Economics meets statistics:
Expected utility and catastrophic risk*

Masako Ikefuji
Department of Environmental and Business Economics
University of Southern Denmark

Roger J. A. Laeven
Amsterdam School of Economics
University of Amsterdam, EURANDOM and CentER

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

Chris Muris
Department of Economics
Simon Fraser University

July 24, 2013

*We are very grateful to Graciela Chichilnisky, John Einmahl, Reyer Gerlagh, Christian Groth, Johan Eyckmans, Sjak Smulders, Peter Wakker, Aart de Zeeuw, Amos Zemel, and seminar and conference participants at Kyushu University, Tilburg University, the University of Copenhagen, EAERE, and the Tilburg Sustainability Center for helpful comments. This research was funded in part by the Japan Society for the Promotion of Science (JSPS) under grant C-22530177 (Ikefuji) and by the Netherlands Organization for Scientific Research (NWO) under grant Vidi-2009 (Laeven). E-mail addresses: ikefuji@sam.sdu.dk (Ikefuji), R.J.A.Laeven@uva.nl (Laeven), magnus@uvt.nl (Magnus), cmuris@sfu.ca (Muris).

Corresponding author:
Roger J. A. Laeven
Amsterdam School of Economics
University of Amsterdam
The Netherlands
Phone: +31 20 525 4219
e-mail: `R.J.A.Laeven@uva.nl`

Abstract: We derive necessary and sufficient conditions on the utility function of the expected utility model to avoid fragility of a cost-benefit analysis to distributional assumptions. The conditions ensure that expected (marginal) utility remains finite also under heavy-tailed distributional assumptions. We apply these conditions to the context of economy-climate catastrophe. We specify a stylized two-period stochastic economy-climate model. We show that, under expected power utility, the model is fragile to heavy-tailed distributional assumptions and, based on our derived conditions, we solve the model for two cases with compatible economic and statistical assumptions.

JEL Classification: D61; D81; Q5.

Keywords: Expected utility; Catastrophe; Cost-benefit analysis; Economy-climate models; Economy-climate policy; Power utility; Heavy tails.

1 Introduction

An economist, when asked to model decision making under risk or 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 power 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. 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. Related questions about the validity of expected utility analysis in a catastrophic climate change context were analyzed by Chichilnisky (2000) and Tol (2003), and, more recently, by Horowitz and Lange (2009), Karp (2009), Arrow (2009), Nordhaus (2009, 2011), Pindyck (2011), Buchholz and Schymura (2012), Chanel and Chichilnisky (2013) and Hwang, Reynès and Tol (2013).

In this paper we contribute to this literature on the question of how to conduct expected utility analysis in the presence of catastrophic risks by deriving general theoretical compatibility results on the utility function of the expected utility model, leaving probability distributions unrestricted, and we illustrate these results in a stylized two-period economy-climate model with catastrophic risk. Our theoretical results are context-free. But we have the context of catastrophe in an economy-climate model in mind. 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, e.g., financial distress, traffic accidents, 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.

A good model 'in the center' is not necessarily good 'at the edges'. Models are approximations, not truths, and approximations may not work well if we move too far away from the point of approximation. In our context of catastrophe in an economy-climate model, the widely adopted family of power utility functions, often appropriate when one considers large inputs remote from zero, may not work well for decision making under heavy-tailed risks with non-negligible support beyond the usual domain of inputs.

The price to reduce catastrophic risk is finite. Are we willing to spend everything to avoid children being killed at a dangerous street? Or 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 of its current wealth to avoid or mitigate catastrophic risks) is not credible, not even from a normative 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 an economy-climate catastrophe 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.

We start our analysis by deriving necessary and sufficient conditions on the utility function of the expected utility model to avoid fragility of a cost-benefit analysis to distributional assumptions. The conditions we derive ensure that expected utility and expected marginal utility remain finite also under heavy-tailed distributional assumptions, and are context-independent, hence of independent interest. They guarantee a valid axiomatization of expected utility and avoid incredible consequences in a cost-benefit analysis. Our analysis exploits the budget restriction of the optimal consumption problem to derive as weak conditions as possible on the utility function.

Next, we apply our general results to the particular setting of economy-climate catastrophe. To allow for catastrophic risk, we specify a stylized stochastic economy-climate model, adapting the canonical Nordhaus' (2008) deterministic dynamic integrated climate and economy (DICE) model by allowing for simple stochasticity, in the spirit of Weitzman (2009). Next, we solve a two-period version of the model, first with power utility and light-tailed distributional assumptions. Since the assumption of expected power utility is incompatible with heavy-tailed distributional assumptions, we then restrict attention to utility functions that satisfy the derived compatibility conditions, and solve our stochastic economy-climate model with the well-known exponential utility function and also with the less well-known (but more suitable) 'Pareto' utility function, under both light- and heavy-tailed distributional assumptions.

The paper is organized as follows. Section 2 studies expected (marginal) utility and catastrophic risk in a general setting, deriving results on the trade-off between permitted distributional assumptions and the existence of expected (marginal) utility, which are of interest in their own right. In Section 3 we propose a simplified version of Nordhaus' economy-climate model, introduce uncertainty in the spirit of Weitzman to obtain a stylized stochastic integrated assessment model of climate economics, and specialize the model to two periods only. In Section 4, we present (partial) results for power utility, which is incompatible with heavy tails, and for exponential and Pareto utility, which are compatible with heavy tails. Section 5 concludes. There are two appendices: Appendix A provides Kuhn-Tucker conditions and Appendix B contains proofs of the propositions.

2 Expected utility and catastrophic risk

We formulate our cost-benefit analysis as a decision under uncertainty problem, in Savage (1954) style. 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 **2.1** *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, $V(X) \geq V(Y)$.*[1]

We henceforth assume that U is defined for $x \geq 0$, twice differentiable, and such that $U'(x) > 0$ and $U''(x) < 0$ for $x > 0$.

Since the axiomatization of EU by Von Neumann and Morgenstern (1944) and Savage (1954), numerous objections have been raised against it. These objections relate primarily to empirical evidence that the behavior of agents under risk and uncertainty does not agree with EU. Despite important developments in non-expected utility theory, EU remains the dominant

[1] 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).

normative decision theory (Broome, 1991; Sims, 2001), and the current paper stays within the framework of EU. Our results presented below in a sense provide compatibility conditions under which expected utility theory may reliably provide normatively appealing results, also in the presence of catastrophic risks. Of course, 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 e.g., Chichilnisky (2000). This is beyond our scope.

Definition **2.2** *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:* $\mathrm{E}\left(e^{\gamma\epsilon}\right) = \infty$ *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 or financial catastrophes. We state the following result, which dates back to Menger (1934).

Proposition **2.1** *If EU is to discriminate univocally among all possible alternative outcome distributions, the utility function must be bounded.*

Proposition 2.1 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. Only a combination of utility function and outcome distribution that leads to finite expected utility is covered by the axiomatic justification of EU.

Now consider a representative agent with time-additive EU preferences and time-preference parameter $\rho > 0$. 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

$$C_1^*(\epsilon_1) \leq B \exp(A\epsilon_1), \qquad B, A > 0, \tag{1}$$

which need not be best-possible. Here C_1^* is optimal consumption at $t = 1$.[2] We exploit (1) to derive compatibility conditions on the utility function.

We normalize (without loss of generality) the agent's consumption by setting $C_0 = 1$, and we define the pricing kernel (intertemporal marginal rate of substitution) as

$$P(C_1^*) = \frac{U'(C_1^*)}{(1+\rho)U'(1)}. \tag{2}$$

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.

Let $\mathrm{RRA}(x) = -xU''(x)/U'(x)$ and $\mathrm{ARA}(x) = -U''(x)/U'(x)$ denote relative and absolute risk aversion, respectively, and let

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

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

PROPOSITION **2.2** *Assume that EU holds and that the budget feasibility restriction (1) 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.

If the EU maximizer has decreasing absolute risk aversion and increasing relative risk aversion, as is commonly assumed (Eeckhoudt and Gollier, 1995, Section 4.2, Hypotheses 4.1 and 4.2), a complete and elegant characterization of boundedness of the expected pricing kernel can be obtained, as follows.

PROPOSITION **2.3** *Assume that EU holds and that the budget feasibility restriction (1) applies. Assume furthermore that $RRA(x)$ exists and is non-negative and non-decreasing for all $x \geq 0$ 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 $\gamma > 0$.*

[2] In our economy-climate model of Section 3.1, and in the two-period setup of Section 3.3, $B = e^{-\tau^2/2} Y_1/(1 + \xi H_1^2)$ and $A = \tau$.

REMARK **2.1** *Notice that, when* $\int_0^\gamma ARA(x)\,dx = \infty$ *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 2.3; it is sufficient that* ϵ_1 *is heavy-tailed to the left. Then* $\mathrm{E}(P) = \infty$ *by Proposition 2.2(a). If* $\alpha^* = 0$ *then heavy-tailedness alone is not sufficient, but we can always find an* ϵ_1 *such that* $\mathrm{E}(P) = \infty$*. When* $\int_0^\gamma ARA(x)\,dx = \infty$ *then* $\beta^* = \infty$*. But when* $\int_0^\gamma ARA(x)\,dx < \infty$*, both* $\beta^* < \infty$ *and* $\beta^* = \infty$ *can occur.*

REMARK **2.2** *An example of an ARA satisfying* $\int_0^\gamma ARA(x)\,dx = \infty$ *and* $\alpha^* > 0$ *is that of power utility. An example of an ARA satisfying* $\int_0^\gamma ARA(x)\,dx = \infty$ *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.*

An example of an ARA satisfying $\int_0^\gamma ARA(x)\,dx < \infty$ *and* $\beta^* = \infty$ *occurs when* $ARA(x) = x^{-\delta}$ $(0 < \delta < 1)$*. An example of an ARA satisfying* $\int_0^\gamma ARA(x)\,dx < \infty$ *and* $\beta^* < \infty$ *occurs in the case of exponential utility in which* $ARA(x) = \beta$ $(0 \leq \beta < \infty)$*. A sufficient condition for* $\int_0^\gamma ARA(x)\,dx < \infty$ *to hold is that there exists* $0 \leq \delta < 1$ *such that* $\limsup_{x \downarrow 0} x^\delta ARA(x) < \infty$*.*

The above propositions provide necessary and sufficient conditions on the utility function to ensure that expected utility and expected marginal utility (hence the expected pricing kernel) are finite, also in the presence of heavy tails. These compatibility results are generally applicable to standard multi-period welfare maximization problems. The importance of the results lies in the fact that (i) if (minus) expected utility is infinite, the axiomatic justification of EU is not valid, and (ii) 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, which is not credible in most settings.

3 Economy-climate catastrophe

The results of the previous section can be applied to many situations involving catastrophic risks. We choose economy-climate catastrophe as our illustration. In this section we first present a simplified deterministic Nordhaus (2008)-type economy-climate model; then introduce simple stochasticity, in the spirit of Weitzman (2009), but in a sufficiently general manner allowing application in other contexts as well; and finally specialize the infinite-horizon model to a two-period model.

Weitzman (2009) recently noted, in a highly stylized setting of extreme climate change, that heavy-tailed uncertainty and power utility are incompatible, since this combination of uncertainty and preferences implies an infinite expected pricing kernel. To avoid this, Weitzman introduces a lower bound on consumption, argues that this lower bound is related to a parameter that resembles the value of a statistical life (VSL), and proves that the expected 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.

Incompatible pairs of utility functions and distribution functions indeed exist, in the sense that the expected pricing kernel or other important policy variables become infinite. In fact, Section 2 presents necessary and sufficient conditions on the utility functions for the expected pricing kernel to exist, also under heavy tails. But we object to the dismal theorem for the following reason. As we proved formally in Section 2 and shall illustrate numerically in Section 4, the dismal theorem is based on an ex ante incompatible (invalid) model specification. It is avoided when the economic model (utility function) is compatible with the statistical model (heavy tails). Note that only then Savage's axiomatization of EU is valid.

3.1 A simple deterministic economy-climate model

Our framework is a simple economy-climate model in the spirit of Nordhaus and Yang (1996) and Nordhaus (2008). It retains all essential features of Nordhaus (2008) but is numerically easier to handle, especially under uncertainty.

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

$$(1-\omega_t)D_t Y_t = C_t + I_t. \tag{3}$$

The damage function D_t depends only on temperature and satisfies $0 < D_t \leq 1$, where $D_t = 1$ represents the optimal temperature for the economy. Deviations from the optimal temperature cause damage. We specify D_t as

$$D_t = \frac{1}{1+\xi H_t^2} \qquad (\xi > 0).$$

For very high and very low temperatures D_t approaches zero. The optimal value $D_t = 1$ occurs at $H_t = 0$, the temperature in 1900, as in Nordhaus. A fraction ω_t of $D_t Y_t$ is spent on abatement, and we specify the abatement cost fraction as

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

When μ_t increases then so does ω_t, and a larger fraction of GDP will be spent on abatement. 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

$$C_t \geq 0, \quad I_t \geq 0, \quad 0 \leq \mu_t \leq 1. \tag{4}$$

In Appendix A we prove that $\mu_1 \geq 0$ and $C_1 \geq 0$ are automatically satisfied. The other restrictions need to be imposed. Then, all variables will have the correct signs and all fractions will lie between zero and one.

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

$$W_t = L_t U(C_t/L_t).$$

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

$$U(x) = \frac{x^{1-\alpha} - 1}{1 - \alpha} \qquad (\alpha > 0), \tag{5}$$

where α denotes the elasticity of marginal utility of consumption. This is the so-called *power* function. Many authors, including Nordhaus (2008), select this function and choose $\alpha = 2$ in which case $U(x) = 1 - 1/x$. Also popular is $\alpha = 1$; see Kelly and Kolstad (1999) and Stern (2007).

Table 1: 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

Calibrating the parameters and initial values (presented in our background document, see Ikefuji *et al.*, 2013b), our simple model[3] (hereafter,

[3]GAMS code available at http://www.janmagnus.nl/items/risk.pdf

SICE = simplified DICE) produces optimal values over sixty periods that are very close to the values obtained in Nordhaus, as shown in Table 1.

3.2 Stochasticity

We now introduce simple uncertainty in the SICE model, in the spirit of Weitzman (2009), thus obtaining a stylized stochastic integrated assessment model of climate economics. There is much uncertainty in the economics of climate change (Manne and Richels, 1992; Nordhaus, 1994; Roughgarden and Schneider, 1999; Kelly and Kolstad, 1999; Keller *et al.*, 2004; Mastrandrea and Schneider, 2004; Leach, 2007; Weitzman, 2009, Ackerman *et al.*, 2010). We model uncertainty through stochasticity. In the literature, stochasticity is typically introduced through the damage function (see, for example, Cai, Judd and Lontzek, 2012). We follow this literature by introducing stochasticity through Equation (3), which we now write as

$$f_t Y_t = C_t + I_t, \tag{6}$$

where f_t depends not only on ω_t and H_t (as in (3)), but also on a random shock ϵ_t. In particular, we specify

$$f_t = (1-\omega_t)\bar{d}_t D_t, \qquad \bar{d}_t = e^{-\tau^2/2}\, e^{\tau\epsilon_t}, \tag{7}$$

where ϵ_t denotes a random error with mean zero and variance one.

This specification should be interpreted as the reduced form resulting from various types of uncertainty, in particular damage and mitigation uncertainty. The potential damage due to adverse climate change is one component of the aggregate stochasticity affecting the economy, as in Weitzman (2009), and all stochasticity is dictated by the probability law of ϵ, which plays the role of $\log C$ in the reduced-form of Weitzman. We emphasize that extreme climate change is just one example of a catastrophe. Another example would be a financial crisis, where we could take f_t to depend on policy, financial institution, and risk.

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, more uncertainty (higher τ) implies more probability mass of $\bar{d}_t$ close to zero. If, however, we move only one step away from the normal distribution and assume e.g., that ϵ_t follows a Student distribution with *any* (finite) degrees of freedom, then the expectation is infinite (Geweke, 2001). 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 ($\bar{d}_t$, $\bar{d}_t Y_t$).

3.3 A two-period model

So far we have assumed an infinite horizon. We now specialize to two periods, as in Weitzman (2009). The two-period model captures the essence of our problem while remaining numerically tractable in the presence of uncertainty.

If the policy maker has a (finite) 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

$$S_T = \nu_0 + \nu_1 K_T - \nu_2 M_T \qquad (\nu_1 > 0, \nu_2 > 0), \tag{8}$$

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; see our background document Ikefuji *et al.* (2013b). We shall therefore introduce nonlinear scrap value functions, whose specific form depends on the form of the utility function; see Ikefuji *et al.* (2013b) for further details on our treatment of scrap value functions.

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 policy restrictions (4) are explicitly imposed, so that we maximize a restriction of expected welfare; see Appendix A. Randomness results from $\bar{d}_1$ only, because $\bar{d}_0$ at the beginning of period 0 is known to us (we set $\bar{d}_0 = 1$, equal to its expectation), and $\bar{d}_2$ 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.

Realizing that $\bar{d}_1$ at the beginning of period 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 $W = W(\mu_0, C_0, \mu_1, C_1, \epsilon_1)$ with respect to (μ_1, C_1) conditional on (μ_0, C_0, ϵ_1) and under the restriction (4). 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 $\overline{W}(\mu_0, C_0) = \mathrm{E}\left(W^*(\mu_0, C_0, \epsilon_1)\right)$ is computed, if it exists. Finally, $\overline{W}$ is maximized with respect to (μ_0, C_0).

4 Compatibility and solution

We now have a simplified Nordhaus model with Weitzman-type stochasticity in a two-period framework. We illustrate our theoretical compatibility results in this stylized economy-climate model with catastrophic risk.

4.1 Compatibility results

In this model, we consider three utility functions (power, exponential, Pareto) and two distributions (normal, Student). Power utility, as given in (5), takes the simple form $U(x) = 1 - 1/x$ for $\alpha = 2$. The following proposition states that if the random errors ϵ_t are generated by a normal $\mathrm{N}(0, 1)$ distribution, then the expectation of welfare exists for power utility, but 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. It illustrates the consequences of violating the conditions of Proposition 2.1.

PROPOSITION **4.1** *With power utility, expected welfare exists under normality of ϵ but not under a Student distribution.*

It follows that the much-used power utility function is incompatible 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.

Motivated by the conditions derived in Section 2 and by the fundamental insight that the economic model and the statistical model must be compatible, and also because we wish to leave distributional assumptions unrestricted at this stage, we consider two alternative utility functions: the exponential function and the Pareto 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 $\mathrm{ARA}(x) = \beta$ and $\mathrm{RRA}(x) = \beta x$, and the *Pareto* utility function by

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

with $\mathrm{ARA}(x) = (k+1)/(x+\lambda)$ and $\mathrm{RRA}(x) = (k+1)x/(x+\lambda)$. The Pareto function was proposed in Ikefuji *et al.* (2013a), where it is shown that this function enjoys a combination of appealing properties especially relevant in heavy-tailed risk analysis. We choose the parameters as follows: $\beta = 25$ in the exponential function, and $k = 1.5$ and $\lambda = 0.02$ in the Pareto function. This choice of parameters is determined by the point x^*, where we want the three utility 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 considered below). Then, given that $\alpha = 2$, we find $\beta = 2/x^* = 25$, and, for any $k > 1$, $\lambda = (k-1)x^*/2$.

The power function is unbounded, hence violates the conditions of Proposition 2.1, and has constant and positive RRA, hence violates the conditions of Proposition 2.3. Both the exponential and the Pareto function are bounded from above and below, hence satisfy the conditions of Proposition 2.1. The exponential function has constant and positive ARA, hence satisfies the conditions of Proposition 2.3, while the RRA is unbounded for large x. In contrast, the RRA in the Pareto function is bounded between 0 and $k+1$, and it further satisfies $\mathrm{RRA}(0) = 0$ and $\mathrm{ARA}(0) < \infty$, hence satisfies the conditions of Proposition 2.3. Notice that the fact that $\mathrm{RRA}(0) = 0$ (as is the case for the exponential and the Pareto utility functions) does *not* imply that the representative agent is risk-neutral at $x = 0$. In particular, we have $\mathrm{ARA}(0) = \beta$ for the exponential function and $\mathrm{ARA}(0) = (k+1)/\lambda$ for the Pareto function.

4.2 Numerical solution

In Table 2 we present the optimal values of the policy and other selected variables obtained from maximizing expected welfare. (Our background document contains the complete tables.) The results allow for uncertainty, consider the short run (two periods) rather than the long run (sixty periods), and also take scrap values into account.

We need values for the exogenous variables L_t, A_t, σ_t, and ψ_t; these are given in our background document. We note that $Y_0 = 556.67$ and $d_0 = 0.9985$ are constant over different scenarios and functions, and that

Table 2: Three utility functions with normal and Student(10) distributions

	Power			Exponential					Pareto				
		Normal			Normal		Student(10)			Normal		Student(10)	
τ	0.0	0.3	0.7	0.0	0.3	0.7	0.3	0.7	0.0	0.3	0.7	0.3	0.7
Policy instruments													
μ_0	0.0933	0.0920	0.0874	0.1175	0.1166	0.1135	0.1166	0.1135	0.0924	0.0910	0.0859	0.0910	0.0861
μ_1	0.1137	0.1141	0.1142	0.1473	0.1515	0.1697	0.1516	0.1700	0.1124	0.1135	0.1175	0.1135	0.1175
C_0	431.07	424.90	399.24	428.74	425.43	413.86	425.41	413.52	430.76	424.31	399.97	424.33	400.67
C_1	547.20	549.08	550.63	551.45	584.64	527.76	∞	∞	548.53	552.59	563.14	∞	∞
I_0	124.72	130.89	156.56	127.00	130.32	141.90	130.35	142.24	125.03	131.48	155.83	131.46	155.12
I_1	151.51	157.03	184.45	149.95	156.68	190.03	∞	∞	150.56	154.23	171.09	∞	∞
Stocks													
K_1	172.49	178.66	204.33	174.78	178.10	189.67	178.12	190.01	172.80	179.25	203.60	179.23	202.89
K_2	211.66	219.32	255.70	210.90	218.78	256.16	∞	∞	210.81	216.73	242.08	∞	∞
M_1	834.25	834.34	834.69	832.44	832.51	832.74	832.51	832.74	834.32	834.42	834.80	834.42	834.79
M_2	868.31	869.18	872.74	863.94	864.06	864.07	864.06	864.09	868.53	869.39	872.45	869.38	872.36
H_1	0.8843	0.8844	0.8848	0.8823	0.8824	0.8826	0.8824	0.8826	0.8844	0.8845	0.8850	0.8845	0.8849
H_2	1.0400	1.0411	1.0452	1.0337	1.0339	1.0341	1.0339	1.0341	1.0403	1.0413	1.0450	1.0413	1.0449
Probability of catastrophe													
π_b		5E−10	9E−03		6E−08	2E−02	7E−05	2E−02		5E−10	9E−03	2E−05	1E−02

the values of μ_0, C_0, I_0, E_0, ω_0, K_1, M_1, and H_1 are optimal 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.

We also need sensible values for the uncertainty parameter τ. The stochasticity, as given in (7), captures uncertainty about GDP that is due in part to uncertainty about climate economics. 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 addition to the value of τ we need to consider the question whether or not the uncertainty introduced is indeed heavy-tailed. A (partial) answer to this question is contained in a recent paper by Ursúa (2010) who claims that the growth rate of GDP indeed features heavy tails. In Figure 1 we plot the density

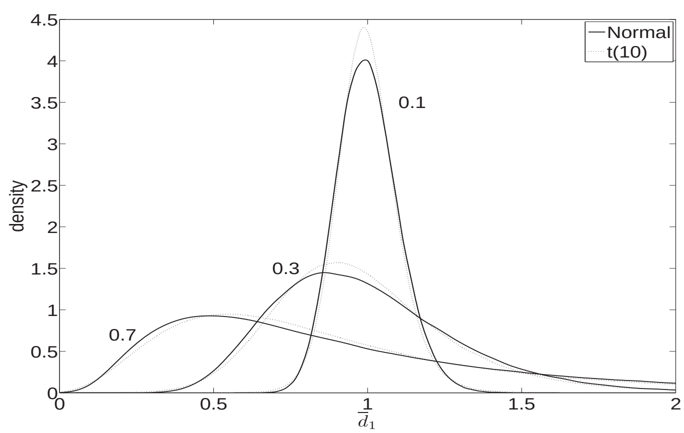

Figure 1: Density of $\bar{d}_t$ for $\tau = 0.1$, 0.3, and 0.7

of $\bar{d}_t$ for three values of τ: 0.1, 0.3, and 0.7, both when ϵ_t follows a $\mathrm{N}(0,1)$ distribution (solid line) and when $\epsilon_t = \sqrt{4/5}\,u$, where u follows a Student distribution (as adopted in Weitzman, 2009) with 10 degrees of freedom (which implies a 'tail index' that is broadly consistent with the empirical analysis in Ursúa, 2010). Notice that $\mathrm{E}(\epsilon_t) = 0$ and $\mathrm{var}(\epsilon_t) = 1$ in both cases. When $\tau = 0.1$, we see that almost 100% of the distribution of $\bar{d}_t$ 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, 0.5, and 0.7.

4.3 Light tails

The first panel of Table 2 gives the results for power utility. For $\tau = 0$ there is no uncertainty, but for $\tau > 0$ there is uncertainty, and the larger is τ the higher is the uncertainty. The increase in I_0 with τ can be explained by precautionary savings. The restriction on I_1 (cf. (4)) 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 at the expense of less abatement and consumption in period 0. 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.

Turning now to the alternative utility functions, we first maximize (deterministic, hence $\tau = 0$) welfare over sixty periods (600 years) for both exponential and Pareto utility. A selection of the resulting optimal values is shown in Table 3. When we compare the results with those in Table 1, we

Table 3: Comparison of stocks in Exponential and Pareto models

	2005		2055		2105		2155	
	Expo	Pareto	Expo	Pareto	Expo	Pareto	Expo	Pareto
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

see that the optimal stock values from the Pareto function closely resemble the optimal stock values from the power function, but not those from the exponential function. In contrast to power and Pareto, where RRA flattens out, the RRA for the exponential distribution continues to increase, 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. Consequently, M and H are high compared to power and Pareto. When $x < x^*$, RRA (Pareto) is close to RRA (exponential), so that more is consumed and less invested when the Pareto function is used instead of the power function. But when $x > x^*$, RRA (Pareto) 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 Pareto than for power utility.

Since exponential utility is calibrated to be close to power utility at $x = x^*$, the two-period results for the two utility functions do not differ greatly; see the second panel of Table 2. As the uncertainty parameter τ increases, M_2 does 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.

4.4 Heavy tails

Suppose next that the underlying distribution has heavier tails: Student instead of normal. Under power utility, expected welfare does not exist any more. But under bounded utility, expected welfare always exists. Although the effect of the excess kurtosis on expected welfare is large and discontinuous, the effect on the optimal values is relatively small *in the center of the distribution*. This is good, because the Student distribution with 10 degrees of freedom is in fact quite close to the normal distribution as Figure 1 reveals, and 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.

We would expect that Pareto and power are relatively close in the observed data range. This is indeed the case as a comparison of the first and third panels reveals. There is little difference between the two panels in the case of no uncertainty, and also when τ increases. The effect of excess kurtosis is again small, as it should be.

In the observed data range, that is in the center of the distribution, the three utility functions (power, exponential, Pareto) yield similar optimal policy values. Apparently the center of the distribution is quite robust with respect to the specification of the utility function and the error distribution. This is important, because discontinuities occur caused by the non-existence of moments. These discontinuities, however, do not cause large shocks in the center of the distribution. The small difference between power, exponential, and Pareto utility on the one hand, and the normal and Student distribution on the other within the observed data range does not mean that the choice between them does not matter in practice. The important differences between them are revealed when low levels of per capita consumption become relevant, that is, in near-catastrophe cases.

4.5 Near-catastrophe

To study near-catastrophe we must define what we mean by a catastrophic event. 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. These definitions are based on background material provided in Ikefuji *et al.* (2013b).

The last line in Table 2 gives the estimated values of π_b, the probability of type-B catastrophe. If we compare the probabilities of catastrophe of the power and exponential distribution for the normal distribution with $\tau = 0.3$, they differ by approximately a factor of 100. For exponential utility, moving from a normal distribution to a $t(10)$ distribution with $\tau = 0.3$ increases the probability of catastrophe by a factor of almost 1300. For Pareto utility, moving from a normal to a $t(10)$ distribution changes the probability of catastrophe even by a factor of 43000. We conclude that results at the mean are similar across models, which is in part a consequence of the manner in which the models are calibrated. But the large differences between the models, both in terms of distributional assumptions (normal versus Student) and in terms of utility function (power, exponential, Pareto), become clear once we consider the tails of the distribution.

Table 4: Pareto utility and Student distribution

	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

4.6 Robustness

To assess the sensitivity of the optimal policy to parameter changes, we have done extensive robustness checks and some representative results of this analysis are reported in columns *a*–*g* of Table 4, see Ikefuji *et al.* (2013b) for further results.

Let us take column *c*, with intermediate $\tau = 0.5$, as our benchmark. 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), a finding that is consistent with Chichilnisky (2000). In columns a, f we consider $\tau = 0.3,\ 0.7$. Here the probabilities of catastrophe become significantly smaller, larger. For example, we have $\pi_c = 0.0005$ in column f as opposed to $\pi_c = 0.00003$ in column c. The choice of volatility τ does affect the policy, and hence is important; cf. also columns b, d. In column g we adjust the curvature of the Pareto utility function. The probabilities are hardly affected but there will be more consumption, less investment, and in particular more (perhaps too much) abatement; cf. also column d. On the basis of these and other robustness checks we conclude that policy c is quite robust and sensibly sensitive to small changes in the underlying assumptions and parameter values.

5 Concluding remarks

We have derived necessary and sufficient conditions in the EU model to avoid fragility of the model to distributional assumptions in a cost-benefit analysis. We have applied these conditions to economy-climate catastrophe. Our strategy in this application has been to specify a stylized stochastic economy-climate model, building on Nordhaus' deterministic economy-climate model while allowing for simple stochasticity in the spirit of Weitzman. Under expected power utility, the model is shown to be fragile to distributional assumptions. Based on our generic results regarding the relationship between the richness of the class of utility functions and the generality of the permitted distributional assumptions, we have restricted ourselves to utility functions that are compatible with our distributional assumptions. Thus we guarantee a valid axiomatization of EU and avoid the unacceptable conclusion that society should sacrifice an unlimited amount of consumption to reduce the probability of an economy-climate catastrophe by even a small amount.

Much of the analysis in our paper is not limited to economy-climate catastrophe. Similar analyses 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, the application in particular, and indicate possible generalizations. First, in Section 4.1 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. Second, for simplicity and clarity of presentation, we have restricted our numerical analysis to only two periods. Conceptually, much of our analysis will remain intact when considering more than two periods, and, as such, it would be interesting to evaluate the alternative utility functions in the multi-short-period stochastic model of Cai, Judd, and Lontzek (2012). 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* stochasticity in scrap values after the last period is a subject for further 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

Consider the economy-climate model of Section 3.1 in a two-period set-up. 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 the proof of Proposition 4.1.

B Proofs of the propositions

Proof of Proposition 2.1: See Menger (1934, p. 468) in the context of St. Petersburg-type lotteries, and also Arrow (1974, p. 136) 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.

Proof of Proposition 2.2: 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 (2) 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

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

using (1) 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 $0 \leq \mathrm{ARA}(x) \leq \beta^*$ and hence

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

Proof of Proposition 2.3: 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,

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

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

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

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

$$\begin{aligned}
\int_{C_1^* \leq 1} g(\min(C_1^*, 1)) \, dF(\epsilon_1) &\geq \int_{\epsilon_1 \leq \epsilon_1^*} g(B e^{A\epsilon_1}) \, dF(\epsilon_1) \\
&= \int_{\epsilon_1 \leq \epsilon_1^*} (u+1) \, dF(\epsilon_1) = \mathrm{E}(u) + \Pr(\epsilon_1 \leq \epsilon_1^*).
\end{aligned}$$

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 $\mathrm{E}(P) = \infty$.

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

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

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

Proof of Proposition 4.1: We shall prove the proposition both for the linear scrap and the non-linear scrap case. In both cases the inequality constraints (4) 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

$$\begin{gathered}
C_1^* \leq C_1^* + I_1^* = (1-\omega_1^*) d_1 Y_1 \leq B_1 e^{\tau\epsilon_1}, \qquad (11) \\
I_1^* \leq C_1^* + I_1^* \leq B_1 e^{\tau\epsilon_1}, \\
(1-\delta)K_1 \leq K_2^* \leq (1-\delta)K_1 + B_1 e^{\tau\epsilon_1},
\end{gathered}$$

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 (11) 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

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

References

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

Arrow, K.J. (2009). A note on uncertainty and discounting in models of economic growth, *Journal of Risk and Uncertainty*, 38, 87–94.

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

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

Buchholz, W. and M. Schymura (2012). Expected utility theory and the tyranny of catastrophic risks, *Ecological Economics*, 77, 234–239.

Cai, Y., K.L. Judd, and T.S. Lontzek (2012). DSICE: A dynamic stochastic integrated model of climate and economy, Mimeo, Stanford University.

Chanel, O. and G. Chichilnisky (2013). Valuing life: Experimental evidence using sensitivity to rare events, *Ecological Economics*, 85, 198–205

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

Eeckhoudt, L., and C. Gollier (1995). *Risk: Evaluation, management and sharing*. Hertfordshire: Harvester Wheatsheaf.

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

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.

Hwang, C., F. Reynès, and R.S.J. Tol (2013). Climate policy under fat-tailed risk: An application of DICE, *Environmental and Resource Economics*, in press.

Horowitz, J. and A. Lange (2009). What's wrong with infinity — A note on Weitzman's dismal theorem, Mimeo, University of Maryland.

Ikefuji, M., R.J.A. Laeven, J.R. Magnus, and C. Muris (2013a). Pareto utility, *Theory and Decision*, 75, 43-57.

Ikefuji, M., R.J.A. Laeven, J.R. Magnus, and C. Muris (2013b). Economics meets statistics: Background document, http://www.janmagnus.nl/items/risk.pdf.

Karp, L. (2009). Sacrifice, discounting and climate policy: five questions, Mimeo, University of California, Berkeley.

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.

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. (2009). An analysis of the dismal theorem, Cowles Foundation Discussion Paper 1686, Yale University.

Nordhaus, W.D. (2011). The economics of tail events with an application to climate change, *Review of Environmental Economics and Policy*, 5, 240–257.

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

Pindyck, R.S. (2011). Fat tails, thin tails, and climate change policy, *Review of Environmental Economics and Policy*, 5, 258–274.

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.

Tol, R.S.J. (2003). Is the uncertainty about climate change too large for expected cost-benefit analysis?, *Climatic Change*, 56, 265–289.

Ursúa, J.F. (2010). Long-run volatility, Mimeo, Harvard University.

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.