Symposium:
Fat Tails and the Economics of Climate Change

The Economics of Tail Events with an Application to Climate Change

William D. Nordhaus*

The Economic Importance of Tail Events

From time to time, something occurs that is outside the range of what is normally expected. For reasons that will soon become clear, I call this a tail event. Some tail events are unremarkable, such as an e-mail about a large inheritance that awaits you in Nigeria. Others may change the course of history. Momentous tail events include the detonation of the first atomic weapon over Hiroshima in 1945, the sharp rise in oil prices in 1973, the 23 percent fall in stock prices in October 19, 1987, the destruction of the World Trade Center towers in 2001, and the meltdown of the world financial system in 2007–2008. 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.

Tail events are more than statistical curiosities. In some cases, they may be so important that they dominate the way we think about our options and our strategies. Obviously, tail events dominate thinking about nuclear weapons. Less obvious is how to deal with tail events in economics. One example of how tail risk has changed economic policy is in the area of finance. In response to the meltdown of the banking system in 2007–2008, the theoretical approach to bank regulation has moved toward containing "systemic risk" rather than individual bank risk.

Is there a general theory of economic policy concerning tail events? In an important paper, Weitzman (2009) has proposed what he calls a dismal theorem. He summarizes the theorem as follows: "[T]he catastrophe-insurance aspect of such a fat-tailed unlimited-exposure situation, which can never be fully learned away, can dominate the social-discounting aspect, the pure-risk aspect, and the consumption-smoothing aspect."[1] The general idea is that under limited conditions concerning the structure of uncertainty and societal preferences, the expected loss from certain risks such as climate change is infinite and that standard economic analysis cannot be applied.

*Department of Economics, Yale University, New Haven, CT 06520; e-mail: william.nordhaus@yale.edu.

I am grateful for comments on these issues from William Brainard, Gary Yohe, Richard Tol, and Martin Weitzman, as well as the editor and two anonymous referees.

[1]The quotations here and throughout this article are from Weitzman (2009).

Review of Environmental Economics and Policy, volume 5, issue 2, summer 2011, pp. 240–257
doi:10.1093/reep/rer004

Weitzman discusses many implications of the dismal theorem, but one of the most striking is its destructive effect on benefit–cost analysis, particularly for climate change. "The burden of proof in climate-change CBA [benefit–cost analysis] 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 does not play a significant role in climate-change CBA when it is combined with a specification that assigns high disutility to high temperatures."

These points potentially have significant implications for both economic analyses of climate change and policy in general. However, because the assumptions underlying the dismal theorem are very restrictive, it is important to examine carefully whether standard tools of economic analysis, such as benefit–cost analysis and expected utility theory, can be usefully employed in areas exhibiting great uncertainty, such as climate change. The purpose of this article, which is part of a symposium on Fat Tails and the Economics of Climate Change, is to put the dismal theorem in context and to analyze its applicability with respect to climate change.[2]

The article is organized as follows. I begin with a discussion of the statistical phenomenon known as fat tails, which occurs when there are occasionally extremely large deviations from the normal range of variations in a variable such as stock price changes or earthquake size. If people are accustomed to a normal level of background variability, they may be very surprised, and sometimes badly hurt, by these tail events. Next, I explain some key statistical concepts that provide a foundation for examining the dismal theorem. I then explore the implications of fat tails for the evaluation of economic outcomes, focusing in particular on Weitzman's proposed dismal theorem, which argues that standard benefit–cost analysis cannot be performed when the distribution of outcomes has fat tails and our preferences show strong aversion to risk.

This is followed by a detailed examination of the scope and applicability of the dismal theorem. I begin by putting Weitzman's analysis in the context of the earlier literature on catastrophic environmental outcomes. Next, I discuss an example used by Weitzman in his examination of the economics of climate change: the uncertainty, the potential fat tails, and the catastrophic declines in consumption related to the temperature response to increased accumulations of greenhouse gases. I suggest that one of the issues is unbounded disutility (i.e., utility going to minus infinity) in the tails, and I consider whether the assumption of unbounded utility is consistent with attitudes about other potentially catastrophic events.

I conclude that tail events are indeed important phenomena that require careful analysis and attention. At the same time, I find that there is no universal rule that can be applied to determine when benefit–cost analysis should or should not be applied. Rather, the applicability of standard economic tools, such as benefit–cost and expected utility analysis, will depend upon the uncertainty surrounding specific issues and phenomena, as well as attitudes toward risk.

[2]The other articles in this symposium are Weitzman (2011) and Pindyck (2011).

The Problem of Fat Tails

The analysis underlying the dismal theorem relies on the idea that low-probability, high-consequence events can dominate the impacts and societal concerns for many issues, of which climate change is a signal example. This is the phenomenon known as "fat tails."

To illustrate the problem of fat tails, it is helpful to first picture a probability distribution such as the common bell curve or normal distribution. The normal distribution has most observations clustering around the center, with few showing highly divergent results. Take the height of American women as an example. This variable has close to a normal distribution, with a mean of sixty-four inches and a standard deviation of three inches. Based on the properties of a normal distribution, 95 percent of American women will be between 58 and 70 inches tall. How likely is it that you will observe an eleven-foot-tall woman? This is about twenty-three standard deviations from the mean, which is an exceedingly small number for a normal distribution (about 10^{-230}). Indeed, the world's tallest woman is reported to be about eight feet tall.

Other probability distributions have the property that from time to time a very unlikely looking event occurs. When this is the case, we say that we have witnessed a "tail event" and that the tail of the distribution is "fat" rather than medium as in the case of the bell curve.

Multi-Sigma Events

People sometimes refer to "four-sigma" or "six-sigma" events. These are shorthand terms for how many standard deviations from the average something is. Returning to women's height, if you see a woman who is six feet tall, that is a three-sigma event. In a normal distribution, a three-sigma positive shock (or an observation three standard deviations above the mean) will occur about once every two hundred observations. So, this suggests that only one in two hundred women will be taller than six feet.

Quite a different case is illustrated by daily changes in stock prices. Prices on U.S. stock markets fell approximately 23 percent on October 19, 1987. An estimate of the daily standard deviation of price change over the 1950–1986 period shows a standard deviation of 1 percent. If stock price changes follow a normal distribution, then we would see a 5 percent change in prices once every 14,000 years and a 7.2-sigma change about once in the life of the universe. However, twenty-three-sigma events, like eleven-foot people, simply do not occur for a normal distribution.

Yet, these large deviations occur much more frequently than would be predicted by the normal distribution. Let's look at the long-term history of price changes in the U.S. stock market. I have calculated the monthly returns for stock prices for the 140-year period from 1871 to 2010. Table 1 illustrates the phenomenon of fat tails for the stock market, whereby the actual maximum and minimum increases over this 140-year period are much larger than would be predicted by a normal distribution. In fact, the maximum is a "ten-sigma event," which would almost never happen with a normal distribution (probability less than once in the life of the universe). So, people who think that financial markets follow the bell curve will, from time to time, be very surprised.

It has been known for many years that there are large deviations from the normal or bell curve distribution for the stock market as well as for many other phenomena. Statisticians

Table 1 Comparison of normal and actual distribution of stock price changes, 1871–2010

Largest increase in 140 years	
Actual	**40.7**
Normal distribution	**14.3**
Largest decrease in 140 years	
Actual	**-30.8**
Normal distribution	**-13.7**

Notes: I looked at 140 years of stock price changes measured in percent per month. I then took a normal distribution with the same mean and standard deviation as the actual data. Increases are logarithmic.
Source: Calculations based on data from Shiller (http://www.econ.yale.edu/~shiller/) and DRI database (derived from CD-ROM from Yale Library).

have developed both probability distributions that have fat tails and techniques to estimate these other distributions. A particularly interesting probability distribution that may have fat tails is one that is known as the "power law."[3] This refers to a distribution in which the probability is proportional to a value to a power or an exponent. One example of this is the power law for earthquakes, which I will return to later. Other examples include the sizes of cities and firms, solar flares, moon craters, wars, incomes, wealth, and commodity prices.

The Element of Surprise

One way to think about tail events is to note that you can be extremely surprised by an outcome when a process has fat tails. This "surprise" can be measured by asking how deviant of an observation might turn up when you have many observations. For example, suppose that you have looked at housing prices for fifty years and observe that they have never declined. So, you place your bets on housing prices continuing to rise. But then you get a draw from a fat-tailed distribution and housing prices fall—not just a little, but 20 or 40 percent. This is what happened in the U.S. housing market after 2006.

Now, suppose you were in the oil market in the early 1970s. Suppose further you were a trader who eschewed any economic theories and just looked at historical data (which is not an absurd approach given the unpredictability of oil prices). Oil prices had been pretty stable, and the "sigma" was about 5 percent on a monthly basis. But then, in 1973, we had, by historical standards, a thirty-seven-sigma event (see Figure 1, which shows the high-sigma stock market and oil price surprises events on a monthly basis). No wonder many people thought the economic world as we knew it was coming to an end in 1973. If we could have a thirty-seven-sigma event, then we can rule out almost nothing (except phenomena inconsistent with the laws of nature).

Enter fat-tailed distributions.

There has been much historical and statistical research on "surprises"—or fat tails—over many years. The conclusion of this research—on oil prices, stock prices, earthquake size, war fatalities, and many other phenomena—is that we have been surprised for the wrong reasons. We have often observed much larger deviations from the norm than would be predicted by standard statistical analysis. Take the example of the events depicted in Figure 1. We might have thought that the "spikes" in Figure 1 were near-zero-probability events, and perhaps never even

[3]Power law distributions were introduced into economics by Mandelbrot (1963) and are widely used in the natural and social sciences.

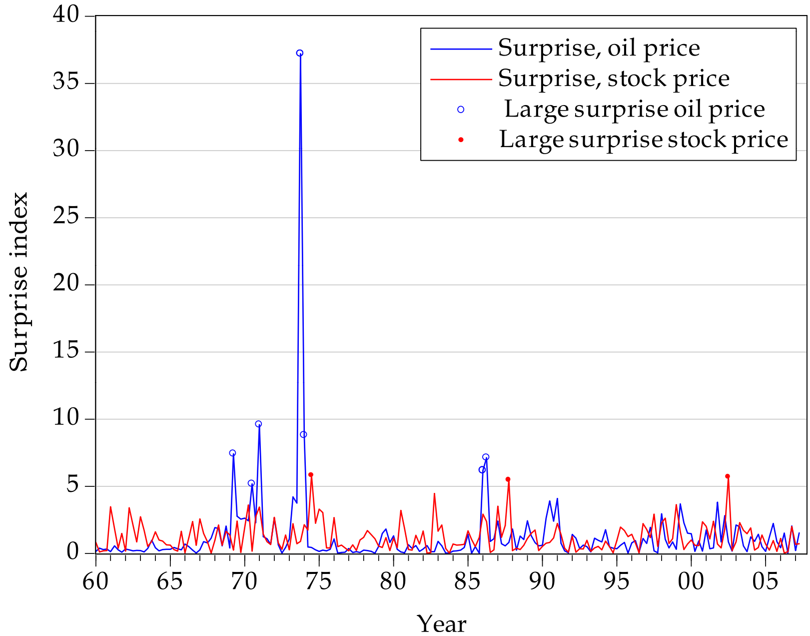

Figure 1 Surprise index for oil prices and stock prices, 1960–2007.
Notes: The "surprise index" is measured as a three-month change in the logarithm of the price divided by a twenty-year moving average volatility, where each is measured monthly. The circles show the periods when the surprise was more than five moving standard deviations, with open circles for oil prices and solid circles for stock prices.
Source: Nordhaus (2007b).

contemplated that such tremendous changes in stock or oil prices could occur. But it turns out that our implicit, standard statistical reasoning was wrong. Thus, we were "surprised" because rather than the distributions being normal, they had fat tails. This means that the probability of the "way-out" events occurring was much greater than predicted by the normal distribution.

How Can We Know Which Distributions Are Fat Tailed?

This research solved one problem but raised another. The new problem is, which are the fat-tailed distributions and which are the thin-tailed ones? When should we be on the lookout for high-sigma events? And, how can we answer these questions before the high-sigma event occurs?

The basic proposition underlying the dismal theorem is that with "fat-tailed" distributions, decision analyses may lead to very unintuitive results. This arises because distributions with fat tails are ones for which the probabilities of rare events decline relatively slowly as the event moves far away from its central tendency. This means that it can be hard to detect fat-tailed distributions and very hard to know how fat the tails are.

I have already mentioned the fat-tailed probability distribution associated with the power law. This is known in statistics as the Pareto distribution, after the Italian economist Vilfredo Pareto (who also introduced the important concept of a Pareto optimum). The important point about this distribution is that the probability of high-sigma events declines slowly relative to distributions like the normal. After a slight statistical detour, I will provide some examples of the power law distribution.

Key Statistical Concepts

I pause here briefly to explain some key concepts for understanding fat tails and the dismal theorem. The dismal theorem depends upon what mathematicians call limiting behavior. In particular, we need to consider what happens to the expected utility of outcomes under catastrophic conditions. The standard economic approach, used here and by Weitzman, assumes that catastrophic outcomes are ones in which "consumption" declines sharply, perhaps to zero or some minimum level. For example, most people think that the Great Depression was a catastrophic economic event; consumption in the United States fell by about 20 percent during this period.

The Utility of Consumption

I first consider the utility of consumption, $U(C)$. Like Weitzman and most everyone else, I assume that people are risk averse. This means that the marginal utility of consumption, $MU(C)$, rises as C declines. Next, I consider the probability that a particular level of consumption will occur as $P(C)$. Viewed in these terms, the dismal theorem is really about what happens as C gets indefinitely small or perhaps approaches some minimum subsistence level. Take the product of the probability of a particular outcome times its marginal utility, which is $P(C) \times MU(C)$. The question is whether P declines more rapidly than MU increases as C goes toward the catastrophic minimum. In Weitzman's dismal theorem, because of the fat tails, P declines slowly for low values of C; and, because of strong risk aversion, MU increases sharply for low values of C. As a result, $P(C) \times MU(C)$ increases sharply as C declines, and the expected marginal utility tends to minus infinity as C goes to zero.

Note that the dismal theorem relies on two conditions as C declines: P must not go to zero and MU must be indefinitely large as C declines. If either of these conditions fails, then the dismal theorem fails, and we are back to standard economics and standard benefit–cost analysis.

Definitions of Fat Tails

There is no generally accepted definition of the term "fat tails," which is also sometimes called "heavy tails." One set of definitions (Schuster 1984) divides distributions into three classes. A thin-tailed distribution has a finite upper limit (such as the uniform distribution), a medium-tailed distribution has exponentially declining tails (such as the normal distribution), and a fat-tailed distribution has power law tails (such as the Pareto distribution).[4]

To further illustrate these distinctions, it is helpful to go a step further and present the mathematics. To make the exposition simpler, I use a specification that is slightly different from Weitzman's. Begin with the Pareto distribution, where the probability of an event is $P = k_1 X^{-(1+a)}$, where a is the Pareto "shape parameter," which reflects the importance of tail events; X, the variable of interest; and k_1, a constant that ensures that the sum of probabilities is 1.[5] If a is very small, then the tail is very fat and the variable has a highly dispersed distribution. As a gets larger, the tail looks more like a normal distribution.

[4]This definition is proposed in Schuster (1984). Weitzman uses a slightly different definition, but the difference is not essential to the discussion here.

[5]Note that when we examine consumption, we look at $X = 1/C$. We do this because we are interested in the lower tail of C (i.e., declines in consumption). However, when we examine something such as earthquakes, high is bad. Also, this equation might hold only in the tail, as in the case of earthquakes.

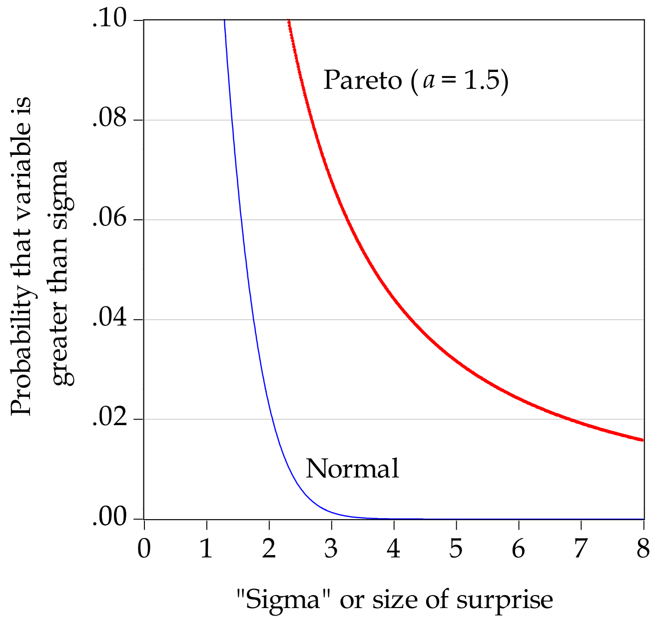

Figure 2 Illustration of tails for a normal distribution and a Pareto distribution with scale parameter $a = 1.5$. *Notes*: Each curve shows the probability that the variable will be greater than the sigma shown on the horizontal axis.

Figure 2 shows an example for a normal distribution and a Pareto distribution with a slope parameter of $a = 1.5$. The curves show the probability that an event will be at least N sigma, where sigma is a measure of dispersion. Note that the probability of a surprise is very small for normal distributions once a four- or five-sigma threshold is reached. For this version of the Pareto distribution, which is found for many economic and physical variables, the four-sigma probability is still substantial.

Risk Aversion

Any analysis of the implications of uncertainty and fat tails must take into account attitudes toward risk, or what is technically known as risk aversion. The notion of risk aversion is employed in many fields of decision sciences. It basically says that we will pay to avoid risk. A useful concept is the rate of relative risk aversion, which can be illustrated using finance. If we are risk averse, then we may hold some low-risk securities even though they have a lower return than a high-risk security. For example, we might hold a portfolio with half in bonds and half in stocks, even though the bonds yield only 2 percent, while the stocks on average yield 6 percent. If we have constant relative risk aversion, we will continue to hold the same share of stocks and bonds as we become richer (or more recently, poorer).

It is common in economic studies to assume a constant rate of relative risk aversion (CRRA), and this assumption underlies Weitzman's work on the dismal theorem. A CRRA utility function is one in which the marginal utility of consumption rises proportionally with the fall in consumption. As it turns out, both the Pareto distribution and the CRRA utility function have the same algebraic structure. The CRRA utility function has the form $U = k_2 C^{(1-b)}$, where the parameter $b > 0$ is the CRRA and k_2 an inessential constant. Thus, $b = 0$ implies risk neutrality; the higher the value of *b*, the higher the degree of risk aversion.

This concludes the statistical interlude.

Implications of the Dismal Theorem

I now consider the implications of Weitzman's dismal theorem. The basic implication is straightforward: in the presence of both fat-tailed uncertain outcomes and strong risk aversion, we cannot rely upon our standard tools of expected utility analysis. The reason is that the probability of extreme and catastrophic events does not decline sufficiently rapidly to compensate for our aversion to encountering these catastrophic events.

More precisely, using the mathematical terms presented in the previous section, the dismal theorem depends upon the values of the fatness of the tail (a) and the rate of relative risk aversion (b). If we assume that the probability distribution is Pareto and the utility function is CRRA, then the dismal theorem holds when $b > a + 1$. This means that our standard tools of economic analysis are in deep trouble either when risk aversion is very high or when the tail is especially fat.

This leads to a *simplified dismal theorem:* when risk aversion (large *b*) is very high or the tails are very fat (small *a*), so that $b > a + 1$, our standard tools of expected utility analysis break down because expected marginal utility is negative infinity.

Weitzman motivates his dismal theorem using an advanced technique from decision sciences known as Bayesian learning. However, the intuition is fairly straightforward. Suppose that some important parameter is unknown. Weitzman uses an example from climate science, the temperature sensitivity coefficient (TSC).[6] The TSC is defined as the equilibrium increase in global mean surface temperature caused by a doubling of atmospheric concentrations of CO_2. He assumes that there is some genuine uncertainty about the TSC perhaps because of geophysical nonlinearities. This means that the "true" TSC is a parameter, call it K, but it has some true imprecision or variability, call it S.

Perhaps the true TSC follows the normal distribution shown in Figure 2. We cannot learn the true K and S, however, because the historical record on temperature and CO_2 increases is too limited, and because we cannot wait for thousands of years for the true climate parameter to be revealed. So, we might need to rely on a small sample of observations, say from a historical period. Because we must rely on a limited sample, we cannot even reliably estimate the degree of imprecision of the important parameter, *K*. Thus, with small samples, we have only a rough idea about the TSC and its imprecision. We can get a better idea with a larger sample, but as long as the number of observations is finite, we can never reduce the estimated variability as low as S.

This point about the difficulty of determining the imprecision of our statistical estimates is generally not terribly important because if the distribution is thin tailed or medium tailed, such as the normal distribution, the imprecision will have limited impact on our decisions. However, if the distribution can extend far to the right (as for large earthquakes) or far to

[6] I defer a detailed discussion of the TSC to a later section and make a more general point here.

the left (as for very low levels of consumption), then the degree of imprecision may be an important factor in our decisions.

Weitzman makes a second point concerning the TSC, but this one is less general and depends upon his exact assumptions. He notes that if the *true* distribution of the TSC is normal, then the *estimated* distribution (known as the t distribution) has a fat-tailed distribution. However, this point depends on the number of observations. If the number of observations is very small, then the distribution is very fat, but as the number of observations gets large, the distribution begins to resemble the medium-tailed normal distribution. Thus, whether the estimated distribution has a fat or a thin tail depends on the number of observations in our sample. This means that the dismal theorem does not hold in all circumstances. However, it is general enough to make us pause to consider its implications before moving on to analyze its applicability.

So, to summarize the findings here concerning when the simplified dismal theorem would apply in practice: it requires some combination of extreme risk aversion and very fat tails. When either of these conditions holds, we can find ourselves in a situation where the tails of the distribution tend to drive our evaluations and policy analyses.

Earthquakes as an Example

Let's take the example of earthquakes. Seismologists have determined that the size of earthquakes tends to follow a power law distribution, known as the Gutenberg–Richter law. Most estimates indicate that the energy (i.e., magnitude) of earthquakes has a Pareto parameter of about $a = 1$. Figure 3 illustrates the power law for earthquakes based on data from the U.S. Geological Survey on the largest recorded earthquakes, but this is a typical result (see Christensen *et al.* 2002 as an example). The variable on the horizontal axis shows the logarithm of the size of the earthquake, where size is measured as energy, while the vertical axis shows the logarithm of the fraction of earthquakes that were at least as large as the number shown on the horizontal axis. This figure is similar to Figure 2 and indicates how fat the tail of the distribution is.

Figure 3 offers a clear example of a fat-tailed distribution. It shows that the probability of a powerful earthquake declines very slowly for more and more powerful earthquakes. To further illustrate, suppose we thought that the energy of earthquakes in Japan was normally distributed. Then, using the past two hundred years of data, we would calculate that an earthquake as large as the March 2011 quake would occur every 10^{-13} years (more or less). But if we were to use the Pareto distribution shown in Figure 3, we would find that an earthquake of that size or larger will occur every one hundred years or so.

It is useful to examine the case of earthquakes in the context of the dismal theorem because this is an area where the economics is quite intuitive. Suppose that the costs—or damages—of an earthquake are proportional to its size (energy). Such a situation could arise because of structural damage, or because of the size and power of a tsunami. In this case, the costs will also have a fat-tailed distribution, as in Figure 3. This means that the surprises that accompany the distribution of the earthquake's size will be accompanied by similar surprises about its damage.

Because of the fat-tailed nature of earthquakes, there is likely to be a big element of surprise when outliers occur, as vividly illustrated by the March 2011 earthquake and tsunami in Japan.

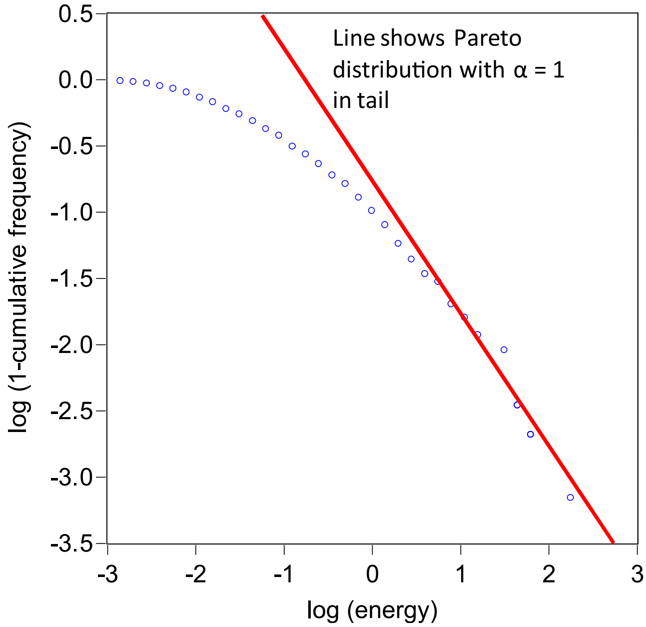

Figure 3 The power law distribution for earthquakes.
Notes: The horizontal axis is the logarithm to the base 10 of the earthquake magnitude (measured as energy). The vertical axis indicates the logarithm of the fraction of earthquakes that were at least as large as the earthquake energy shown on the horizontal axis. The slope of the power law distribution is the Pareto slope. Estimates place the Pareto parameter at $a \approx 1$ at the upper tail.
Source: Calculations based on historical data from the U.S. Geological Survey at http://earthquake.usgs.gov/earthquakes/world/historical_mag_big.php.

The estimated magnitude of this earthquake was 9.0, which means that only four earthquakes in world record history have exceeded it. The largest prior earthquake in Japan had a recorded magnitude of 8.5. This also implies that the energy of the 2011 earthquake was six times as powerful as the largest Japanese earthquake ever measured. This is the power law with a vengeance and explains in part why the earthquake was so devastating and surprising.

We can also apply our surprise theory here. If earthquake and tsunami magnitudes were normally distributed, then an unusually large earthquake would be *only slight larger* than historical experience. However, with fat-tailed distributions, outliers can be much larger than historical experience (as in Figure 1)—in fact, six times larger in the case of the 2011 Japanese earthquake. If you thought earthquakes had a normal distribution, such a tail event would be equivalent to observing a twenty-foot-tall woman striding down the street.

Does the Dismal Theorem Apply as a General Rule?

I turn next to the question of whether the dismal theory applies as a general rule by analyzing the conditions under which tail events dominate the outcomes and may therefore undermine standard policy analysis. Weitzman focuses on the case of climate change, and this is where I will concentrate my analysis. But the analysis would apply to a wide variety of other issues and events that have the same logical structure, such as earthquakes, nuclear accidents, oil spills, bank failures, and asteroid hits.

The applicability of the dismal theorem depends upon some very restrictive assumptions. First, it is necessary that the probability distribution of consumption has a sufficiently fat tail so that tail events become quantitatively important. Second, it is necessary that risk aversion be sufficiently powerful that these tail events have a significant impact on the overall utility. And under both of these assumptions, the distribution must go to the limit of catastrophe and not have a finite bound or upper or lower limit. I begin by putting the dismal theorem in context by comparing Weitzman's analysis with some earlier studies of catastrophic outcomes. I then examine the extent of the dismal theorem's relevance.

Early Studies of Catastrophic Outcomes

The analysis of catastrophic (or at least very serious) environmental outcomes has a long history. One of the early influential approaches was Ciriacy-Wantrup's (1952) theory of a "safe minimum standard," or a level of an ecosystem or pollution beyond which the outcome would be catastrophic. This theory has led to the precautionary principle.

Cropper (1976) presented a rigorous treatment of the issue, dealing specifically with the optimal pollution or resource use in the presence of a catastrophic threshold. Cropper's analysis differs from the Weitzman analysis in two major respects. First, Cropper's study examines the optimal policy and is therefore a true benefit–cost analysis, whereas there are no policies embedded in the Weitzman analysis. Second, Cropper assumes that utility is never infinitely negative; so, in contrast to the Weitzman analysis, the case of infinitely large disutility does not arise. In Cropper's analysis, while there may be catastrophic outcomes, the problem is well defined and we can use standard economic tools to identify the optimal policy.

An important study by Kolstad (1996) analyzed the interaction of uncertainty and learning in the context of global warming. Kolstad's concern was how the optimal policy is affected by the irreversibility of policy as well as the potential for learning about future damages. Using a parameterized model of global warming rather than a theoretical model, Kolstad found that the average damages are more important than the uncertainty about damages; more specifically, in his results, the uncertainty about damages is second order relative to the first-order mean damages. Additionally, he found that the prospect of learning would reduce the optimal control rate relative to the deterministic case. Note that Kolstad's analysis assumes that damages are never catastrophic. Moreover, like Cropper, the utility function in Kolstad's analysis shows only modest risk aversion (with a rate of relative risk aversion of one). The Cropper and Kolstad studies indicate that the presence of uncertainty and very large damages is not sufficient per se to undermine standard economic analysis.

Studies by the present author have also examined the implications of learning and uncertainty in the context of global warming. The most recent comprehensive study (Nordhaus 2007a) was an aggregative model of global warming and the economy, known as the DICE-2007 model. This study examined the implications of uncertainty about eight major variables on the optimal climate change policy and other variables. The major finding of the uncertainty analysis was consistent with Kolstad's findings. I found that the best-guess or certainty-equivalent policy is a good approximation for the policy in which a full expected utility framework is used. There appear to be no empirical grounds in the DICE model for paying a major risk premium for future uncertainties beyond what would be justified by the averages (Nordhaus 2007a). Thus, as in the earlier studies discussed above, the assumptions underlying the DICE model differ significantly from those in the Weitzman analysis. More specifically, they do not include a fat-tailed distribution of outcomes, they include policy, and consumption is not catastrophically affected by climate change.

The Absence of Policy in the Dismal Theorem

One of the major points in the dismal theorem is that doing standard economic benefit–cost analysis is not possible. Weitzman (2009) argues that "the policy relevance of any CBA [benefit-cost analysis] coming out of such a thin-tail-based model might then remain under a very dark cloud until this fat-tail issue is addressed seriously and resolved empirically." In fact, however, there is no benefit–cost analysis in the analysis underlying the dismal theorem, and indeed, there are no policies. How then might we extend the Weitzman (2009) analysis to address questions of policy?

As in the early studies reviewed above, catastrophic environmental outcomes should be considered in the context of environmental policies. For example, should we slow greenhouse gas emissions? Should CO_2 concentrations remain below some threshold level? Should we ensure a minimum stock of bluefin tuna? What is the safe exposure to radioactive wastes?

If we include a policy lever, the analysis becomes quite different from the analysis underlying Weitzman's dismal theorem. Suppose we have the same model as discussed above (i.e., with CRRA utility and a power law distribution), but temperature is affected by a policy variable, which might represent pollution control. We calculate the costs and benefits of different levels of pollution control. The benefit–cost optimum occurs where expected utility is maximized with respect to the policy.

The interesting point here is that there is no particular relationship between the impact of fat tails on expected utility and the impact of fat tails on the optimal policy. For example, it might well occur that the outcome of an asteroid collision has very fat tails; but if we cannot prevent collisions, then policy is tail-irrelevant. In other cases, policy might turn a benign situation into a catastrophic one. For example, suppose some foolish leader decided that the best policy for addressing a relatively benign environmental issue was to threaten to drop a nuclear weapon, which could escalate into global conflagration. The point is that when we introduce policies, the analysis underlying the dismal theorem no longer applies directly.

Parameter Uncertainty and the TSC

One important element in the analysis of tail events is the distribution of the uncertain parameter. Let's return to Weitzman's example of a parameter used in climate change analysis, the temperature-sensitivity coefficient (TSC). The idea is that consumption may be drastically affected by the extent of climate change, and the extent of climate change depends upon how sensitive climate is to changes in CO_2 and other greenhouse gases. That sensitivity is the TSC. For our purposes here, we assume that consumption declines are a function of the TSC. If the TSC is very large, then the consumption decline is also very large. This implies that if we take strong steps to curb CO_2 emissions, then the temperature increase and the consumption declines will be smaller. I will examine the methods used by Weitzman to estimate the distribution of the TSC because these methods illustrate the uncertainties and the difficulties of determining the exact probability distributions of uncertain variables such as the TSC.

Climate models reviewed in the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) found that the central estimate of the TSC was 3.2°C for an equilibrium doubling of CO_2 concentrations (IPCC 2007). Weitzman's estimate of the TSC is different from the IPCC estimate. He relies on a "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 weakenings of natural carbon sinks." This would refer to a long-run TSC, perhaps on the scale of many centuries in which the different feedbacks could take place. Weitzman argues that the precision of the IPCC estimate is extremely uncertain, and so he examines a number of scientific studies of the distribution of the TSC.[7] In describing how he arrives at his estimate of the TSC, Weitzman states, "I just naively assume that all 22 studies have equal credibility and . . . can be simplistically aggregated." His approach is to take the *average* of the 95th percentiles of the different distributions and use that as his estimate. Weitzman concludes that the 95th percentile of the TSC is 10°C and that the 99th percentile is 20°C.

Many people would agree that a 5 percent chance of a 10°C change, or a 1 percent chance of a 20°C change, would be a very dire prospect for human societies. However, there is a serious flaw in the technique Weitzman has used to estimate the TSC. This is an important point not only because it affects the substance of Weitzman's analysis but also because it shows how difficult it is to determine the variability or reliability of estimates.

Weitzman's estimates of the TSC are in the spirit of a meta-analysis of existing statistical studies of the TSC. The problem with this procedure is that if we have studies with any statistical independence, it is not appropriate to use the average of the 95th or the 99th percentile of different studies as the estimate of the percentiles of the underlying distribution. Those numbers might be reasonable estimates of the 95th or the 99th percentile of the next study but are not good estimates of the percentiles of the underlying distribution. The appropriate procedure is to start with the underlying distributions from different studies, then combine them into a meta-distribution, and finally calculate the percentiles from the combined distribution. The approach used by Weitzman will be correct only if the studies are drawn from exactly the same data so that the distributions have a perfect correlation. However, this is not the case, as an examination of the sources, methods, and distributions makes clear.

A numerical example will help to illustrate this point. 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. Assume that Y has a mean of zero and a standard deviation of one. It just happens that the observations come in two groups, with group A containing the first ten observations and group B containing the next ten thousand observations. If we take 10,010 random draws of Y, then 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 0.01956. Following Weitzman's procedure, we would average these to get an overall standard deviation of 0.359. However, the correct approach would be to combine the two samples into a complete sample distribution, which yields an expected 95th percentile of the estimated mean of 0.01955. Thus, Weitzman's procedure will generally provide an estimate of variability that is biased upward.

A final important issue concerning the assumptions implicit in the dismal theorem is that it assumes that the uncertain parameter has no upper bound. However, if there is an upper

[7] I have simplified the analysis by omitting the distinction Weitzman makes between a normal TSC and an augmented TSC. However, this omission does not affect the central argument either in Weitzman (2009) or here.

bound, and if consumption only goes to zero in the limit, then the dismal theorem does not hold. This is because the worst that can happen is that the parameter takes the upper bound, which means that at worst the expected value of the parameter is equal to the utility of consumption at that upper bound. To address and settle this issue concerning the dismal theorem, it would be helpful to establish an upper bound for the TSC. Unfortunately, to date, little work has been undertaken in this area.

Further Thoughts on the Implications of Catastrophic Declines in Consumption

The dismal theorem requires that the marginal utility of consumption tends to infinity for near-zero consumption. This requirement has the unattractive and unrealistic implication that societies would pay unlimited amounts to prevent zero consumption even if its probability is infinitesimal. For example, assume that there is a very, very tiny probability that a killer asteroid might hit Earth and that we can deflect that asteroid for a huge expenditure. The CRRA utility function implies that we would spend virtually all of world income *no matter how small the probability.* That is, even if the probability were 10^{-10}, 10^{-20}, or even $10^{-1,000,000}$, we would still spend most of our income to avoid these infinitesimally low-probability outcomes (short of going extinct to prevent extinction).

An alternative would be to assume that near-zero consumption is extremely but not infinitely undesirable. This is analogous to assuming in the health literature that the value of avoiding an individual's statistical death is finite. To be realistic, societies tolerate a tiny probability of zero consumption if preventing zero consumption is ruinously expensive. I consider some possible bounds below.

A final question is, what exactly do we mean by "zero consumption"? Weitzman defines zero consumption as being "the value of statistical civilization as we know it, or perhaps even the value of statistical life on Earth (as we know it)." Zero consumption is actually an ambiguous concept. Is zero consumption (1) declining average consumption of a fixed number of people, (2) high average consumption of a declining number of people, or (3) high average consumption of thriving civilizations for a statistically declining period? These are very different descriptions of the end of civilization or of our species.

Valuing "Zero Consumption"

How should we think about societal valuation of "zero consumption?" Take the third of the possible approaches to "zero consumption" above (which implies an end to human civilizations as we know them). This is the number that Weitzman takes to be unboundedly negative.

Do people really decide about catastrophic events by putting infinite disutility on them? Clearly not. This issue has been contemplated from time to time. It arose about two decades ago in the context of "nuclear winter," which was the theory that the detonation of a large number of nuclear weapons would lower global temperatures so much as to kill off most if not all of humanity.[8] More recently, there has been a spirited debate about "strangelets" and

[8]One of the most influential early studies was by Turco *et al.* (1983). This study was generally disregarded after further studies. However, recent work has done new modeling and found disturbing results (see Mills *et al.* 2008).

black holes triggered by heavy-ion collisions in large colliders (see Dar, De Rújula, and Heinz 1999; Jaffea *et al.* 2000). Strangelets are hypothetical particles that might conceivably be created in heavy-ion colliders and would turn the earth into a hot lump of strange matter in the blink of an eye. Most knowledgeable scientists would regard these as catastrophic events with positive (if very low) probabilities. However, the low probabilities of these catastrophic outcomes have not induced people to dismantle all but a few nuclear weapons or to stop the experiments in colliders.

Again, an example will help to clarify this issue. A well-established catastrophic risk is killer asteroids. An asteroid such as the one at the K/T (Cretaceous–Tertiary) boundary, which had a diameter of around ten kilometers, would probably be sufficiently large to destroy human civilizations. Such asteroids are estimated to have a probability of Earth collision of about 10^{-8} per year (Chapman and Morrison 1994). If we were to follow the dismal theorem, we should be devoting an unlimited fraction of our resources to reduce that probability by even a small amount. Yet, at present, the U.S. government is spending about $4 million per year to track hazardous asteroids. This program is designed to detect only 90 percent of large potentially hazardous objects within the next decade.[9] This appears to be a pretty relaxed approach to addressing such catastrophic risks if we are indeed highly risk averse. In this case, the revealed social utility function does not seem to place a very high premium on preventing catastrophic outcomes.

To summarize, societies do not appear to behave as if catastrophic outcomes have infinite disutility. Perhaps the dismal theorem is really a warning against mechanically applying a specific utility function to situations where consumption might be very small. This would make it as much a story about extrapolating utility functions as it is about shuddering at the prospect of infinitesimally likely events.

Other Concerns about Catastrophic Events

The dismal theorem holds that we cannot rule out catastrophic impacts of climate change with 100 percent certainty. If we broaden our horizons, we would find that these results apply in a wide variety of circumstances in which we are highly uncertain about the technology or societal impacts of human activities. Areas in which experts have warned about potentially catastrophic outcomes include biotechnology, strangelets, runaway computer systems, nuclear proliferation, rogue weeds and bugs, nanotechnology, emerging tropical diseases, alien invaders, asteroids, and so on. Like global warming, all these outcomes have deep uncertainty in the sense that we really cannot be sure about the shape of the probability distribution. Indeed, these outcomes may have greater uncertainty than global warming because there are fewer well-understood constants in the biological and technological world than in the geophysical world. Thus, if we were to accept the dismal theorem, we would likely drown in a sea of anxiety at the prospect of the infinity of infinitely bad outcomes.

Weitzman dismisses such pervasive anxieties about these other catastrophic outcomes, arguing that they are "extremely unlikely." However, other scientists have come to very different conclusions. One example is Freeman Dyson, who optimistically believes that we are on the threshold of developing new technologies that can scrub carbon from the atmosphere at

[9]See NASA (2007).

low cost (see Dyson 2008). In another example, Ray Kurzweil (2005) argues that we need to protect ourselves from the "GNR" (genetics, nanotechnology, robotics) revolution but believes that low-cost and clean energy will be attainable in two or three decades. We clearly need an economic and a statistical approach that can be generally applied to potentially catastrophic events.

The Role of Learning

There is an important difference among these many potentially catastrophic outcomes in terms of the potential for learning (Yohe and Tol 2010). For some catastrophes, we have no possibility of learning and making midcourse corrections. Edward Teller suggested that the Trinity test of an atomic bomb in 1945 might generate enough heat to ignite the atmosphere; this question could only be definitively answered by the test, but by then, it would be too late to defer the test. Strangelets are also in this category. There is no point in revising our views about strangelets in the microsecond after we discover that the calculations of the physicists were wrong and the Earth and all its life is turning into a small lump of hot strange matter. No midcourse correction would be possible. Rogue bugs may be in the same category as strangelets with respect to learning: Once they have escaped, they cannot be contained in the lab.

Climate change, by contrast, is a situation where we can learn as we go along. Every theory that allows for a very high climate sensitivity also predicts that we should see a very large warming now, with a rapid gradient over the next half century (IPCC 2007). This means that we can learn, and then act when we learn, and perhaps even do some geoengineering while we learn some more or get our abatement policies or low-carbon technologies in place. In other words, if the dismal theorem were to apply, it would apply primarily to areas where we have no reasonable chance of learning and taking midcourse corrections after learning that things are heading toward a catastrophic outcome.

Conclusions

There is increasing appreciation of the importance of tail events. Weitzman's dismal theorem holds that under strict conditions concerning the structure of uncertainty and preferences, society has an indefinitely large expected loss from high-consequence, low-probability events. In such situations, standard tools such as expected utility analysis cannot easily be applied. Thus, the dismal theorem is important because it helps identify when tail events have significance for our actions.

It must be emphasized, however, that the dismal theorem holds only under very limited conditions. The theorem requires strong risk aversion, a very fat tail for the uncertain variables, and the inability of society to learn and act in a timely fashion. Moreover, these properties must extend to indefinitely low consumption and indefinitely high values of the uncertain parameters.

Even if the dismal theorem does not hold in its strict form, there is still a constructive finding here. We see that it is very difficult to obtain precise information about the likelihood of extreme outcomes from observational data when the distributions have fat tails. This difficulty arises from ignorance about both the exact form of the distribution (e.g., normal, Pareto, or exponential) and the exact parameters of the distribution (supposing it to be Pareto, is the Pareto parameter 1, 1.5, 2, ...?). Returning to Figure 2, for example, we can see that a "four-sigma surprise" has a one in twenty-five probability of occurring with the Pareto distribution, but a near-zero probability with the normal distribution.

In many cases, the data speak softly or not at all about the likelihood of extreme events. This means that reasonable people may have quite different views about the likelihood of extreme events, such as the catastrophic outcomes of climate change, and that there are no data to adjudicate such disputes. This humbling thought applies more broadly, however, as there are indeed deep uncertainties about virtually every issue that humanity faces, and the only way these uncertainties can be resolved is through continued careful consideration and analysis of all data and theories.

References

Chapman, Clark R., and David Morrison. 1994. Impacts on the Earth by asteroids and comets: Assessing the hazard. *Nature* 367: 33–40.

Christensen, Kim, Leon Danon, Tim Scanlon, and Per Bak. 2002. Unified scaling law for earthquakes. *PNAS* 99 Suppl. no. 1:2509–13.

Ciriacy-Wantrup, S. V. 1952. *Resource conservation*. Berkeley: University of California Press.

Cropper, Maureen. 1976. Regulating activities with catastrophic environmental effects. *Journal of Environmental Economics and Management* 3: l–15.

Dar, Arnon, A. De Rújula, and Ulrich Heinz. 1999. Will relativistic heavy-ion colliders destroy our planet? *Physics Letters B* 470: 142–48.

Dyson, Freeman. 2008. The question of global warming. *The New York Review of Books* 55 (10), June 12.

Intergovernmental Panel on Climate Change (IPCC). 2007. Intergovernmental panel on climate change, climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC, Chapter 10.

Jaffea, R. L., W. Buszaa, J. Sandweiss, and F. Wilczek. 2000. Review of speculative 'disaster scenarios' at RHIC. *Review of Modern Physics* 72: 1125–40.

Kolstad, Charles D. 1996. Learning and stock effects in environmental regulation: The case of greenhouse gas emissions. *Journal of Environmental Economics and Management* 31: 1–18.

Kurzweil, Ray. 2005. *The singularity is near*. New York: Penguin.

Mandelbrot, Benoit. 1963. The variation of certain speculative prices. *The Journal of Business* 36 (4): 394–419.

Mills, Michael J., Owen B. Toon, Richard P. Turco, Douglas E. Kinnison, and Rolando R. Garcia. 2008. Massive global ozone loss predicted following regional nuclear conflict.. *Proceedings of the National Academy of Sciences of the United States of America* 105: 5307–12.

NASA. 2007. National Aeronautics and Space Administration, Near-Earth object survey and deflection: Analysis of alternatives. Report to Congress. http://www.nasa.gov/pdf/171331main_NEO_report_march07.pdf.

Nordhaus, William. 2007a. *A question of balance: Weighing the options on global warming policies*. New Haven, CT: Yale University Press.

———. 2007b. Who's afraid of a big bad oil shock? *Brookings Papers on Economic Activity* 2: 219–38.

Pindyck, Robert S. 2011. Fat tails, thin tails, and climate change policy. *Review of Environmental Economics and Policy,* 5 (2): 258–74. 10.1093/reep/rer005.

Schuster, Eugene F. 1984. Classification of probability laws by tail behavior. *Journal of the American Statistical Association* 79 (388):936–39.

Turco, R. P., O. B. Toon, T. P. Ackerman, J. B. Pollack, and Carl Sagan. 1983. Nuclear winter: Global consequences of multiple nuclear explosions. *Science* 222: 1283–92.

Weitzman, Martin L. 2009. On modeling and interpreting the economics of catastrophic climate change. *The Review of Economics and Statistics* 91 (1): 1–19.

———. 2011. Fat-tailed uncertainty in the economics of catastrophic climate change. *Review of Environmental Economics and Policy.* 5 (2): 275–92. 10.1093/reep/rer006.

Yohe, Gary, and Richard Tol. 2010. Precaution and a dismal theorem: Implications for climate policy and climate research. *In Risk management in commodity markets*, ed. Geman, Helyette, 91–9. New York: Wiley.