines

The Dismal Theorem concerns evaluating situations where consumption approaches "zero," again with the application being to the impacts of future climate change. From an empirical point of view, as with the TSC, it would seem difficult to determine a reliable probability distribution for the decline in consumption. In the case of climate change, e.g., there are severe difficulties in estimating even the central values. Impact analyses are extremely crude and generally apply to high-income countries and to global mean temperature that increases up to $2\frac{1}{2}$ °C, with few studies of impacts above that level.¹⁰ Determining tail behavior is even more difficult because it requires understanding the distribution of extreme outcomes for climate change as well as economic response and would involve projecting these well outside the range of current experience or estimates.

From the point of view of determining the importance of tail events and the applicability of the Dismal Theorem, the key issue is whether the tail distribution of consumption changes is fat or thin. Given the sparse evidence

¹⁰ See the survey by Tol (2009), which covers most of the recent studies. The Fourth Assessment Report on Impacts by the IPCC (IPCC 2007) did not attempt to provide a comprehensive estimate of impacts.

from climate, one useful approach would be to examine the history of extreme consumption changes in economic history. We can think of these as “analog economic catastrophes.” Output shocks come from a wide variety of sources including international and civil wars, trade shocks, revolutions, regime changes, droughts, and similar events. While climate shocks are different from other extreme shocks, they raise similar issues of social response, adaptability to extreme stress, availability of international trade and aid, and resort to conflict under conditions of economic stress. These are also useful analog because it is difficult to see how climate-induced economic impacts could compete with World War II, genocides, and prolonged civil strife in its economic impacts.

I therefore examined historical consumption changes to provide some idea of the distribution of economic responses to a variety of shocks. For this purpose, I examined two data sources to determine both the frequency of extreme output shocks and the shape of the tail of the distribution. The first source of data is the study by Emi Nakamura, Jon Steinsson, Robert Barro, and Jose Ursua (NSBU) on the frequency of extreme events. This study looks at the prevalence of extreme consumption shocks of 24 relatively developed countries over the last century.¹¹ Their definition of disaster is relatively complicated, but it involves consumption declines averaging 30% from peak to trough. They find a probability of entering a disaster is 1.7% per year and further find that disasters last on average for 6.5 years.

The NSBU sample is unrepresentative of all countries for the climate-change analog, so I also examined the disaster history of all countries. For this purpose, I collected data for 188 countries for the period 1950–2007 from the Penn World Table 6.3. I then examined declines in per capita output over all 10-year periods in the sample period (although different lengths of time produced similar results). The idea is that a major shock from war, depression, drought, or other calamities would typically last in the order of a decade. This exercise produced 6543 (overlapping) observations.

We begin by defining an economic disaster as one that produces an output decline of at least 15%. This produces average output declines of 30%, which is the same as NSBU. Under this definition, the number of disaster episodes is 11% of all periods for the sample. That is to say, 11% of years of all countries were ones that belonged to periods in which cumulative output declines were at least 15%. Moreover, major economic declines of more than 50% occurred in 1.1% of country-years. The former number is almost 10 times higher than those in NSBU. The NSBU sample may be biased to produce less extreme results exactly because their sample is primarily successful in countries with long statistical histories.

Table 1 shows the largest output disasters in the sample along with an attribution to the major cause of the disaster. Wars and commodity price

¹¹ Nakamura, Steinsson, Barro, and Ursua (2010). Another model using this shorter data set and examining the distribution is Barro and Tao (2009).

Table 1: Top economic disasters among all countries, 1950–2007

Country	Output decline	Peak year	Trough year	Source
Liberia	89.4	1984	1994	Wars
Dem. Rep. Congo	72.3	1990	2000	Wars
Iraq	70.6	1981	1991	Wars
Afghanistan	70.0	1984	1994	Wars
Kuwait	67.0	1972	1982	Oil market
Libya	66.2	1980	1990	Oil market
Sierra Leone	60.7	1989	1999	Wars
Lebanon	56.0	1987	1997	War
Equatorial Guinea	55.3	1977	1987	Wars
Iran	54.6	1976	1986	Revolution
Saudi Arabia	54.4	1977	1987	Oil market
Brunei	51.5	1979	1989	Oil market
United Arab Emirates	50.2	1977	1987	Oil market
Guyana	49.0	1976	1986	Economic plagues
Cambodia	48.8	1970	1980	Wars
Qatar	47.3	1976	1986	Oil market
Nicaragua	47.3	1983	1993	Economic plagues
Zambia	46.4	1974	1984	Wars
Montenegro	46.2	1990	2000	Wars
Bahrain	43.6	1977	1987	Oil market

This shows the top 20 economic disasters in the post–World War II period. An economic disaster is defined as a period in which per capita output fell by more than 15 logarithmic percentage points in a 10-year period. Note that the output-decline figures are arithmetic declines in real GDP per capita. Data are from Penn World Tables, version 6.3. Sources are attributed by the author.

turmoil are heavily represented in the table of top disasters. Table 2 shows 25 of the largest natural disasters of the same period and shows how they fit into the rank of economic disasters. Only one of the natural disasters was also a top economic disaster (Iran), but that is likely to be a result of revolution rather than environmental disaster.

We also examined the distribution of output declines to measure the tail properties of the distribution. Figure 2 shows the tail of the distribution for the observations with output declines more than 15 logarithmic percent. The Pareto estimates from this sample have an estimated exponent of $\beta = 3.7$ to 4.5 depending upon the lower threshold of the consumption decline. For the entire disaster sample, the Pareto parameter is $\beta = 4.22 (\pm 0.16)$.¹²

¹² Estimates of the Pareto parameter and its standard error are based on the Hill MLE estimator (Hill 1975). The sample sizes are 50 and 30 for the two sample periods reported below. Note that the estimate for the NSBU sample is close to that reported in Barro and Jin (2009), but the number for the larger group of countries has a much heavier tail.

Table 2: Top natural disasters of 1950–2007 and their economic consequence

Country	Year	Disaster	Killed	10-year growth rate	Rank of all periods	Percentile of economic disasters [low is worst]
China	1959	Flood	2,000,000	0.06	1860	28.4
Ethiopia	1972	Famine	600,000	-0.16	658	10.1
India	1967	Drought	500,000	0.21	3484	53.2
India	1966	Drought	500,000	0.24	3878	59.3
India	1965	Drought	500,000	0.18	3153	48.2
Ethiopia	1984	Drought	300,000	-0.09	945	14.4
Bangladesh	1970	Cycl.Hurr.Typh	300,000	-0.10	867	13.3
China	1976	Earthquake	242,000	0.58	6081	92.9
Ethiopia	1974	Drought	200,000	-0.16	665	10.2
Sudan	1984	Drought	150,000	0.10	2254	34.4
Bangladesh	1991	Cycl.Hurr.Typh	138,866	0.12	2468	37.7
Mozambique	1985	Drought	100,000	-0.08	968	14.8
Ethiopia	1973	Drought	100,000	-0.07	1029	15.7
Peru	1970	Earthquake	66,794	0.38	5176	79.1
Sudan	1974	Drought	62,500	0.06	1866	28.5
Sudan	1973	Drought	62,500	-0.27	400	6.1
Ethiopia	1972	Drought	62,500	-0.08	995	15.2
Iran	1990	Earthquake	36,000	0.35	4961	75.8
Bangladesh	1965	Cycl.Hurr.Typh	36,000	-0.11	861	13.2
China	1954	Flood	30,000	0.19	3185	48.7
Bangladesh	1974	Flood	28,700	0.12	2482	37.9
Guatemala	1976	Earthquake	23,000	-0.05	1125	17.2
Colombia	1985	Volcano	21,800	0.22	3587	54.8
Iran	1978	Earthquake	20,000	-0.48	168	2.6
China	1974	Earthquake	20,000	0.54	5951	91.0

The table shows the top 25 natural disasters ranked by the number of fatalities. The last three columns show the 10-year growth rate following the disaster, the rank out of 6,543 country-periods (with a low number being the worst), and the percentile rank (with low being worst). Only one natural disaster was associated with a top economic disaster, but that one (Iran) was primarily due to revolution and oil-market disturbances. None of the natural disasters led to major economic disasters in this test. (Source: EM-DAT International Disaster Data Base 2011)

For “super-disasters” with output declines of more than 70% (which include the wiggly end of the tail in Figure 2), the Pareto parameter is $\beta = 2.36$ (± 0.75).¹³

¹³ Note that this approach will underestimate the standard errors because we use overlapping samples of years. The observations in little tail to the left are all data for Liberia during its economic disaster.

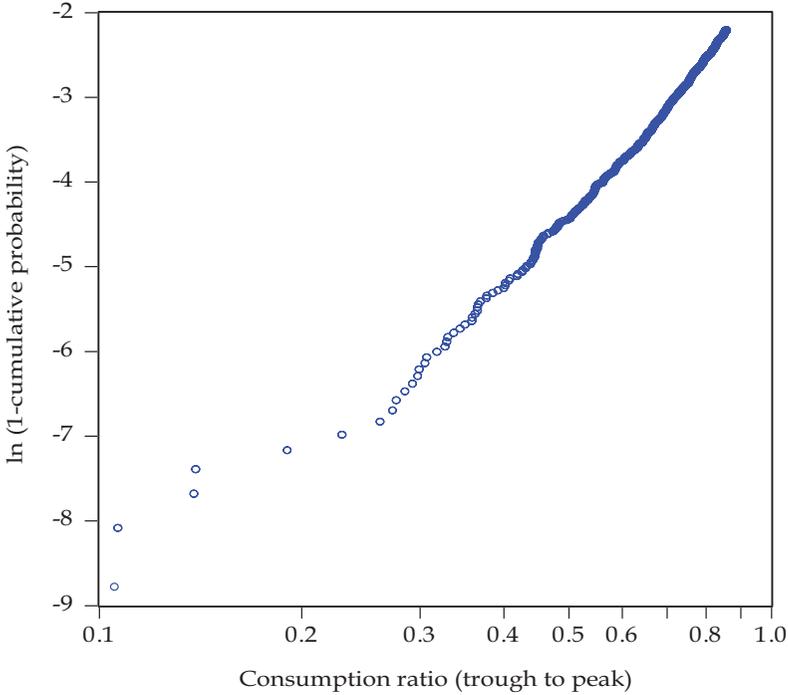


Figure 2: Cumulative distribution of consumption declines.

Note: Consumption ratio is logarithmic scale. The slope is the Pareto parameter described in the text. Sample is only consumption disasters. The Pareto parameter for this sample is $\beta = 4.22 (\pm 0.16)$.

We used two different procedures to determine the robustness of the estimates of the Pareto parameter. Using the PWT sample just described, we examined a set of disasters that look at the largest declines for countries for all periods between 5 and 25 years. The Pareto parameter for this sample looking at the largest 30 disasters is $\beta = 2.31 (\pm 0.43)$. We also looked at the largest 50 disasters in the NSBU data set. These produced a Pareto parameter of $\beta = 3.87 (\pm 0.55)$.¹⁴

This excursion into the incidence of economic disasters suggests that they are actually relatively frequent in recent history, particularly in poor countries. About one-tenth of all years are spent in disastrous circumstances. The distribution of these events does not suggest that the tail is sufficiently fat to satisfy the Dismal Theorem, however. We suggested above that the condition for the expected marginal utility to be bounded is $\beta + 2 > \alpha$. With a

¹⁴ This estimate produces a smaller estimate (fatter tail) than that in Barro and Jin (2009), possibly because it examines only the largest disasters.

Pareto parameter of $\beta = 3$, unbounded expected marginal utility would require a coefficient of relative risk aversion of at least 5. This sets a high bar for a finding of strong tail dominance using the framework outlined above. With a Pareto parameter of 4, societies would need to have a very high risk premium for strong tail dominance to apply.

4. Not So Dismal Conclusions

We have examined the conditions under which tail behavior is likely to dominate economic outcomes or policies. Tail dominance occurs when outcomes or policies are not robust to including an increasingly large domain of uncertain outcomes. In the extreme case contemplated by Weitzman's Dismal Theorem, tail dominance is so strong that no outcome or policy exists in the sense that the expected value of output or policy does not converge as the domain of uncertainty is increased.

I have examined an alternative version of the conditions for strong tail dominance in which preferences display constant relative risk aversion and consumption follows a Pareto distribution. I showed that in this example strong tail dominance holds when the rate of relative risk aversion is greater than three or two plus the Pareto parameter depending upon whether we are examining expected utility or expected marginal utility. These conditions also require that consumption must approach zero and that the marginal utility at zero consumption is unbounded.

The discussion deals with issues of risk preferences and the distribution of economic disasters. One clear point is that strong tail dominance as described by the Dismal Theorem or the variant examined above requires very strong assumptions. The distribution of economic catastrophes over the last six decades indicates that there are indeed severe and frequent output declines, but the tail of the declines is not sufficiently fat to trigger strong tail dominance.

Even though the loaded gun of strong tail dominance has not been discovered to date, the results of the Dismal Theorem are sufficiently powerful to serve as a reminder that we must constantly be alert to this possibility. Perhaps climate change is not a Dismal event for the globe, or even for countries. But there are a sufficiently large number of other possibilities from exotic events such as asteroids and robotic enslavement to more mundane events such as tsunamis, nuclear meltdowns, and financial collapses to motivate careful attention to tail events.

References

- ABE, K. (1981) Magnitude of large shallow earthquakes from 1904 to 1980, *Physics of the Earth and Planetary Interiors* **27**, 2–92.
- BARRO, R. J., and T. JIN (2009) On the size distribution of macroeconomic disasters, NBER Working Paper 15247.

- CHRISTENSEN, K., L. DANON, T. SCANLON, and P. BAK (2002) Unified scaling law for earthquakes, *Proceedings of the National Academy of Sciences* **99**, 2509–2513.
- COOPER, H. M., L. V. HEDGES, and J. C. VALENTINE (2009) *The Handbook of Research Synthesis and Meta-Analysis*, 2nd ed. New York: Russell Sage Foundation.
- EM-DAT International Disaster Data Base (2011) available from: <http://www.emdat.be/>.
- FELLER, W. (1950) *An Introduction to Probability Theory and Its Applications*. New York: Wiley.
- GEWEKE, J. (2001) A note on some limitations of CRRA utility, *Economic Letters* **71**, 341–345.
- GREGORY, J. M., R. J. STOUFFER, S. C. B. RAPER, P. A. STOTT, and N. A. RAYNER (2002) An observationally based estimate of the climate sensitivity, *Journal of Climate* **15**, 3117–3121.
- GUTENBERG, B., and C. F. RICHTER (1954) *Seismicity of the Earth and Associated Phenomena*, 2nd ed. Princeton, N.J.: Princeton University Press.
- HILL, B. (1975) A simple general approach to inference about the tail of a distribution, *The Annals of Statistics* **3**, 1163–1173.
- HOPE, C. (2006) The marginal impact of CO₂ from PAGE 2002: An integrated assessment model incorporating the IPCC's five reasons for concern, *Integrated Assessment* **6**, 19–56.
- Intergovernmental Panel on Climate Change (2007) *Climate Change 2007: The Physical Science Basis*. Contribution of Working Group I to the Fourth Assessment Report of the IPCC, available at: <http://ipcc-wg1.ucar.edu/wg1/wg1-report.html>
- LENTON, T. M., H. HELD, E. KRIEGLER, J. W. HALL, W. LUCHT, S. RAHMSTORE, and H. J. SCHELLNHUBER (2008) Tipping elements in the Earth's climate system, *Nature* **105**, 1786–1793.
- LÉVY, P. (1923) Sur les lois stables en calcul des probabilités, *Compte Rendu Hebdomadaire des Séances de l'Académie des Sciences Paris* **176**, 1284–1286.
- LÉVY, P. (1924) Théorie des erreurs. La loi de Gauss et les lois exceptionnelles, *Bulletin de la Société Mathématique de France* **52**, 49–85.
- MANDELBROT, B. (1963) The variation of certain speculative prices, *The Journal of Business* **36**, 394–419.
- NAKAMURA, E., J. STEINSSON, R. BARRO, and J. URSUA (2010) Crises and recoveries in an empirical model of consumption disasters, NBER Working Paper 15920.
- National Research Council (2000) *Toxicological Effects of Methylmercury*. Committee on the Toxicological Effects of Methylmercury, Board on Environmental Studies and Toxicology, Commission on Life Sciences. Washington, DC: National Academy Press.
- NORDHAUS, W. (2007) *A Question of Balance. Weighing the Options on Global Warming Policies*. New Haven, CT: Yale University Press.
- Penn World Table 6.3 (2010) available from: <http://pwt.econ.upenn.edu/>.
- PISARENKO, V. F., and D. SORNETTE (2003) Characterization of the frequency of extreme earthquake events by the generalized Pareto distribution, *Pure Applied Geophysics* **160**, 2343–2364.

- SAMUELSON, P. A. (1967) Efficient portfolio selection for Pareto-Lévy investments, *Journal of Financial and Quantitative Analysis* **2**, 107–122.
- SCHUSTER, E. F. (1984) Classification of probability laws by tail behavior, *Journal of the American Statistical Association* **79**, 936–939.
- TOL, R. (2009) The economic effects of climate change, *Journal of Economic Perspectives* **23**(2), 29–51.
- United States Geological Survey (2011) *Historic World Earthquakes*. Available from: http://earthquake.usgs.gov/earthquakes/world/historical_mag_big.php
- WEITZMAN, M. L. (2009) On modeling and interpreting the economics of catastrophic climate change, *The Review of Economics and Statistics* **91**, 1–19.

