Expected Utility and Catastrophic Risk in a Stochastic Economy-Climate Model

Ikefuji, M.; Laeven, R.J.A.; Magnus, J.R.; Muris, C.H.M.

Publication date:
2010

CentER Discussion Paper

No. 2010-122

EXPECTED UTILITY AND CATASTROPHIC RISK IN A STOCHASTIC ECONOMY-CLIMATE MODEL

By Masako Ikefuji, Roger J.A. Laeven,
Jan R. Magnus, Chris Muris

November 2010

ISSN 0924-7815

TILBURG UNIVERSITY

Expected utility and catastrophic risk in a stochastic economy-climate model*

November 26, 2010

Masako Ikefuji
Institute of Social and Economic Research, Osaka University

Roger J.A. Laeven
Department of Econometrics & Operations Research, Tilburg University

Jan R. Magnus
Department of Econometrics & Operations Research, Tilburg University

Chris Muris
CentER, Tilburg University

*We are grateful to Graciela Chichilnisky, John Einmahl, Reyer Gerlagh, Johan Eyckmans, Sjak Smulders, Peter Wakker, Aart de Zeeuw, Amos Zemel, and seminar participants at the Tilburg Sustainability Center for helpful comments. This research was funded in part by the JSPS under grant C-22530177 (Ikefuji) and by the NWO under grant Vidi-2009 (Laeven). E-mail addresses: ikefuji@iser.osaka-u.ac.jp (Ikefuji), r.j.a.laeven@uvt.nl (Laeven), magnus@uvt.nl (Magnus), chrismuris@gmail.com (Muris).

Corresponding author:
Roger J.A. Laeven
Department of Econometrics and Operations Research
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands
e-mail: r.j.a.laeven@uvt.nl

Abstract: In the context of extreme climate change, we ask how to conduct expected utility analysis in the presence of catastrophic risks. Economists typically model decision making under risk and uncertainty by expected utility with constant relative risk aversion (power utility); statisticians typically model economic catastrophes by probability distributions with heavy tails. Unfortunately, the expected utility framework is fragile with respect to heavy-tailed distributional assumptions. We specify a stochastic economy-climate model with power utility and explicitly demonstrate this fragility. We derive necessary and sufficient compatibility conditions on the utility function to avoid fragility and solve our stochastic economy-climate model for two examples of such compatible utility functions. We further develop and implement a procedure to learn the input parameters of our model and show that the model thus specified produces quite robust optimal policies. The numerical results indicate that higher levels of uncertainty (heavier tails) lead to less abatement and consumption, and to more investment, but this effect is not unlimited.

JEL Classification: D81; Q5.

Keywords: Economy-climate models; Catastrophe; Expected utility; Heavy tails; Power utility.

1 Introduction

An economist, when asked to model decision making under risk and uncertainty for normative purposes, would typically work within the expected utility framework with constant relative risk aversion (that is, power utility). A statistician, on the other hand, would model economic catastrophes through probability distributions with heavy tails. Unfortunately, expected utility is fragile with respect to heavy-tailed distributional assumptions: expected utility may fail to exist or it may imply conclusions that are 'incredible'.

Economists have long been aware of this tension between the expected utility paradigm and distributional assumptions (Menger, 1934), and the discussions in Arrow (1974), Ryan (1974), and Fishburn (1976) deal explicitly with the trade-off between the richness of the class of utility functions and the generality of the permitted distributional assumptions. Compelling examples in Geweke (2001) corroborate the fragility of the existence of expected power utility with respect to minor changes in distributional assumptions.

The combination of heavy-tailed distributions and the power utility family may not only imply infinite expected utility, but also infinite expected *marginal* utility, and hence, via the intertemporal marginal rate of substitution (the pricing kernel), lead to unacceptable conclusions in cost-benefit analyses. For example, with heavy-tailed log-consumption and power utility, the representative agent should postpone *any* unit of current consumption to mitigate future catastrophes. The latter aspect was recently emphasized by Weitzman (2009) in the context of catastrophic climate change. Weitzman also argues that attempts to avoid this unacceptable conclusion will necessarily be non-robust.

In this paper we study the fundamental question of how to conduct expected utility analysis in the presence of catastrophic risks, in the context of extreme climate change. Our paper is built on four beliefs, which will recur in our analysis:

Catastrophic risks are important. To study risks that can lead to catastrophe is important in many areas, for example financial (trader, insurer, bank) distress, traffic accidents (bridge collapse, airplane crash, flight control system failure), dike bursts, killer asteroids, nuclear power plant disasters, and extreme climate change. Such low-probability high-impact events should not be ignored in cost-benefit analyses for policy making. In the context of extreme climate change: catastrophic climate changes, unlikely as they may be, should be accounted for in expected-welfare calculations for policy making.

A good model 'in the center' is not necessarily good 'at the edges'. Suppose we have estimated a function $C = a+bY$, relating consumption to disposable income. The dots in Figure 1 represent the data and the line gives the resulting OLS prediction $\hat{C} = \hat{a} + \hat{b}Y$. For incomes in the center, roughly

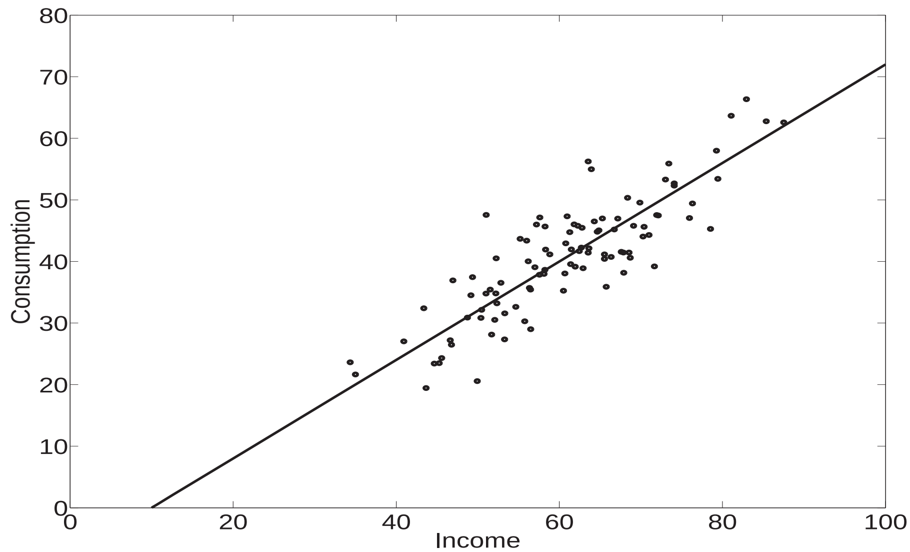

Figure 1: A consumption function

between 40 and 80, the consumption function can be well approximated by the regression line. How useful is this result for very low (or very high) incomes? Not very useful. For very low incomes, predicted consumption would be negative! This does not mean that a linear consumption function is useless. But it is only useful in the center of the domain. This is simply because models are not truths but approximations, and approximations may not work well if we move too far away from the point of approximation. In our context, the widely adopted class of power utility functions, often appropriate when one considers large inputs remote from zero as is common in macroeconomics and finance, may not work well for decision making under heavy-tailed risks with non-negligible support beyond the usual domains. Moreover, estimates of the coefficient of risk aversion are very sensitive to the particular domain of inputs that a utility function operates on (Rabin, 2000, footnote 10).

The price to reduce catastrophic risk is finite. Are we willing to spend everything to avoid children being killed at a dangerous street? Or to avoid the dikes to burst? Or a power plant to explode? Or a killer asteroid to hit the Earth? Or climate to change rapidly? No, we are not. To assume the opposite (that a society would be willing to offer all its current wealth to avoid or mitigate catastrophic risks) is not credible, not even from a normative (prescriptive, rational) perspective. In our context, there is a limit to the amount of current consumption that the representative agent is willing to give up in order to obtain one additional *certain* unit of future consumption, no matter how extreme and irreversible climate change may be. In other words: the expected pricing kernel is finite.

Light-tailed risks may result in heavy-tailed risk. When x is normally distributed (light tails) then $1/x$ has no moments (heavy tails). Also, when x is normally distributed then e^x has finite moments, but when x follows a Student distribution then e^x has no moments. In the context of extreme climate change: temperature has fluctuations but one would not expect heavy tails in its distribution. This does not, however, imply that functions of temperature cannot have heavy tails. For example, it may well be reasonable to use heavy-tailed distributional assumptions to model future (log) consumption.

There is an important literature on stochastic economy-climate models (see, for example, Keller *et al.*, 2004, Mastrandrea and Schneider, 2004, and the references therein). However, the integrated assessment models of climate economics are predominantly deterministic and rarely incorporate catastrophic risk (Ackerman *et al.*, 2010). To allow for uncertainty and extreme climate change, we start by specifying a stochastic economy-climate model that builds on Nordhaus' (2008) deterministic dynamic integrated climate and economy (DICE) model. We solve the model first with power utility and light-tailed distributional assumptions and prove that the assumption of expected power utility is incompatible with heavy-tailed distributional assumptions. We then address the question of how to conduct expected utility analysis in the presence of catastrophic risks. In Appendix C we provide necessary and sufficient conditions on the utility function, so that expected utility and expected marginal utility (hence the pricing kernel) are finite, also under heavy-tailed distributional assumptions. Restricting attention to utility functions that satisfy these compatibility conditions, we propose the two-parameter 'Burr' function as a particularly appealing utility function in our setting. We solve our stochastic economy-climate model with Burr utility (and also with the well-known exponential utility) under both light-tailed and heavy-tailed distributional assumptions.

Completing the resulting model requires specifying a number of model parameters as inputs. These parameters cannot 'simply' be determined by conventional statistical inference based on historical data. One reason is that temperature and other variables will be affected by economic policy decisions; another that economic parameters should be set so as to reflect rational decision-making behavior under circumstances that have never been encountered before. We discuss how to set the model parameters in a process towards agreement, using experts' priors on parameter values, and learning about parameters from resulting optimal model output. The key to the learning and agreement process is the translation of model parameters that are relatively difficult to interpret into quantities that allow a more straightforward interpretation. We find that our optimal policies are quite robust with respect to minor (and reasonable) changes to the input parameters.

Our numerical analysis indicates that allowing for heavy-tailed distributional assumptions in extreme climate change modeling leads to a reduction of current abatement and consumption and to an increase in current investment, when compared to a deterministic analysis. The increase in current investment may be interpreted via precautionary savings. Most notably and contrary to Weitzman (2009), while the differences are visible, they are not unlimited.

The paper is organized as follows. In Section 2 we propose a simplified version of Nordhaus' economy-climate model. There are two new features to this model: scrap value functions and, more importantly, uncertainty. In Section 3 we specialize this model to two periods only, and maximize expected welfare with non-linear scrap value functions. We also prove that expected welfare exists under normality but not under a Student distribution (Appendix B). Motivated by the fact that the expectation of the pricing kernel is finite for all outcome distributions whenever the concavity index (index of absolute risk aversion) $\mathrm{ARA}(x)$ is bounded (which we formally prove in Appendix C), we discuss such utility functions in Section 4: the well-known exponential function and the less well-known 'Burr' function. Section 5 discusses how we can learn the parameters of our model and calibrate policy using information such as the probability of catastrophe, and reports on robustness tests. Section 6 concludes. There are three appendices. Appendix A provides the Kuhn-Tucker conditions; Appendix B contains the proof of Proposition 3.1; and Appendix C discusses expected utility and tail uncertainty in a more general setting.

2 A simple stochastic economy-climate model

Our framework is a simple economy-climate model in the spirit of Nordhaus and Yang (1996) and Nordhaus (2008).

2.1 Emissions, temperature, and the economy

Everybody works. In period t, the labor force L_t together with the capital stock K_t generate GDP Y_t through a Cobb-Douglas production function

$$Y_t = A_t K_t^{\gamma} L_t^{1-\gamma} \qquad (0 < \gamma < 1),$$

where A_t represents technological efficiency and γ is the elasticity of capital. Capital is accumulated through

$$K_{t+1} = (1-\delta)K_t + I_t \qquad (0 < \delta < 1),$$

where I_t denotes investment and δ is the depreciation rate of capital. Production generates carbon dioxide (CO2) emissions E_t:

$$E_t = \sigma_t(1-\mu_t)Y_t,$$

where σ_t denotes the emissions-to-output ratio for CO2, and μ_t is the abatement fraction for CO2. The associated CO2 concentration M_t accumulates through

$$M_{t+1} = (1-\phi)M_t + E_t \qquad (0 < \phi < 1),$$

where ϕ is the depreciation rate of CO2 (rate of removal from the atmosphere). Temperature H_t develops according to

$$H_{t+1} = \eta_0 + \eta_1 H_t + \eta_2 \log(M_{t+1}) \qquad (\eta_1 > 0,\ \eta_2 > 0).$$

In each period t, the fraction of GDP not spent on abatement or 'damage' is either consumed (C_t) or invested (I_t) along the budget constraint

$$(1-\omega_t)d_t Y_t = C_t + I_t. \tag{1}$$

The temperature-impact function d_t depends only on temperature and satisfies $0 < d_t \leq \bar{d}_t$, where $\bar{d}_t$ represents the optimal temperature for the economy. Deviations from the optimal temperature cause damage. We specify d_t as

$$d_t = \frac{\bar{d}_t}{1+\xi H_t^2} \qquad (\xi > 0).$$

For very high and very low temperatures d_t approaches zero. The optimal value of d_t occurs at $H_t = 0$ (the temperature in 1900, as in Nordhaus) when $d_t = \bar{d}_t$. Hence, 'net' output $d_t Y_t$ is a fraction, not of Y_t as in Nordhaus, but of $\bar{d}_t Y_t$, the output achievable under optimal climate conditions. A fraction ω_t of $d_t Y_t$ is spent on abatement, and we specify the abatement cost fraction as

$$\omega_t = \psi_t \mu_t^{\theta} \qquad (\theta > 1).$$

If μ_t increases then so does ω_t, and a larger fraction of GDP will be spent on abatement. These equations capture the essence of the Nordhaus (2008) DICE model.

The model includes stock variables L_t, K_t, M_t, and H_t, fractions ω_t and μ_t, and scale variables A_t, d_t, σ_t, and ψ_t, all measured at the beginning of period t; and flow variables Y_t, C_t, I_t, and E_t, all measured in period t (not in year t). Notice that L_t is a stock, not a flow. As in Nordhaus (2008) one period is ten years. We choose the exogenous variables such that $L_t > 0$, $A_t > 0$, $\sigma_t > 0$, and $0 < \psi_t < 1$. The policy variables must satisfy

$$C_t \geq 0, \quad I_t \geq 0, \quad 0 \leq \mu_t \leq 1. \tag{2}$$

With these restrictions all variables will have the correct signs and all fractions will lie between zero and one.

2.2 Utility and welfare

Given a utility function U we define welfare in period t as

$$W_t = L_t U(C_t/L_t). \tag{3}$$

If the policy maker has an infinite horizon, then he/she will maximize total discounted welfare,

$$W = \sum_{t=0}^{\infty} \frac{W_t}{(1+\rho)^t} \qquad (0 < \rho < 1),$$

where ρ denotes the discount rate. Letting x denote per capita consumption, the utility function $U(x)$ is assumed to be defined and strictly concave for all $x > 0$. There are many such functions, but a popular choice is

$$U(x) = \frac{x^{1-\alpha} - 1}{1 - \alpha} \qquad (\alpha > 0), \tag{4}$$

where α denotes the elasticity of marginal utility of consumption. This is the so-called power function. If we define

$$\mathrm{ARA}(x) = -\frac{U''(x)}{U'(x)}, \qquad \mathrm{RRA}(x) = -\frac{xU''(x)}{U'(x)}, \tag{5}$$

then its coefficient of absolute risk aversion $\mathrm{ARA}(x) = \alpha/x$ is decreasing and its coefficient of relative risk aversion $\mathrm{RRA}(x) = \alpha$ is constant. The power function may be popular, but it does have drawbacks. In particular, we have RRA(0) > 0, which implies that the expected pricing kernel may not exist in the presence of heavy tails (Appendix C). Later we shall therefore consider other utility functions as well.

The power function is bounded from below, but not from above when $0 < \alpha < 1$; and it is bounded from above, but not from below when $\alpha > 1$. When $\alpha = 1$ we have $U(x) = \log(x)$, which is unbounded from below and above. Many authors, including Nordhaus (2008), choose $\alpha = 2$. Also popular is $\alpha = 1$ (Kelly and Kolstad, 1999; Stern, 2007).

Table 1: Parameter values for simplified DICE (SICE) model

Parameter	Value	Description
Endogenous stocks: initial levels		
K_0	137	Capital stock, begin of period 0
M_0	808.9	CO2 concentration, begin of period 0
H_0	0.731	Temperature, begin of period 0
Technology		
γ	0.30	Elasticity of capital in production function
δ	0.6513	Depreciation rate on capital, per decade
Pollution, damage, and abatement		
ϕ	0.0524	Depreciation rate on CO2 concentration, per decade
ξ	0.0028388	Quadratic term, temperature-impact function
θ	2.80	Exponent in abatement function
Temperature		
η_0	-5.9839	Constant term, temperature equation
η_1	0.7708	Previous period impact, temperature equation
η_2	0.9373	CO2 concentration impact, temperature equation
Discount rate		
ρ	0.1605	Welfare discount rate, per decade

Our interest is in maximizing welfare W with respect to the policy bundles (C_t, I_t, μ_t) for $t = 0, 1, 2, \ldots$. In Table 1 we present the parameters and initial values used. These values are chosen such that our results closely resemble the results obtained by Nordhaus (2008), when applied within the same 60-period (600-year) DICE framework. We choose the exogenous variables L_t, A_t, σ_t, and ψ_t as in Nordhaus (2008), and we let $\bar{d}_t = 1$ and $\alpha = 2$.

Our GAMS code (http://center.uvt.nl/staff/magnus/catastrophe) then produces optimal values over sixty periods that are very close to the values obtained in Nordhaus, as shown in Table 2. Hence it appears that our simplified version of the DICE model (hereafter, SICE = simplified DICE)

Table 2: Comparison of stocks in Nordhaus (DICE) and our (SICE) models

	2005		2055		2105		2155	
	DICE	SICE	DICE	SICE	DICE	SICE	DICE	SICE
K	137	137	353	354	707	711	1317	1324
M	809	809	1048	988	1270	1233	1428	1430
H	0.7	0.7	1.8	1.5	2.7	2.4	3.3	3.2

works as the original version.

2.3 Uncertainty

So far we have ignored uncertainty. There is however much uncertainty in the economics of climate change (Manne and Richels, 1992; Nordhaus, 1994; Weitzman, 2009). There is model uncertainty, parameter uncertainty, and uncertainty about the possible reduction of parametric variability over time (updating); see Kelly and Kolstad (1999) and Leach (2007). We model uncertainty through stochasticity. In the literature, stochasticity is typically introduced through the damage function (Roughgarden and Schneider, 1999; Mastrandrea and Schneider, 2004) or through a random shock in temperature (Kelly and Kolstad, 1999; Leach, 2007). We follow this literature by introducing stochasticity through the temperature-impact function d_t, more precisely through $\bar{d}_t$, the impact under optimal temperature. We are uncertain about the optimal temperature, because we are uncertain about the correctness of the functional form of d_t, about the values of the parameters, and about the underlying temperature equation. We capture these three sources of uncertainty by writing

$$\bar{d}_t = e^{-\tau^2/2} \, e^{\tau\epsilon_t},$$

where ϵ_t denotes a random error with mean zero and variance one. This implies that 'net GDP' is given by

$$d_t Y_t = \frac{e^{-\tau^2/2} Z_t}{1 + \xi H_t^2}, \qquad Z_t = A_t K_t^{\gamma} L_t^{1-\gamma} e^{\tau\epsilon_t}, \tag{6}$$

so that random noise enters the Cobb-Douglas production function in the usual 'linear' way when we write $\log(Z_t/L_t) = \log A_t + \gamma \log(K_t/L_t) + \tau\epsilon_t$.

If ϵ_t follows a normal distribution $\mathrm{N}(0,1)$, then the moments of $\bar{d}_t$ exist, and we have $\mathrm{E}(\bar{d}_t) = 1$ and $\mathrm{var}(\bar{d}_t) = e^{\tau^2} - 1$. Since the distribution of $\bar{d}_t$ is heavily skewed, its expectation is larger than its median, and hence more uncertainty (higher τ) implies more probability mass of $\bar{d}_t$ close to zero, and a higher probability of damage. If, however, we move only one step away from the normal distribution and assume that ϵ_t follows a Student distribution with *any* (finite) degrees of freedom, then the expectation is infinite (Geweke, 2001). (Heavy-tailed distributions (see Appendix C for a formal definition) such as the Student distribution are natural in the context of extreme climate change.) This fact predicts that expected welfare may be very sensitive to distributional assumptions: random noise with finite moments (Student distribution) may turn into random variables without moments ($\bar{d}_t$, d_tY_t).

2.4 Scrap values

If the policy maker has a T-period policy horizon, then we write welfare as

$$W = \sum_{t=0}^{T-1} \frac{L_t U(x_t)}{(1+\rho)^t} + \frac{1}{(1+\rho)^T} \sum_{t=0}^{\infty} \frac{L_{T+t} U(x_{T+t})}{(1+\rho)^t},$$

where $x_t = C_t/L_t$ denotes per capita consumption in period t. If $\{x_t^*\}$ denotes the optimal path for $\{x_t\}$, then we define the scrap value as

$$S_T = \sum_{t=0}^{\infty} \frac{L_{T+t} U(x^*_{T+t})}{(1+\rho)^t}.$$

Maximizing W is then equivalent to maximizing

$$\sum_{t=0}^{T-1} \frac{L_t U(x_t)}{(1+\rho)^t} + \frac{S_T}{(1+\rho)^T}.$$

The scrap value S_T will depend on the state variables at time T, in particular K_T and M_T, and this functional relationship is the scrap value function: $S_T = S(K_t, M_T)$. If T is large we may ignore the scrap value S_T because of the large discount factor $(1+\rho)^T$. But if T is small, then we need to model S_T explicitly, thus emphasizing the fact that the policy maker has the double objective of maximizing discounted welfare over a finite number of periods T, while also leaving a reasonable economy for the next policy maker, based on the remaining capital stock and CO2 concentration.

The simplest approximation to S_T is the linear function

$$S_T = \nu_0 + \nu_1 K_T - \nu_2 M_T \qquad (\nu_1 > 0, \nu_2 > 0), \tag{7}$$

where ν_1 and ν_2 denote the scrap prices of capital and pollution at the beginning of period T. This scrap value function captures the idea that the next government will be happier if there is more capital and less pollution at the beginning of its policy period. But the linear scrap value function has some problems. These are discussed in Ikefuji *et al.* (2010a), where we also propose the non-linear scrap value function,

$$S_T = \nu_0 - \frac{\nu_1 K_0}{p}\left(\frac{K_T}{K_0}\right)^{-p} - \frac{\nu_2 M_0}{q}\left(\frac{M_T}{M_0}\right)^{q}, \tag{8}$$

where $\nu_1 > 0$, $\nu_2 > 0$, $p > 0$, and $q > 1$. This function is strictly concave, bounded from above, and approaching $-\infty$ when either $M_T \to \infty$ or $K_T \to 0$. In addition, it has the property that if we linearize $S(K_T, M_T)$ around (K_0, M_0) we find

$$S_T \approx \text{constant} + \nu_1 K_T - \nu_2 M_T,$$

so that ν_1 and ν_2 can be interpreted as scrap prices, just as in the linear case.

3 A two-period model with CRRA preferences

The simplest version of the model occurs when $T = 2$ in which case we have only two periods. We can write welfare in this case as

$$W = W(\mu_0, C_0, \mu_1, C_1, \epsilon_1) = W_0 + \frac{W_1}{1+\rho} + \frac{S_2}{(1+\rho)^2}.$$

The two-period model, which we will consider henceforth, captures the essence of our problem while remaining numerically tractable in the presence of uncertainty. In this section the utility function is given by $U(x) = 1 - 1/x$ (power utility with $\alpha = 2$), the random errors ϵ_t are generated by a normal $\mathrm{N}(0, 1)$ distribution, and the policy restrictions (2) are explicitly imposed, so that we maximize a restriction of expected welfare; see Appendix A. Randomness results from d_1 only, because the temperature-impact d_0 at the beginning of period 0 is known to us (we set $\bar{d}_0 = 1$, equal to its expectation), and d_2 at the end of period 1 does not appear in the welfare function. Hence, the only source of randomness is caused by the error ϵ_1. The policy maker has to choose the policy bundles (C_0, I_0, μ_0) at the beginning of period 0 and (C_1, I_1, μ_1) at the beginning of period 1 that will maximize expected welfare.

We need values for the exogenous variables L_t, A_t, σ_t, and ψ_t. These are given in Table 3. Since a linear scrap value function is not realistic, because the combination of a half-bounded utility function and an unbounded scrap

Table 3: Exogenous variables in the two-period model

Parameter	Value	Description
Population		
L_0	6514	Population, begin of period 0
L_1	7130	Population, begin of period 1
Technology		
A_0	0.2722	Total factor productivity, begin of period 0
A_1	0.3000	Total factor productivity, begin of period 1
Pollution		
σ_0	0.1342	CO2 emissions-to-output ratio, period 0
σ_1	0.1253	CO2 emissions-to-output ratio, period 1
Abatement		
ψ_0	0.0561	Coefficient in abatement function, period 0
ψ_1	0.0511	Coefficient in abatement function, period 1

value function is theoretically not possible, we consider the non-linear scrap value function proposed in (8) with

$$\nu_1 = 286.15, \quad \nu_2 = 3.60, \quad p = 0.20, \quad q = 2.0.$$

Finally, we need sensible values for the uncertainty parameter τ. The stochasticity, as given in (6), captures uncertainty about GDP that is due to uncertainty about climate change. Historical variation in GDP may therefore serve as an initial upper bound proxy for τ. Barro (2009) calibrates the standard deviation of log GDP to a value of 0.02 on an annual basis. Over a 10-year horizon this would correspond to about 0.06, under normality. Barro, however, only considers rich (OECD) countries, which means that for our purposes this value needs to be scaled up.

In Figure 2 we plot the density of $\bar{d}_1$ for three values of τ: 0.1, 0.3, and 0.7, both when ϵ_1 follows a $\mathrm{N}(0,1)$ distribution (solid line) and when $\epsilon_1 = \sqrt{4/5}\,t$, where t follows a Student distribution with 10 degrees of freedom. Notice that $\mathrm{E}(\epsilon_1) = 0$ and $\mathrm{var}(\epsilon_1) = 1$ in both cases. When $\tau = 0.1$, we see that almost 100% of the distribution of $\bar{d}_1$ lies in the interval $(0.5, 2.0)$, both for the $\mathrm{N}(0,1)$ distribution and for the $t(10)$ distribution. When $\tau = 0.3$, 97.8% (97.2% for the Student distribution) lies in the interval $(0.5, 2.0)$; and, when $\tau = 0.7$, only 64.9% (67.2% for the Student distribution) lies in this interval. We conclude that $\tau = 0.7$ may serve as a credible upper bound for the uncertainty range, and hence we report our results for $\tau = 0.0$, 0.3, and 0.7.

Realizing that at the beginning of period 1 the temperature-impact d_1 is observed based on the realization of ϵ_1, the policy maker will maximize expected welfare in three steps as follows. First, he/she maximizes welfare

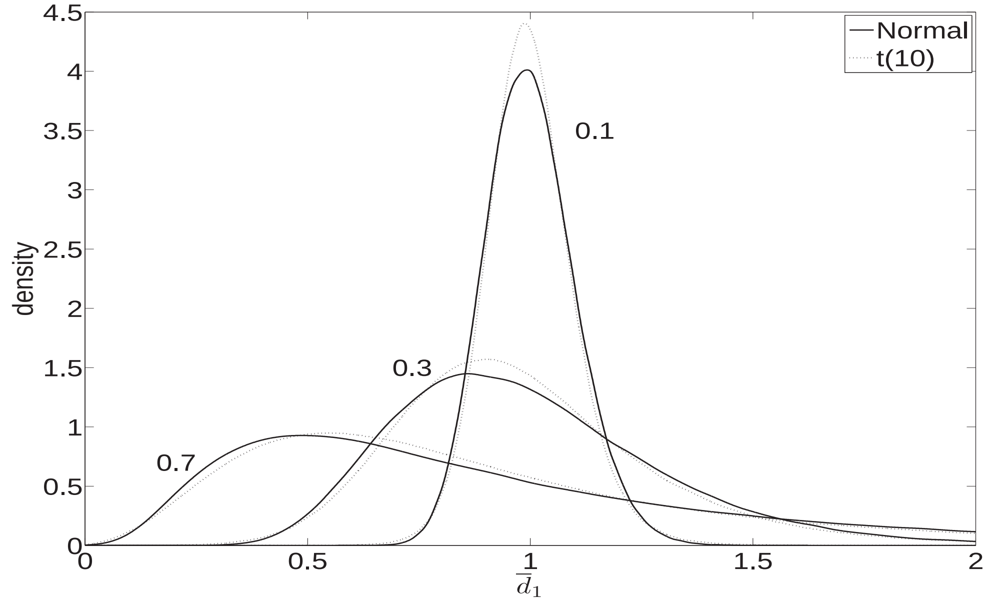

Figure 2: Density of $\bar{d}_1$ for $\tau = 0.1$, 0.3, and 0.7

$W = W(\mu_0, C_0, \mu_1, C_1, \epsilon_1)$ with respect to (μ_1, C_1) conditional on (μ_0, C_0, ϵ_1) and under the restriction (2). This gives (μ_1^*, C_1^*) and concentrated welfare

$$W^*(\mu_0, C_0, \epsilon_1) = W(\mu_0, C_0, \mu_1^*, C_1^*, \epsilon_1).$$

Then the expectation $\overline{W}(\mu_0, C_0) = \mathrm{E}\,(W^*(\mu_0, C_0, \epsilon_1))$ is computed, if it exists. Finally, $\overline{W}$ is maximized with respect to (μ_0, C_0). With the parameter values and exogenous variables given in Tables 1 and 3, and the four parameter values in the non-linear scrap value function given above, we obtain the results presented in Table 4. We note here and in subsequent tables that $Y_0 = 556.67$ and $d_0 = 0.9985$ are constant over different scenarios and functions, and that the values of μ_0, C_0, I_0, E_0, ω_0, K_1, M_1, and H_1 are optimal

Table 4: Power utility under normality

τ	0.0	0.3	0.7		0.0	0.3	0.7
Policy instruments				*Stocks*			
μ_0	0.0933	0.0920	0.0874	K_1	172.49	178.66	204.33
μ_1	0.1137	0.1141	0.1142	K_2	211.66	219.32	255.70
C_0	431.07	424.90	399.24	M_1	834.25	834.34	834.69
C_1	547.20	549.08	550.63	M_2	868.31	869.18	872.74
I_0	124.72	130.89	156.56	H_1	0.8843	0.8844	0.8848
I_1	151.51	157.03	184.45	H_2	1.0400	1.0411	1.0452

values. In contrast, μ_1, C_1, I_1, Y_1, E_1, ω_1, d_1, K_2, M_2, and H_2 are optimal *functions* of ϵ_1. What we present in the tables are their expectations.

For $\tau = 0$ there is no uncertainty. For $\tau > 0$ there is uncertainty, and all policy variables are affected when τ increases. More uncertainty results in less abatement, less consumption, and more investment in period 0, and to more abatement, consumption, and investment in period 1. In period 1, the changes in abatement and consumption are negligible. The increase in I_0 with τ can be explained by precautionary savings. The restriction on I_1 can be viewed as a penalty for negative investment. To avoid this penalty, the policy maker can increase the budget in period 1 by investing more in period 0. As the amount of uncertainty increases, the probability of negative investment increases, *ceteris paribus*. In response, the policy maker increases investment at the expense of abatement and consumption in period 0. The increase in I_0 leads to higher output in period 1, which explains the increases in I_1, K_2, M_2, and H_2. The decrease in μ_0 leads to higher emissions in period 0, and increases carbon concentration and temperature in period 1. An additional reason why investment in period 1 increases with uncertainty is that positive shocks translate into possibly unlimited upward shocks in I_1, but negative shocks will never cause I_1 to drop below zero.

We need to show that the expectation of welfare exists for power utility. The following proposition states not only this but also that, if we move one step away from normality and assume a Student distribution with any finite degrees of freedom, then the expectation does not exist.

PROPOSITION **3.1** *With power utility, expected welfare exists under normality but not under a Student distribution.*

The proof of Proposition 3.1 is in Appendix B. It follows that the much-used power utility function is inconsistent with expected utility theory with heavy tails, not because utility theory itself is at fault but because power utility is inappropriate when tails are heavy.

4 Catastrophic risk and compatibility

4.1 Expected utility and catastrophic risk

Since the axiomatization of expected utility (EU) by Von Neumann and Morgenstern (1944) and Savage (1954) numerous objections have been raised against it. Most of these relate to empirical evidence that the behavior of agents under risk and uncertainty does not agree with EU. Indeed there is much evidence that for descriptive applications the Von Neumann and

Morgenstern axioms are violated systematically. Motivated by such empirical evidence, various alternative theories have emerged, usually coined 'non-expected utility' theories; see Sugden (1997) for a review. Starting with the Allais paradox in the 1950s, problems involving low-probability high-impact outcomes have played a central role in non-expected utility. This is particularly, but not exclusively, evident in the rank-dependent class of models such as Kahneman and Tversky's cumulative prospect theory.

One option in our context would be to dismiss EU and replace it by a non-expected utility theory. While this could conceivably, at least partially, solve the problem of maximizing EU for catastrophic risks (in particular, the existence of moments), it might also aggravate the problems. Non-expected utility theories, most notably prospect theory, account for the fact that people are limited in their ability to comprehend and evaluate extreme probabilities, so that highly unlikely events are either ignored or overweighted. But, for normative purposes, it is dangerous to ignore or overweight highly unlikely events, and policy makers should choose a framework where this is avoided.

Despite important developments in non-expected utility theory, EU remains the dominant *normative* decision theory (Broome, 1991; Sims, 2001; Dhami and Al-Nowaihi, 2010), and the current paper stays within the framework of EU. Our results presented below corroborate the fact that expected utility theory may reliably provide normatively appealing results, also in the presence of catastrophic risks. Nevertheless, one may legitimately question whether EU is the appropriate normative theory for decision making under catastrophic risks and continue a search for better theories; see also Chichilnisky (2000).

In Appendix C we derive necessary and sufficient conditions on the utility function to ensure that expected utility and expected marginal utility (hence also the expected pricing kernel) are finite, also in the presence of heavy tails. These results are generally applicable to standard multi-period welfare maximization problems. This is important, because if the expected pricing kernel is infinite, then the amount of consumption in period 0 which the representative agent is willing to give up in order to obtain one additional certain unit of consumption in period 1 is infinite. This is not credible and, following our discussion in the Introduction, we argue that such a view of life is unreasonable and has extreme irrational implications in our setting. The price we are willing to pay to avoid a global economy-climate catastrophe is finite.

Propositions C.2 and C.3 underline the importance of a compatible specification of the concavity index, especially in the presence of catastrophic risk. We note in this context that the concavity index is to be specified as input, and need not be estimated, though it may be learned, that is, implicitly elicited (see Section 5 below). In what follows, we will provide translations of the concavity index parameters into quantities that allow a more straightforward interpretation. These translations can then serve as handles to set the input parameters. Quantities with a relatively simple meaning will be the key to the process of learning about the concavity index parameters.

4.2 Compatibility: Non-normality and Burr utility

Motivated by the conditions derived in Appendix C and by the fundamental insight that the economic model and the statistical model must be compatible, and because we wish to leave distributional assumptions unrestricted at this stage, we consider two bounded utility functions: the exponential function and the 'Burr' function. (Other choices are permitted but may require restrictions on distributional assumptions.) The exponential utility function is given by

$$U(x) = 1 - e^{-\beta x} \qquad (\beta > 0) \tag{9}$$

with $\text{ARA}(x) = \beta$ and $\text{RRA}(x) = \beta x$, and the Burr utility function by

$$U(x) = 1 - \left(\frac{\lambda}{x+\lambda}\right)^k \qquad (k > 0,\ \lambda > 0) \tag{10}$$

with $\text{ARA}(x) = (k+1)/(x+\lambda)$ and $\text{RRA}(x) = (k+1)x/(x+\lambda)$. Both functions are members of the HARA class of utility functions. The Burr function, based on Burr (1942) and Burr and Cislak (1968), was proposed in Ikefuji *et al.* (2010b), where it is also shown that this function is particularly appropriate as an approximation to a bounded utility function and enjoys a combination of appealing properties especially relevant in heavy-tailed risk analysis. This is exemplified in Figure 3, where we plot RRA and ARA for

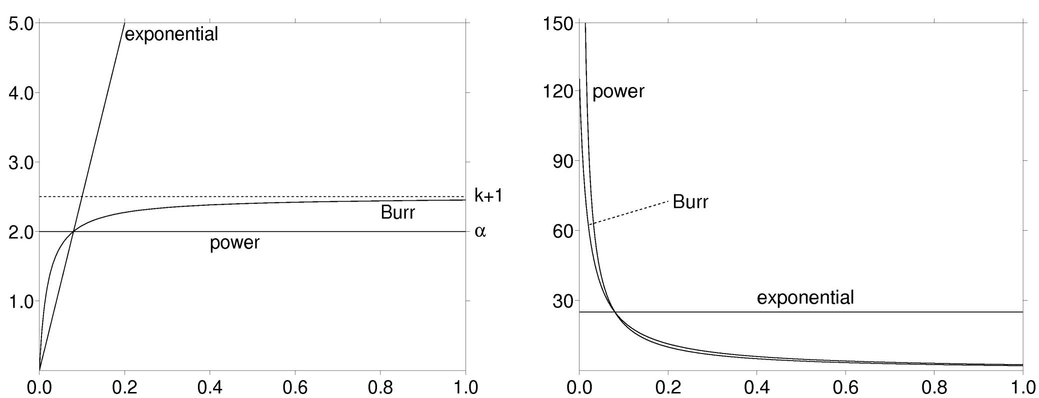

Figure 3: RRA (left) and ARA (right) for three utility functions

the power function ($\alpha = 2$), the exponential function ($\beta = 25$), and the Burr function ($k = 1.5$, $\lambda = 0.02$). The parameter choice is determined by the point x^*, where we want the three functions to be close. Suppose we want the functions to be close at $x^* = 0.08$, which is approximately the value of C_0/L_0 and C_1/L_1. Then, given that $\alpha = 2$, we find $\beta = 2/x^* = 25$, and, for any $k > 1$, $\lambda = (k-1)x^*/2$. Intertemporal preferences are jointly determined by the RRA parameters (α, β, k, λ) and the discount rate ρ. In our case we keep ρ constant and choose the RRA-ARA parameters appropriately according to the above closeness criterium.

The power function has RRA(0) > 0 and ARA(0) $= \infty$, while the RRA in the exponential function is unbounded for large x. In contrast, the RRA in the Burr function is bounded between 0 and $k+1$, it satisfies RRA(0) $= 0$, and ARA(0) is finite (125 in the figure). Notice that the fact that RRA(0) $= 0$ (as is the case for the exponential and the Burr utility functions) does *not* imply that the representative agent is risk-neutral at $x = 0$. For example, we have ARA(0) $= \beta$ for the exponential function and ARA(0) $= (k+1)/\lambda$ for the Burr function. Also, the derivative RRA$'$ is zero for the power function, constant for the exponential function, and monotonically decreasing from $(k+1)/\lambda$ to zero for the Burr function. Hence the slope of the Burr function at $x = 0$ is finite.

The Burr function is attractive because it lies in-between the power and exponential functions. It exhibits exponential-like features when x is close to zero (with $\text{RRA}(0) = 0, \text{ARA}(0) < \infty$, and $U(0) > -\infty$), hence satisfies the compatibility conditions derived in Appendix C; and power-like features in the middle and on the right-side of the distribution. Relative concepts are useful away from zero but not close to zero, and this is why power utility does not work well in near-catastrophe scenarios. Indeed, in the insurance literature exponential utility is typically used (Gerber, 1979, Chapter 5).

Table 5: Comparison of stocks in Exponential and Burr models

	2005		2055		2105		2155	
	Expo	Burr	Expo	Burr	Expo	Burr	Expo	Burr
K	137	137	286	343	388	666	456	1220
M	809	809	1012	993	1328	1258	1727	1512
H	0.7	0.7	1.6	1.5	2.6	2.5	3.7	3.3

Using GAMS and without uncertainty, we maximize welfare over sixty periods (600 years) for both the exponential and Burr utility functions, and a selection of the resulting optimal values is shown in Table 5. When we compare the results with those in Table 2, we see that the optimal stock values from the Burr function closely resemble the optimal stock values from the power function, but not those from the exponential function. In contrast to power and Burr, where RRA flattens out, the RRA for the exponential distribution continues to increase (Figure 3), and hence the growth rate of marginal utility continues to increase as well. As x increases, consumption will therefore increase, and investment and abatement will decrease. As a result, C/Y is relatively large for exponential utility. The low growth rate of capital (for exponential utility) leads to a low growth rate of output. However, since more consumption leads to less abatement, the growth rate of CO2 concentration is high even when the amount of production is low. Consequently, M and H are high compared to power and Burr. When $x < x^*$, RRA (Burr) is close to RRA (exponential), so that more is consumed and less invested when the Burr function is used instead of the power function. But when $x > x^*$, RRA (Burr) is close to RRA (power). The optimal path of K is slightly lower and the optimal paths of M and H are slightly higher for Burr than for power utility.

The scrap value function for both utility functions, developed in Ikefuji *et al.* (2010a), is defined as

$$S_T = \nu_0 - \nu_1\zeta_1 \left(1 + K_T/\lambda_1\right)^{-p} + \nu_2\zeta_2 \left(1 + (M_T/\lambda_2)^c\right)^{-q} \tag{11}$$

with

$$\zeta_1 = \frac{\lambda_1}{p} \left(1 + K_0/\lambda_1\right)^{p+1}, \qquad \zeta_2 = \frac{\lambda_2}{cq} \cdot \frac{\left(1 + (M_0/\lambda_2)^c\right)^{q+1}}{(M_0/\lambda_2)^{c-1}},$$

where $p > 0$, $q > 0$, $c > 1$, $\lambda_1 > 0$, and $\lambda_2 > 0$, and we have again normalized $\nu_1 > 0$ and $\nu_2 > 0$ such that $\partial S_T/\partial K = \nu_1$ at $K = K_0$ and $\partial S_T/\partial M = -\nu_2$ at $M = M_0$.

The values of the calibrated parameters for the scrap value functions in the case of exponential utility are $\nu_1 = 8.0282$, $\nu_2 = 0.1487$, and

$$p = 0.85, \quad q = 0.55, \quad c = 1.65, \quad \lambda_1 = 0.0936, \quad \lambda_2 = 0.0415.$$

The optimal values of the policy and other variables obtained from maximizing expected welfare are presented in Table 6. In contrast to Table 5, the results in Table 6 (and Table 7) allow for uncertainty, consider the short run (two periods) rather than the long run (sixty periods), and also take scrap values into account. Since exponential utility is calibrated to be close to power utility at $x = x^*$, the results for the two utility functions do not differ greatly. This is especially true for $\tau = 0$, where only the abatement fraction μ is higher for exponential utility, and therefore temperature H is lower.

When τ increases, I_0 increases less and I_1 increases more for exponential than for power. Moreover, as the uncertainty parameter τ increases, M_2 does

Table 6: Exponential utility: Normal versus Student(10)

		Normal		Student(10)	
τ	0.0	0.3	0.7	0.3	0.7
Policy instruments					
μ_0	0.1175	0.1166	0.1135	0.1166	0.1135
μ_1	0.1473	0.1515	0.1697	0.1516	0.1700
C_0	428.74	425.43	413.86	425.41	413.52
C_1	551.45	584.64	527.76	∞	∞
I_0	127.00	130.32	141.90	130.35	142.24
I_1	149.95	156.68	190.03	∞	∞
Stocks					
K_1	174.78	178.10	189.67	178.12	190.01
K_2	210.90	218.78	256.16	∞	∞
M_1	832.44	832.51	832.74	832.51	832.74
M_2	863.94	864.06	864.07	864.06	864.09
H_1	0.8823	0.8824	0.8826	0.8824	0.8826
H_2	1.0337	1.0339	1.0341	1.0339	1.0341

not change much in the exponential case, while it increases in the power case. The effect of uncertainty on the marginal scrap values is therefore larger in the exponential case than in the power case. As in Table 4, more uncertainty results in less abatement, less consumption, and more investment in period 0, and to more abatement, and investment in period 1.

Suppose now that the underlying distribution has heavier tails: Student instead of normal. We have normalized the variance of the Student distribution to be one, and therefore the first three moments of ϵ_1 are the same as under normality. The kurtosis, however, is now slightly higher: 4 instead of 3 (assuming 10 degrees of freedom). Under power utility, expected welfare does not exist any more. But under bounded utility, expected welfare always exists. The effect of the excess kurtosis on the optimal values is relatively small. It is important to realize that, while the Student distribution features excess kurtosis, it remains quite close to the normal distribution (see also Figure 2). Hence it would be unreasonable if a 'small' change in distributional assumptions would lead to a large possibly 'discontinuous' change in optimal policies.

All variables move in the same direction as before when τ increases. Notice that some variables (C_1, I_1, and K_2) have infinite expectations even though expected welfare is finite. This is no surprise because these variables are unbounded and depend on $\bar{d}_1 = e^{-\tau^2/2}e^{\tau\epsilon_1}$. When ϵ_1 follows a Student distribution, $\mathrm{E}(\bar{d}_1) = \infty$ and this property carries over to the other three variables.

Table 7: Burr utility: Normal versus Student(10)

		Normal		Student(10)	
τ	0.0	0.3	0.7	0.3	0.7
Policy instruments					
μ_0	0.0924	0.0910	0.0859	0.0910	0.0861
μ_1	0.1124	0.1135	0.1175	0.1135	0.1175
C_0	430.76	424.31	399.97	424.33	400.67
C_1	548.53	552.59	563.14	∞	∞
I_0	125.03	131.48	155.83	131.46	155.12
I_1	150.56	154.23	171.09	∞	∞
Stocks					
K_1	172.80	179.25	203.60	179.23	202.89
K_2	210.81	216.73	242.08	∞	∞
M_1	834.32	834.42	834.80	834.42	834.79
M_2	868.53	869.39	872.45	869.38	872.36
H_1	0.8844	0.8845	0.8850	0.8845	0.8849
H_2	1.0403	1.0413	1.0450	1.0413	1.0449

Let us now consider a second bounded utility function, the appealing Burr function. For Burr utility the following parameter values were calibrated for the scrap value functions: $\nu_1 = 2.8884$, $\nu_2 = 0.0366$, and

$$p = 0.6, \quad q = 0.5, \quad c = 1.5, \quad \lambda_1 = 0.1631, \quad \lambda_2 = 0.0502.$$

The optimal values are presented in Table 7. In view of Figure 3 we would expect that Burr and power are relatively close in the observed data range. This is indeed the case as a comparison of Tables 7 and 4 reveals. There is little difference between the two tables in the case of no uncertainty, and also when τ increases. The effect of excess kurtosis is again small, as it should be.

The important difference between power and Burr utility is not revealed in our typically observed data. It is only revealed when low levels of per capita consumption become relevant, that is, in near-catastrophe cases. This is clarified in Figure 4, where we present μ_1 as a function of ϵ_1 for $\tau = 0.3$. The expected value of μ_1 is 0.1141 for power utility under normality (Table 4), and 0.1135 for Burr utility under either normality or Student(10) (Table 7). This is not very different. But for values of ϵ_1 further away from 0 the difference is large.

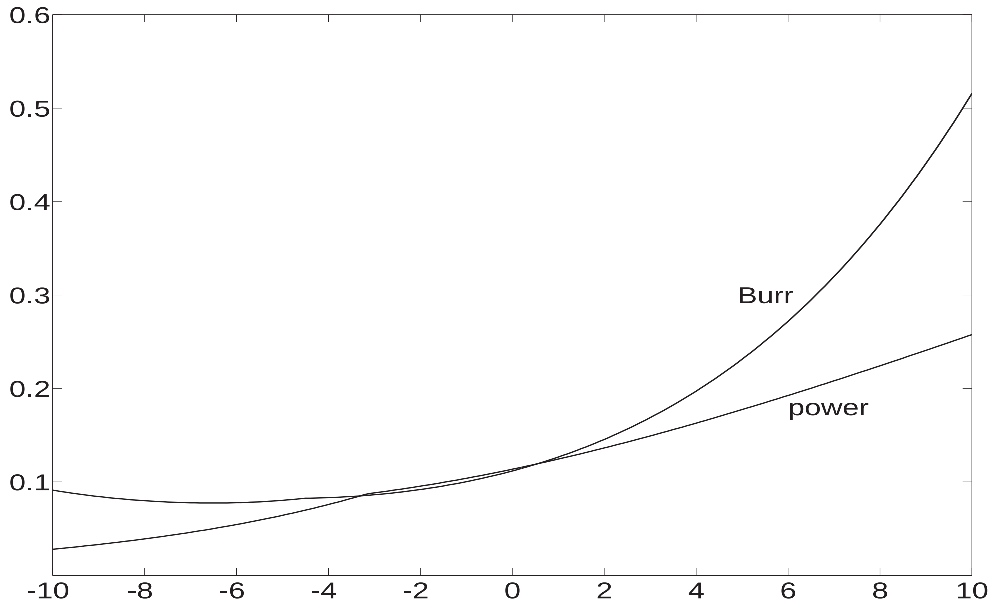

Figure 4: μ_1 as a function of ϵ_1: Burr versus power utility

5 Agreement and robustness

5.1 Learning and agreement

To complete the model we need to specify our input parameters. We show how this can be achieved in a process towards agreement, using experts' priors. The key to this learning and agreement process is the translation of model parameters that are relatively difficult to interpret into quantities that allow a more straightforward interpretation. The process is applicable not only in the current context of extreme climate change, but also in many other policy making settings involving catastrophic risks.

Our parameters cannot be estimated using conventional methods and historical data, but experts will have prior ideas about these parameters. Different experts will have different priors. Model output can be generated on the basis of various priors. Then, in an iterative procedure, one learns about the parameter values from experts' opinions and model output, and an agreeable intersection of model parameters may be reached.

This process is illustrated in Figure 5. In the left panel, we visualize the contributions of two experts. One expert states that the value of input 2 should be bounded as indicated by the two vertical lines. The other expert provides a lower and upper bound for the value of input 1, depending on the value of input 2. The horizontally-shaded area gives the combinations

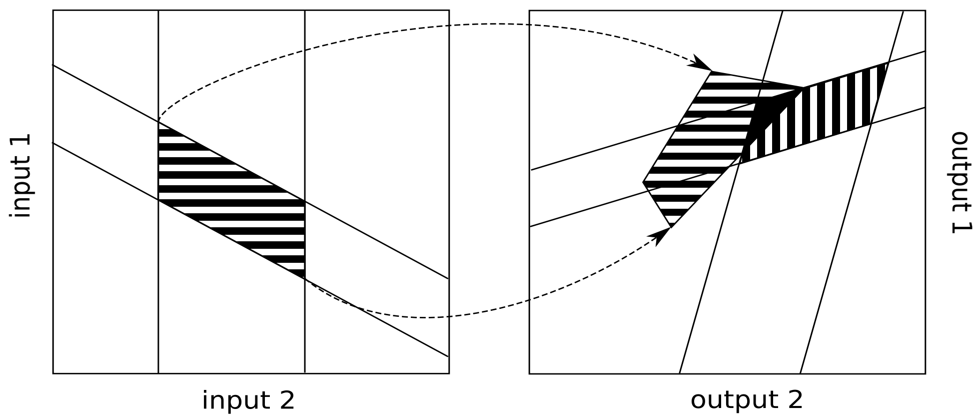

Figure 5: The decision making process

of inputs that are acceptable to both experts. The right panel is more complicated. We first visualize the contributions of two policy makers regarding two output variables. This is the vertically-shaded area, giving the combinations of outputs that are acceptable to both policy makers. Next we map the left panel onto the right panel. For every acceptable combination of inputs the model provides one combination of outputs, that is, one point in the right panel. The horizontally-shaded area in the right panel is the image of the horizontally-shaded area in the left panel. We now have two areas in the right panel: the vertically-shaded area and the horizontally-shaded area. If the two areas do not intersect, then the experts and policy makers must adjust their priors in an iterative process of learning. Once the areas do intersect, agreement is possible. The black triangle then contains all points for which both inputs and outputs are acceptable. Agreement must be reached on the three policy variables (μ_0, C_0, I_0), and we recall that expected welfare is maximized in three steps as described in Section 3, yielding the optimal policy (μ_0^*, C_0^*, I_0^*).

Our analysis requires prior beliefs about various inputs, in particular: form of the utility function (Burr or otherwise), degree of risk-aversion (k, λ), discount rate (ρ), form of the distribution (Student or otherwise), and volatility (τ). If agreement is to be reached, then the policy makers must be willing to adjust their individual priors on each of these inputs, based on the experts' opinions and the generated output.

Since extreme outcomes matter, the normal distribution is not appropriate. We want a distribution which allows heavier tails, such as the Student distribution. Given our treatment of stochasticity, power utility is not compatible with the Student distribution, because the required expectations don't exist. Also, exponential utility has the disadvantage that RRA increases without bound. Burr utility provides a useful compromise: it exhibits exponential-like features when per capita consumption is small, and power-like features otherwise. Let us then confine ourselves to Burr utility, assume that ϵ_1 follows a Student distribution, and take the following parameter values as our benchmark:

$$k = 1.5, \quad \lambda = 0.02 \quad \tau = 0.3, \quad \mathrm{df} = 10, \quad \rho = 0.1605.$$

Note that the value of λ is linked to k through $\lambda = 0.04(k-1)$, as explained in Section 4. The symbol df denotes the degrees of freedom in the Student distribution, and the discount rate of 0.1605 per decade corresponds to an annual discount rate of 0.015.

Table 8: Parameter calibration based on Burr utility and Student distribution

	Agreement				*Robustness*		
	a	*b*	*c*	*d*	*e*	*f*	*g*
Parameter values							
τ	0.3	0.3	0.5	0.3	0.5	0.7	0.5
df	10	25	10	10	25	10	10
k	1.5	1.5	1.5	2.0	1.5	1.5	2.0
Policy instruments, beginning of period 0							
μ_0	0.0910	0.0910	0.0888	0.1192	0.0887	0.0861	0.1163
C_0	424.33	424.31	413.71	438.01	413.50	400.67	427.56
I_0	131.46	131.47	142.08	117.73	142.29	155.12	128.19
Capital stock and expectations							
K_1	179.23	179.25	189.86	165.50	190.06	202.89	175.96
μ_1	0.1135	0.1135	0.1154	0.1604	0.1154	0.1175	0.1655
H_2	1.0413	1.0413	1.0429	1.0309	1.0430	1.0449	1.0323
Probabilities of catastrophe π_ℓ							
π_a	5.0E−03	3.3E−03	5.2E−02	5.1E−03	5.3E−02	1.4E−01	5.2E−02
π_b	2.3E−05	5.9E−07	1.4E−03	2.6E−05	5.5E−04	1.2E−02	1.5E−03
π_c	2.5E−07	5.0E−11	2.8E−05	2.6E−07	8.2E−07	4.9E−04	3.0E−05
Values of statistical subsistence $V_\ell = VSS_\ell/C_0$							
V_a	2.8E+01	2.9E+01	4.8E+00	2.4E+01	4.1E+00	2.9E+00	4.1E+00
V_b	1.3E+04	2.8E+05	3.1E+02	1.1E+04	5.3E+02	5.2E+01	2.6E+02
V_c	1.9E+06	4.2E+09	2.0E+04	1.6E+06	3.6E+05	1.4E+03	1.7E+04

Our benchmark is column a in Table 8. The model outputs are within credible bounds: policy variables at the beginning of period 0 (μ_0, C_0, I_0); stock variables at the beginning of period 1 (K_1 and also $M_1 = 834.42$ and $H_1 = 0.8845$); and expectations ($\mathrm{E}(\mu_1)$, $\mathrm{E}(H_2)$, and also $\mathrm{E}(M_2) = 869.38$). If we consider temperature H_2 as a function of ϵ_1 we find relatively low volatility in comparison to the confidence intervals proposed by the IPCC (2007, Chapter 10). The reason for this is twofold. First, the IPCC determines confidence intervals by considering multiple deterministic climate models, not a single stochastic one as we do. Second and more importantly, the IPCC confidence intervals are based on non-mitigation scenarios, while our model takes policy effects into account. For both reasons, the volatility in temperature found by the IPCC is higher than what we find.

5.2 Probability of catastrophe and value of statistical subsistence

In addition to the 'direct' outputs of our model we also have 'derived' outputs, in particular the probability of catastrophe. These derived outputs are functions of the direct outputs and they represent important policy variables on which prior information is available. Hence, they also require agreement.

We propose to define catastrophe as the event $C_1^* \leq \underline{C}$ for some given value $\underline{C} > 0$. The probability of catastrophe is then given by $\pi = \Pr(C_1^* \leq \underline{C})$. We shall consider three different values of $\underline{C}$: $\underline{C}_a$, $\underline{C}_b$, and $\underline{C}_c$, corresponding to three levels of catastrophe, labeled A, B, and C. Catastrophe A occurs when 20% of the world population live in extreme poverty, and catastrophes B and C occur when 50% and 80% of the world population live in extreme poverty, respectively. (The definitions and priors proposed in this subsection are based on background material provided in Ikefuji *et al.* (2010c) and available at the project's website http://center.uvt.nl/staff/magnus/catastrophe.)

We must agree on acceptable values for the probability π of catastrophe. We have studied acceptable risks in various situations, and we conclude that an acceptable probability for an economy-climate catastrophe in the next 10-year period is in the range 10^{-5}–10^{-6}. Given the definition of catastrophe we propose: $\pi_a = 0.1$, $\pi_b = 0.001$, and $\pi_c = 0.00001$ as reasonable values. There are of course other definitions of catastrophe. Barro and Ursúa (2008) define catastrophe as a peak-to-trough fall in per capita GDP of at least 15%, and find that the probability of this happening is approximately $\pi = 0.017$ per year. This does not relate directly to our π values, because GDP is not the same as consumption, Barro and Ursúa consider one year while we work with 10-year intervals, and, most importantly, because they only consider 21 'rich' countries. So, the numbers are difficult to compare. A situation where 20% of the world live in extreme poverty has in fact occurred before. The percentage of the world's population living in extreme poverty has halved since 1981. So a probability of $\pi_a = 0.1$ seems reasonable.

In the benchmark model we find $\pi_a = 0.005$, $\pi_b = 0.00002$, and $\pi_c = 0.0000003$, which is much lower than the acceptable values. Given the associated costs, it seems unnatural that policies would be chosen that mitigate the probability of a global economy-climate catastrophe far beyond acceptable levels. What can we do about this? One possibility is to make the tails heavier or lighter, that is, to adjust the degrees of freedom. If we set df = 25 then π becomes even smaller. In general, π becomes smaller as the tails become lighter (df increases), as one would expect. For df = ∞ (the normal distribution) we find $\pi_a = 2.3\text{E}{-}03$, $\pi_b = 5.3\text{E}{-}10$, and $\pi_c = 1.5\text{E}{-}24$. Interestingly, the policy variables are hardly affected (column b), not even when df = 200 or df = ∞. If we set df = 3, which is the minimum value where $\text{var}(\epsilon_1)$ exists, then $\pi_a = 0.008$, a little higher than for df = 10, but not enough. So, adjusting the degrees of freedom hardly changes the results.

Perhaps the fact that the heaviness of the tail (degrees of freedom) has little effect on the optimal policy is caused by the Burr utility function. Maybe this function does not distinguish well between light and heavy tails? In fact, this is not so. It follows from Figure 2 (and Section 4.2) that τ has much more impact than df. Hence the Burr function does distinguish between light and heavy tails.

Perhaps we should then adjust the value of τ. In our benchmark we set $\tau = 0.3$ as a reasonable starting point. We could revise τ upwards. We argued in Section 3 and Figure 2 that $\tau = 0.7$ is an upper bound to the volatility. Let us therefore consider the case $\tau = 0.5$. A larger value of τ means more volatility and hence one would expect less consumption and more investment. This is indeed what happens (column c). Also, the probabilities are affected and are now much closer to our prior ideas.

We can also adjust the curvature k (and λ). If k increases, then agents become less risk-averse and, as expected, there is more consumption and less investment (column d). The probabilities are not much different from our benchmark in a, but the values of μ_0 and μ_1 are very high and the capital stock accumulation rate is only 1.9% per year, which is too low.

Finally, we could adjust the discount rate ρ. This is an important issue (see, for example, Gollier, 2002, 2008, and the references therein), with possibly significant (yet not 'discontinuous') impact on the optimal policies. It is, however, beyond the scope of this study.

Based on these comparisons it seems that policy c should be recommended. There is, however, one other derived output which is often discussed, namely the value of statistical life. If we agree on the definition of catastrophe, then we can also define the 'value of a statistical subsistence' (VSS) as the amount of consumption in period 0 that the government is willing to trade off in order to change the probability of catastrophe; see Ikefuji *et al.* (2010c). The VSS is similar to the value of statistical life (VSL), except that it refers to the condition of just having enough food to stay alive (more than \$1/day) rather than to life. We define (Ikefuji *et al.*, 2010c)

$$\mathrm{VSS} = \frac{1}{\partial\pi(C_0)/\partial C_0}$$

evaluated at $C_0 = C_0^*$. We propose $\mathrm{VSS}_a = C_0$, $\mathrm{VSS}_b = 10C_0$, and $\mathrm{VSS}_c = 100C_0$ as reasonable orders of magnitude. We first need to establish that the VSS exists. A little algebra shows that the VSS can be expressed as $\mathrm{VSS} = \Gamma_1/\Gamma_2$ with

$$\Gamma_1 = \mathrm{E}\left(W_1(C_0^*, C_1^*) \,|\, C_1^* > \underline{C}\right) - \mathrm{E}\left(W_1(C_0^*, C_1^*) \,|\, C_1^* \le \underline{C}\right)$$

and

$$\begin{aligned}\Gamma_2 = {} & (1+\rho)\partial W_0(C_0)/\partial C_0 + (1-\pi)\partial\,\mathrm{E}\left(W_1(C_0, C_1^*) \,|\, C_1^* > \underline{C}\right)/\partial C_0 \\ & + \pi\partial\,\mathrm{E}\left(W_1(C_0, C_1^*) \,|\, C_1^* \le \underline{C}\right)/\partial C_0 + \frac{1}{1+\rho}\partial\,\mathrm{E}\left(S_2(C_0, C_1^*)\right)/\partial C_0,\end{aligned}$$

and where π and all partial derivatives are evaluated at $C_0 = C_0^*$. Under Burr utility all expectations exist.

The VSS (and the VSL) is a difficult concept to measure, and the VSS priors may be unreliable. As such it should not carry too much weight as a derived output. Still we notice that the VSSs of our preferred policy *c* are much closer to our reasonable values than the VSSs in columns *a*, *b*, and *d*. The VSSs in column *c* are quite high though. Apparently society is willing to sacrifice $5C_0$ to avoid catastrophe *A* and even $20{,}000C_0$ to avoid catastrophe *C*. Perhaps our 'reasonable' values are too small. In fact, this is a well-known problem. Weitzman (2009) discusses it and he also mentions large values for the VSL without it being clear what the consequences are.

5.3 Robustness

If we believe that column *c* is the best, then we should do some further robustness checks, starting from column *c* rather than column *a*. We have done extensive robustness checks and some representative results of this analysis is reported in columns *e*–*g* of Table 8. If we adjust the degrees of freedom (column *e*), then not much happens. There is little to choose between columns *c* and *e*. The optimal policy (μ_0^*, C_0^*, I_0^*) is hardly affected, which is a good thing, because it means that our policy is not too sensitive to changes in the heaviness of the tail (degrees of freedom). In column *f* we consider $\tau = 0.7$.

Here the probabilities of catastrophe seem to be too large. For example, we have $\pi_c = 0.0005$ and it is doubtful if the government would find this acceptable. The choice of volatility τ does, however, affect the policy, and hence is important. We see that economics and statistics are bound together and difficult to separate. In column g we adjust the curvature of the Burr utility function. The probabilities are hardly affected but there will be more consumption, less investment, and in particular more (perhaps too much) abatement. On the basis of these and other robustness checks we conclude that policy c is robust against small changes in the underlying assumptions and parameter values.

5.4 Weitzman's dismal theorem

The previous discussion is closely related to an important debate initiated by Weitzman (2009). In a highly stylized setting, Weitzman notices that heavy-tailed uncertainty and power utility are incompatible, as this combination of uncertainty and preferences implies an infinite pricing kernel. In order to avoid this, Weitzman introduces a lower bound on consumption. He then shows that this lower bound is related to a parameter that resembles the value of a statistical life, and proves that the pricing kernel approaches infinity as the value of this parameter approaches infinity (the 'dismal theorem'). Weitzman further argues that this 'VSL-like' parameter is hard to know, and interprets this result as follows:

> "...reasonable attempts to constrict the length or the heaviness of the 'bad' tail (or to modify the utility function) still can leave us with uncomfortably big numbers whose exact value depends non-robustly upon artificial constraints or parameters that we really do not understand." (Weitzman, 2009, p. 11)

We agree with Weitzman that incompatible combinations of utility functions and distribution functions exist, in the sense that the pricing kernel or other important policy variables become infinite. In fact we derive necessary and sufficient conditions on the utility functions for the pricing kernel to exist (Appendix C). But we object to the dismal theorem for two reasons.

First, we think that the result is implied by using an incorrectly specified model. A key ingredient in Weitzman's model is the power utility function. This popular utility function is characterized by constant relative risk aversion (CRRA). The assumption of CRRA (hence $\text{RRA}(0) > 0$) is not appropriate when dealing with extremely low levels of consumption, and it is exactly the behavior at these low consumption levels that leads to the dismal theorem. As we have demonstrated in Section 4, Weitzman's result is avoided when the economic model (utility function) is compatible with the statistical model (heavy tails). Utility functions with appropriate risk aversion for low levels of consumption are not subject to the dismal theorem.

Second, more effort can be made to know an input parameter that is 'hard to know', and we describe a (stylized) learning-and-agreement procedure for precisely this purpose in Sections 5.1 and 5.2. Although it is difficult to state upper and lower bounds for the 'VSL-like' input parameter, we can still obtain reasonable constraints on difficult-to-know parameters of interest indirectly. The economic model translates the parameter of interest into output variables with an easier interpretation (such as the optimal policies and the probability of catastrophe). Bounds on these output variables, together with the economic model, imply bounds on the parameter of interest.

6 Conclusions

Our strategy in this paper has been to first specify and analyze a stochastic economy-climate model using the popular power utility function. Section 3 demonstrates explicitly that power utility is fragile with respect to distributional assumptions. This is not unexpected. Weitzman (2009) summarizes this fragility and the non-existence of a robust solution in a 'dismal' theorem. We agree with Weitzman's concerns about the validity of expected utility analysis in settings featuring catastrophic risks. We argue that one should indeed allow for heavy-tailed distributions when modeling catastrophic climate changes, but that, in contrast to Weitzman, heavy-tailed distributional assumptions are not *per se* irreconcilable with expected utility.

Based on general results regarding the relationship between the richness of the class of utility functions and the generality of the permitted distributional assumptions (Section 4 and Appendix C), we then restrict ourselves to utility functions that are compatible with our distributional assumptions. In Section 4 we propose that on the domain that contains the typically observed consumption levels, the utility function behaves power-like (CRRA) as is popular in macroeconomics and finance, while on the remote domain containing extreme adverse consumption shocks, the utility function exhibits exponential-like (CARA) features as is popular in insurance. Thus we avoid the unacceptable conclusion that society should sacrifice an unlimited amount of consumption to reduce the probability of catastrophic climate change by even a small amount. After reaching agreement on the model parameters, the sensitivity analysis conducted in Section 5 shows that our completed model and the resulting optimal policies are quite robust and sensibly sensitive. With quasi exponential-like behavior of the utility function in near-catastrophe situations, extreme sensitivities that would otherwise be present using CRRA preferences are avoided.

Much of the analysis in our paper is not limited to extreme climate change. A similar analysis could apply in other policy making settings involving catastrophic risks, such as the development of new financial incentive schemes to mitigate the risk of extreme systemic failures and resulting financial economic crises, or policies concerning medical risks (pandemic flu and vaccination risks).

Let us finally admit four limitations of our paper, and indicate possible generalizations. First, from Section 4 onwards, we have focussed our attention on bounded utility functions, so as to avoid having to restrict distributional assumptions. In general, one could assume more structure on stochasticity (yet still allow for heavy tails) and broaden the constraints on utility. In particular, unbounded utility (such as HARA with $0 < \alpha \leq 1$; see Ikefuji *et al.*, 2010b) is also permitted under additional assumptions on stochasticity. Second, for simplicity and clarity of presentation, we have restricted our analysis to only two periods. In principle, much of our analysis will remain intact when considering more than two periods. Third, to account for the fact that the policy maker has the double objective of maximizing current consumption, while also leaving a reasonable economy for the next policy maker, we have used scrap values in our analysis. We ignore, however, stochasticity in the scrap value function after the second period. The development of a numerically tractable economy-climate model with multi-period stochasticity *and* scrap values is left for future research. Finally, the equations making up our stochastic economy-climate model are of a simple and stylized nature, and each one of them, including the specification of stochasticity, leaves room for generalizations and extensions.

Appendices

A Kuhn-Tucker conditions under positive investment

Consider the economy-climate model of Section 2 in the two-period set-up of Section 3. Let U be a general well-behaved utility function and let $S^{(1)}$ and $S^{(2)}$ be general well-behaved scrap value functions. At the beginning of period 1 our welfare function, conditional on (C_0, μ_0, ϵ_1), is

$$W = L_1 U(C_1/L_1) + \nu_1 S^{(1)}(K_2) - \nu_2 S^{(2)}(M_2).$$

We have four constraints: $C_1 \geq 0$, $I_1 \geq 0$, $\mu_1 \geq 0$, and $\mu_1 \leq 1$, but only two of these can be binding as we shall see. Hence, we define the Lagrangian $\mathcal{L} = \mathcal{L}(C_1, \mu_1)$ as

$$\mathcal{L} = L_1 U(C_1/L_1) + \nu_1 S^{(1)}(K_2) - \nu_2 S^{(2)}(M_2) + \kappa_1 I_1 + \kappa_2(1 - \mu_1),$$

and we find

$$\frac{\partial \mathcal{L}}{\partial C_1} = U'(C_1/L_1) - (\nu_1 g_1 + \kappa_1)$$

and

$$\frac{\partial \mathcal{L}}{\partial \mu_1} = \left(-(\nu_1 g_1 + \kappa_1)\psi_1 \theta \mu_1^{\theta-1} d_1 + \nu_2 g_2 \sigma_1\right) Y_1 - \kappa_2,$$

where

$$g_1 = g_1(C_1, \mu_1) = \frac{\partial S^{(1)}(K_2)}{\partial K_2}, \qquad g_2 = g_2(\mu_1) = \frac{\partial S^{(2)}(M_2)}{\partial M_2}.$$

This leads to the Kuhn-Tucker conditions:

$$\kappa_1 = U'(C_1/L_1) - \nu_1 g_1 \geq 0,$$
$$I_1 = (1 - \psi_1 \mu_1^{\theta}) d_1 Y_1 - C_1 \geq 0,$$

and

$$\kappa_2 = \left(-U'(C_1/L_1)\psi_1 \theta \mu_1^{\theta-1} d_1 + \nu_2 g_2 \sigma_1\right) Y_1 \geq 0,$$
$$\mu_1 \leq 1,$$

together with the slackness conditions $\kappa_1 I_1 = 0$ and $\kappa_2(1 - \mu_1) = 0$.

Under the assumption that $I_1 > 0$ we have $\kappa_1 = 0$ and we distinguish between two cases, as follows.

Case (1): $\kappa_2 > 0$. We have $\mu_1 = 1$ and $g_2 = g_2(1)$, and we solve two equations in two unknowns:

$$U'(C_1/L_1) = \nu_1 g_1, \qquad g_1 = g_1(C_1, 1),$$

under the restrictions:

$$\frac{C_1}{(1 - \psi_1) Y_1} \leq d_1 < \frac{\nu_2 g_2 \sigma_1}{\nu_1 g_1 \psi_1 \theta}.$$

Case (2): $\kappa_2 = 0$. We solve four equations in four unknowns:

$$U'(C_1/L_1) = \nu_1 g_1, \qquad \mu_1^{\theta-1} d_1 = \frac{\nu_2 g_2 \sigma_1}{\nu_1 g_1 \psi_1 \theta},$$

$$g_1 = g_1(C_1, \mu_1), \qquad g_2 = g_2(\mu_1),$$

under the restrictions:

$$C_1 \leq (1 - \psi_1 \mu_1^\theta) d_1 Y_1, \qquad \mu_1 \leq 1.$$

The following two points are worth noting. First, we see that the restrictions $\mu_1 \geq 0$ and $C_1 \geq 0$ are automatically satisfied, so that they do not need to be imposed. Second, we see that $U'(C_1/L_1) = \nu_1 g_1$ in both cases. This fact will be used in Appendix B.

B Proof of Proposition 3.1

We shall prove the proposition both for the linear scrap and the non-linear scrap case. In both cases the inequality constraints (2) are imposed. Since

$$d_1 Y_1 = B_1 e^{\tau \epsilon_1}, \qquad B_1 = \frac{e^{-\tau^2/2} Y_1}{1 + \xi H_1^2},$$

we obtain

$$\begin{aligned} C_1^* \leq C_1^* + I_1^* = (1 - \omega_1^*) d_1 Y_1 \leq B_1 e^{\tau \epsilon_1}, \\ I_1^* \leq C_1^* + I_1^* \leq B_1 e^{\tau \epsilon_1}, \\ (1-\delta) K_1 \leq K_2^* \leq (1-\delta) K_1 + B_1 e^{\tau \epsilon_1}, \end{aligned} \tag{12}$$

and

$$M_2^* \leq (1 - \phi) M_1 + \sigma_1 Y_1.$$

We distinguish between three cases.

Linear scrap under normality. Linear scrap implies that $S^{(1)}(K_2) = K_2$ and $S^{(2)}(M_2) = M_2$. Since $\mathrm{E}(e^{\tau \epsilon_1})$ exists under normality, it follows that C_1^*, I_1^*, K_2^*, and M_2^* all have finite expectations, and therefore that $\mathrm{E}(W^*)$ exists if and only $\mathrm{E}(1/C_1^*)$ exists. For notational convenience we do not distinguish between the random variable ϵ_1 and its realization. With this slight abuse of notation, we write

$$\begin{aligned} \mathrm{E}(1/C_1^*) &= \int_{-\infty}^{\infty} (1/C_1^*)\, dF(\epsilon_1) = \int_{I_1^*=0} (1/C_1^*)\, dF(\epsilon_1) + \int_{I_1^*>0} (1/C_1^*)\, dF(\epsilon_1) \\ &= (1/B_1) \int_{I_1^*=0} \frac{e^{-\tau \epsilon_1}}{1 - \omega_1^*}\, dF(\epsilon_1) + \int_{I_1^*>0} (1/C_1^*)\, dF(\epsilon_1) \\ &\leq \frac{1}{(1-\psi_1) B_1} \mathrm{E}(e^{-\tau \epsilon_1}) + \int_{I_1^*>0} (1/C_1^*)\, dF(\epsilon_1). \end{aligned}$$

Since $\mathrm{E}(e^{-\tau\epsilon_1})$ is finite, it suffices to show that $\int_{I_1^*>0}(1/C_1^*)\,dF(\epsilon_1)$ is finite. Now, it follows from Appendix A that, under the assumption that $I_1^* > 0$, $U'(C_1^*/L_1) = L_1^2/C_1^{*2} = \nu_1 g_1^* = \nu_1$, because $g_1^* = 1$. Hence,

$$\int_{I_1^*>0}(1/C_1^*)\,dF(\epsilon_1) = \frac{\nu_1^{1/2}}{L_1}\Pr(I_1^* > 0) \leq \frac{\nu_1^{1/2}}{L_1} < \infty.$$

Nonlinear scrap under normality. Nonlinear scrap implies that

$$S^{(1)}(K_2) = -\frac{K_0}{p}\left(\frac{K_2}{K_0}\right)^{-p}, \qquad S^{(2)}(M_2) = \frac{M_0}{q}\left(\frac{M_2}{M_0}\right)^{q}$$

where $p > 0$ and $q > 1$. Since

$$(K_2^*)^{-p} \leq ((1-\delta)K_1)^{-p}$$

and

$$(M_2^*)^q \leq ((1-\phi)M_1 + \sigma_1 Y_1)^q\,,$$

we see that $\mathrm{E}(W^*)$ exists if and only $\mathrm{E}(1/C_1^*)$ exists. As in the linear scrap case, it suffices to show that $\int_{I_1^*>0}(1/C_1^*)\,dF(\epsilon_1)$ is finite. Since

$$g_1 = g_1(K_2) = \frac{\partial S^{(1)}(K_2)}{\partial K_2} = \left(\frac{K_0}{K_2}\right)^{p+1},$$

it follows from Appendix A that, under the assumption that $I_1^* > 0$,

$$U'(C_1^*/L_1) = L_1^2/C_1^{*2} = \nu_1 g_1^* = \nu_1\left(\frac{K_0}{K_2^*}\right)^{p+1} \leq \nu_1\left(\frac{K_0}{(1-\delta)K_1}\right)^{p+1},$$

and hence that

$$\int_{I_1^*>0}(1/C_1^*)\,dF(\epsilon_1) \leq \frac{\nu_1^{1/2}}{L_1}\left(\frac{K_0}{(1-\delta)K_1}\right)^{(p+1)/2} < \infty.$$

Student distribution. From (12) we have $1/C_1^* \geq e^{-\tau\epsilon_1}/B_1$. Under a Student distribution, the right-hand side has no finite expectation, and hence the left-hand side has no finite expectation either. In the non-linear scrap case, this is sufficient to prove the non-existence of $\mathrm{E}(W^*)$ because $S^{(1)}(K_2^*)$ and $S^{(2)}(M_2^*)$ are both bounded. In the linear scrap case, M_2^* is bounded, but K_2^* is not. Now, since

$$C_1^* \leq B_1 e^{\tau\epsilon_1}, \qquad K_2^* \leq (1-\delta)K_1 + B_1 e^{\tau\epsilon_1},$$

we obtain

$$L_1(1-L_1/C_1^*)+\nu_1 K_2^* \leq L_1-(L_1^2/B_1)\,e^{-\tau\epsilon_1}+\nu_1(1-\delta)K_1+\nu_1 B_1 e^{\tau\epsilon_1} \equiv G(\epsilon_1).$$

Since G is monotonically increasing from $-\infty$ to $+\infty$, there exists a unique ϵ_1^* defined by $G(\epsilon_1^*)=0$. Hence, $G(\epsilon_1)\leq 0$ for all $\epsilon_1 \leq \epsilon_1^*$ and

$$\begin{aligned}\mathrm{E}\,|(L_1(1-L_1/C_1^*)+\nu_1 K_2^*)| &\geq \int_{\epsilon_1\leq\epsilon_1^*} |G(\epsilon_1)|\,dF(\epsilon_1)\\ &\geq -L_1-\nu_1(1-\delta)K_1+(L_1^2/B_1)\int_{\epsilon_1\leq\epsilon_1^*} e^{-\tau\epsilon_1}\,dF(\epsilon_1)-\nu_1 B_1 e^{\tau\epsilon_1^*}=\infty.\end{aligned}$$

C Expected utility and tail uncertainty

We now formulate our decision under uncertainty problem in Savage (1954) style, independent of the specific model considered in this paper, so that the results obtained below are generally applicable. We fix a set $\mathcal{S}$ of states of nature and we let $\mathcal{A}$ denote a σ-algebra of subsets of $\mathcal{S}$. One state is the true state. We also fix a set $\mathcal{C}$ of consequences (outcomes, consumption) endowed with a σ-algebra $\mathcal{F}$. Since we are only interested in monetary outcomes, we may take $\mathcal{C}=\mathbb{R}_+$. A decision alternative (policy bundle) X is a measurable mapping from $\mathcal{S}$ to $\mathcal{C}$, so that $X^{-1}(A)\in\mathcal{A}$ for all events $A\in\mathcal{F}$. We assume that the class of all decision alternatives $\mathcal{X}$ is endowed with a preference order $\succeq$.

Definition **C.1** *We say that expected utility (EU) holds if there exists a measurable and strictly increasing function $U:\mathcal{C}\to\mathbb{R}$ on the space of consequences, referred to as the utility function, and a probability measure $\mathbb{P}$ on $\mathcal{A}$, such that the preference order $\succeq$ on $\mathcal{X}$ is represented by a functional V of the form $X\mapsto\int_{\mathcal{S}} U(X(s))\,d\mathbb{P}=V(X)$. Thus, the decision alternative $X\in\mathcal{X}$ is preferred to the decision alternative $Y\in\mathcal{X}$ if, and only if, $V(X)\geq V(Y)$.*

In the Von Neumann and Morgenstern (1944) framework, utility U is subjective, whereas the probability measure $\mathbb{P}$ associated with $\mathcal{A}$ is objective and known beforehand (decision under risk). In the more general framework of Savage (1954) adopted here, the probability measure itself can be, but need not be, subjective (decision under uncertainty).

Definition **C.2** *We say that a risk $\epsilon:\mathcal{S}\to\mathbb{R}$ is heavy-tailed to the left (right) under $\mathbb{P}$ if its moment-generating function is infinite: $\mathrm{E}\,(e^{\gamma\epsilon})=\infty$ for any $\gamma<0$ ($\gamma>0$).*

Examples of heavy-tailed risks are the Student, lognormal, and Pareto distributions. Heavy-tailed risks provide appropriate mathematical models for low-probability high-impact events, such as environmental catastrophes.

PROPOSITION **C.1** *If EU is to discriminate univocally among all possible alternative outcome distributions, the utility function must be bounded.*

Proof: See Menger (1934, p. 468) in the context of St. Petersburg-type lotteries, and also Arrow (1974) and Gilboa (2009, pp. 108-109). Menger (implicitly) assumes boundedness from below and demonstrates that boundedness from above should hold, and it is straightforward to generalize his result to an a priori unrestricted setting. ‖

Proposition C.1 states that the EU functional is finite for all outcome distributions if, and only if, the utility function is bounded. Moreover, the axiomatization of EU is valid for all outcome distributions if, and only if, the utility function is bounded. The implications are non-trivial: boundedness of the utility function must hold not just in exotic situations but also in more familiar and economically relevant settings involving high levels of uncertainty. (See Moscadelli (2004) regarding operational risk.)

In what follows we do not require the utility function to be bounded. We simply assume that the class of feasible outcome distributions is restricted (though the restriction may be void) in such a way that the utility function permits discrimination among them. Recall from (5) that $\mathrm{RRA}(x) = -xU''(x)/U'(x)$ and $\mathrm{ARA}(x) = -U''(x)/U'(x)$, and let

$$\alpha^* = \inf_{x>0} \mathrm{RRA}(x), \qquad \beta^* = \sup_{x>0} \mathrm{ARA}(x).$$

Now consider a representative agent with time-additive EU preferences and time-preference parameter $\rho > 0$. We normalize (without loss of generality) the agent's consumption C by setting $C_0 = 1$, and we define the pricing kernel (intertemporal marginal rate of substitution) as

$$P(C_1^*) = \frac{U'(C_1^*)}{(1+\rho)U'(1)}, \tag{13}$$

where C_1^* is optimal consumption at $t = 1$. Consumption C_1 is commonly restricted to a budget-feasible consumption set which is subject to uncertainty (ϵ_1). We assume that the budget restriction takes the general form

$$C_1^*(\epsilon_1) \leq B \exp(A\epsilon_1), \qquad B, A > 0, \tag{14}$$

which need not be best-possible. (In our economy-climate model of Section 2 as well as in the two-period setup of Section 3, $B = B_1$ and $A = \tau$.) The expectation $\mathrm{E}(P)$ represents the amount of consumption in period 0 that the representative agent is willing to give up in order to obtain one additional certain unit of consumption in period 1.

The following result states that the expectation of the pricing kernel is finite for all outcome distributions whenever the concavity index (Arrow-Pratt index, index of absolute risk aversion) $\mathrm{ARA}(x)$ is bounded.

Proposition **C.2** *Assume that EU holds and that the budget feasibility restriction (14) applies.*
(a) If $\alpha^* > 0$ *and* ϵ_1 *is heavy-tailed to the left under* $\mathbb{P}$*, then* $\mathrm{E}(P) = \infty$*;*
(b) If $\beta^* < \infty$ *and* $\alpha^* = 0$*, then* $\mathrm{E}(P) < \infty$ *for any* ϵ_1*.*

Proof: Let $\alpha^* > 0$. The EU maximizer is then more risk-averse in the sense of Arrow-Pratt than an agent with power (CRRA) utility of index α^*. It follows from (13) that

$$\frac{P'(C_1^*)}{P(C_1^*)} = \frac{U''(C_1^*)}{U'(C_1^*)} = -\mathrm{ARA}(C_1^*).$$

Since $\mathrm{ARA}(x) = \mathrm{RRA}(x)/x \geq \alpha^*/x$, we then have

$$\begin{aligned}\mathrm{E}(P) &= \frac{1}{1+\rho}\,\mathrm{E}\exp\left(-\int_{C_1^*}^{1} d\log P(x)\right) = \frac{1}{1+\rho}\,\mathrm{E}\exp\left(\int_{C_1^*}^{1} \mathrm{ARA}(x)\,dx\right)\\ &\geq \frac{1}{1+\rho}\int_{C_1^*\leq 1} \exp\left(\int_{C_1^*}^{1}(\alpha^*/x)\,dx\right)\,dF(\epsilon_1)\\ &= \frac{1}{1+\rho}\int_{C_1^*\leq 1}(C_1^*)^{-\alpha^*}\,dF(\epsilon_1) \geq \frac{B_1^{-\alpha^*}}{1+\rho}\int_{C_1^*\leq 1} e^{-\tau\alpha^*\epsilon_1}\,dF(\epsilon_1) = \infty,\end{aligned}$$

using (14) and the fact that ϵ_1 is heavy-tailed to the left. This proves part (a). Intuitively, if agent 1 is more risk-averse in the sense of Arrow-Pratt than agent 2, and if it is optimal to postpone all consumption for agent 2, then this will also be optimal for agent 1.

Next let $\alpha^* = 0$ and $\beta^* < \infty$. The EU maximizer is then less risk-averse in the sense of Arrow-Pratt than an agent with exponential (CARA) utility of index β^*. Since $\alpha^* = 0$, we have $0 \le \mathrm{ARA}(x) \le \beta^*$ and hence

$$\begin{aligned}
\mathrm{E}(P) &= \int_{C_1^* \le 1} P \, dF(\epsilon_1) + \int_{C_1^* > 1} P \, dF(\epsilon_1) \\
&\le \frac{1}{1+\rho} \int_{C_1^* \le 1} \exp\left(\int_{C_1^*}^{1} \beta^* \, dx \right) dF(\epsilon_1) \\
&\quad + \frac{1}{1+\rho} \int_{C_1^* > 1} \exp\left(- \int_{1}^{C_1^*} \mathrm{ARA}(x) \, dx \right) dF(\epsilon_1) \\
&\le \frac{e^{\beta^*} \Pr(C_1^* \le 1) + \Pr(C_1^* > 1)}{1+\rho} < \infty. \quad \|
\end{aligned}$$

If the EU maximizer has decreasing absolute risk aversion and increasing relative risk aversion, as is commonly assumed, a complete and elegant characterization of boundedness of the expected pricing kernel can be obtained, as follows.

Proposition **C.3** *Assume that EU holds and that the budget feasibility restriction (14) applies. Assume furthermore that RRA(x) exists and is non-negative and non-decreasing for all $x \ge 0$ and that ARA(x) is non-increasing for all $x > 0$. Then,* $\mathrm{E}(P) < \infty$ *for any* ϵ_1 *if and only if* $\int_0^\gamma ARA(x) \, dx < \infty$ *for some* $\gamma > 0$.

Proof: To prove the 'only if' part, we assume that $\int_0^\gamma \mathrm{ARA}(x) \, dx$ is infinite for every $\gamma > 0$, and then show that there exist $(\mathcal{S}, \mathcal{A}, \mathbb{P})$ and ϵ_1 defined on it such that $\mathrm{E}(P) = \infty$. We note that $\beta^* = \infty$. Define a function $g : (0, 1] \to [1, \infty)$ by

$$g(y) = \exp\left(\int_y^1 \mathrm{ARA}(x) \, dx \right).$$

Then,

$$\mathrm{E}(P) \ge \frac{1}{1+\rho} \int_{C_1^* \le 1} g(\min(C_1^*, 1)) \, dF(\epsilon_1).$$

Recall from (14) that $C_1^* \le B_1 e^{\tau \epsilon_1}$, and let ϵ_1^* be such that $B_1 e^{\tau \epsilon_1^*} = 1$, so that $0 < B_1 e^{\tau \epsilon_1^*} \le 1$ if and only if $\epsilon_1 \le \epsilon_1^*$. Define $u : (-\infty, \infty) \to [0, \infty)$ by

$$u(\epsilon_1) = \begin{cases} g(B_1 e^{\tau \epsilon_1}) - 1 & \text{if } \epsilon_1 \le \epsilon_1^*, \\ 0 & \text{if } \epsilon_1 > \epsilon_1^*. \end{cases}$$

Since $\text{ARA}(1) > 0$, g is monotonically decreasing and we obtain

$$\int_{C_1^* \leq 1} g(\min(C_1^*, 1))\, dF(\epsilon_1) \geq \int_{\epsilon_1 \leq \epsilon_1^*} g(B_1 e^{\tau \epsilon_1})\, dF(\epsilon_1)$$
$$= \int_{\epsilon_1 \leq \epsilon_1^*} (u+1)\, dF(\epsilon_1) = \text{E}(u) + \Pr(\epsilon_1 \leq \epsilon_1^*).$$

Strict monotonicity of g implies its invertibility. Hence we can choose u to be any non-negative random variable whose expectation does not exist (for example, the absolute value of a Cauchy distribution), and then define ϵ_1 through $B_1 e^{\tau \epsilon_1} = g^{-1}(u+1)$. With such a choice of ϵ_1 we have $\text{E}(P) = \infty$.

To prove the 'if'-part we assume that $\int_0^\gamma \text{ARA}(x)\, dx$ is finite. This implies that $\int_0^1 \text{ARA}(x)\, dx$ is finite, so that

$$\begin{aligned} \text{E}(P) &= \frac{1}{1+\rho} \int_{C_1^* \leq 1} \exp\left(\int_{C_1^*}^1 \text{ARA}(x)\, dx \right) dF(\epsilon_1) \\ &\quad + \frac{1}{1+\rho} \int_{C_1^* > 1} \exp\left(- \int_1^{C_1^*} \text{ARA}(x)\, dx \right) dF(\epsilon_1) \\ &\leq \frac{\Pr(C_1^* \leq 1)}{1+\rho} \exp\left(\int_0^1 \text{ARA}(x)\, dx \right) + \frac{\Pr(C_1^* > 1)}{1+\rho} < \infty, \end{aligned}$$

using the fact that $\alpha^* = \text{RRA}(0) = 0$. $\|$

Notice that, when $\int_0^\gamma \text{ARA}(x)\, dx = \infty$ for some $\gamma > 0$, both $\alpha^* > 0$ and $\alpha^* = 0$ can hold. If $\alpha^* > 0$ then we do not need the full force of Proposition C.3; it is sufficient that ϵ_1 is heavy-tailed to the left. Then $\text{E}(P) = \infty$ by Proposition C.2(a). If $\alpha^* = 0$ then heavy-tailedness alone is not sufficient, but we can always find an ϵ_1 such that $\text{E}(P) = \infty$. An example of an ARA satisfying $\int_0^\gamma \text{ARA}(x)\, dx = \infty$ and $\alpha^* = 0$ is a function which behaves as $-1/(x \log x)$ for values of x close to 0 and in addition satisfies the conditions of the proposition.

When $\int_0^\gamma \text{ARA}(x)\, dx = \infty$ then $\beta^* = \infty$. But when $\int_0^\gamma \text{ARA}(x)\, dx < \infty$, both $\beta^* < \infty$ and $\beta^* = \infty$ can occur. For example, when $\text{ARA}(x) = x^{-\delta}$ $(0 < \delta < 1)$ then $\beta^* = \infty$; but when $\text{ARA}(x) = \beta$ $(0 \leq \beta < \infty)$ then β^* is finite. A sufficient condition for $\int_0^\gamma \text{ARA}(x)\, dx < \infty$ to hold is that there exists $0 \leq \delta < 1$ such that $\limsup_{x \downarrow 0} x^\delta \text{ARA}(x) < \infty$.

References

Ackerman, F., E.A. Stanton, and R. Bueno (2010). Fat tails, exponents, extreme uncertainty: Simulating catastrophe in DICE, *Ecological Eco-*

nomics, 69, 1657–1665.

Arrow, K.J. (1974). The use of unbounded utility functions in expected-utility maximization: Response, *Quarterly Journal of Economics*, 88, 136–138.

Barro, R.J. (2009). Rare disasters, asset prices, and welfare costs, *American Economic Review*, 99, 243-264.

Barro, R.J. and J.F. Ursúa (2008). Consumption disasters in the twentieth century, *American Economic Review*, 98, 58–63.

Broome, J.R. (1991). *Weighing Goods*, Basil Blackwell, Oxford, UK.

Burr, I.W. (1942). Cumulative frequency functions, *Annals of Mathematical Statistics*, 13, 215–232.

Burr, I.W. and P.J. Cislak (1968). On a general system of distributions: I Its curve-shape characteristics, II The sample median, *Journal of the American Statistical Association*, 63, 627–635.

Chichilnisky, G. (2000). An axiomatic approach to choice under uncertainty with catastrophic risks, *Resource and Energy Economics*, 22, 221–231.

Dhami, S. and A. Al-Nowaihi (2010). Optimal taxation in the presence of tax evasion: Expected utility versus prospect theory, *Journal of Economic Behavior and Organization*, 75, 313–337.

Fishburn, P.C. (1976). Unbounded utility functions in expected utility theory, *Quarterly Journal of Economics*, 90, 163–168.

Gerber, H.U. (1979). *An Introduction to Mathematical Risk Theory*, S.S. Huebner Foundation Monograph 8, Irwin, Homewood, IL.

Geweke, J. (2001). A note on some limitations of CRRA utility, *Economics Letters*, 71, 341–345.

Gilboa, I. (2009). *Theory of Decision under Uncertainty*, Cambridge University Press, Cambridge, UK.

Gollier, C. (2002). Time horizon and the discount rate, *Journal of Economic Theory*, 107, 463–473.

Gollier, C. (2008). Discounting with fat-tailed economic growth, *Journal of Risk and Uncertainty*, 37, 171-186.

Ikefuji, M., R.J.A. Laeven, J.R. Magnus, and C. Muris (2010a). Scrap value functions in dynamic decision problems, CentER Discussion Paper No. 2010-77, Tilburg University.

Ikefuji, M., R.J.A. Laeven, J.R. Magnus, and C. Muris (2010b). Burr utility, CentER Discussion Paper No. 2010-81, Tilburg University.

Ikefuji, M., R.J.A. Laeven, J.R. Magnus, and C. Muris (2010c). Catastrophe, VSL, and curvature, Technical Notes, Tilburg University.

IPCC (2007). Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller (eds.), *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, UK and New York, NY.

Keller, K., B.M. Bolker, and D.F. Bradford (2004). Uncertain climate thresholds and optimal economic growth, *Journal of Environmental Economics and Management*, 48, 723–741.

Kelly, D.L. and C.D. Kolstad (1999). Bayesian learning, growth, and pollution, *Journal of Economic Dynamics and Control*, 23, 491–518.

Leach, A.J. (2007). The climate change learning curve, *Journal of Economic Dynamics and Control*, 31, 1728–1752.

Manne, A.S. and R.G. Richels (1992). *Buying Greenhouse Insurance: The Economic Costs of CO2 Emission Limits*, MIT Press, Cambridge, MA.

Mastrandrea, M.D. and S.H. Schneider (2004). Probabilistic integrated assessment of 'dangerous' climate change, *Science*, 304, 571–575.

Menger, K. (1934). Das Unsicherheitsmoment in der Wertlehre, *Zeitschrift für Nationalökonomie*, 5, 459–485.

Moscadelli, M. (2004). The modelling of operational risk: Experience with the analysis of the data collected by the Basel Committee, Technical Report 517, Banca d'Italia.

Nordhaus, W.D. (1994). *Managing the Global Commons: The Economics of Climate Change*, MIT Press, Cambridge, MA.

Nordhaus W.D. (2008). *A Question of Balance: Weighing the Options on Global Warming Policies*, *The Challenge of Global Warming: Economic Models and Environmental Policy*, Yale University Press, New Haven, CT.

Nordhaus, W.D. and Z. Yang (1996). A regional dynamic general-equilibrium model of alternative climate-change strategies, *American Economic Review*, 86, 741–765.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem, *Econometrica*, 68, 1281–1292.

Roughgarden, T. and S.H. Schneider (1999). Climate change policy: Quantifying uncertainties for damages and optimal carbon taxes, *Energy Policy*, 27, 415–429.

Ryan, T.M. (1974). The use of unbounded utility functions in expected-utility maximization: Comment, *Quarterly Journal of Economics*, 88, 133–135.

Savage, L.J. (1954). *The Foundations of Statistics*, Wiley, New York, NY.

Sims, C.A. (2001). Pitfalls of a minimax approach to model uncertainty, *American Economic Review*, 91, 51–54.

Stern, N. (2007). *The Economics of Climate Change: The Stern Review*, Cambridge University Press, Cambridge, UK.

Sugden, R. (1997). Alternatives to expected utility theory: Foundations and concepts, in: *Handbook of Expected Utility Theory*, Eds. S. Barberà, P.J. Hammond and C. Seidl, Kluwer, Boston, MA.

Von Neumann, J. and O. Morgenstern (1944, 1947, 1953). *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ.

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