Catastrophe aversion

Social attitudes towards common fates

Nicolas Treich and Christoph Rheinberger

APPORTS DE LA RECHERCHE

n° 2016-02

TOPIC

Human and organizational factors of safety

Les Cahiers de la sécurité industrielle

FONCSI
Fondation pour une culture de sécurité industrielle

The *Foundation for an Industrial Safety Culture* (FonCSI) is a french public-interest research foundation created in 2005. It funds research activities that contribute to improving safety in hazardous organizations and contributes to a better mutual understanding between high-risk industries and civil society.

FonCSI works in an interdisciplinary manner, in France and internationally, and aims to innovate and anticipate tomorrows' problems. It has four main goals:

- ▷ Identify emerging ideas and innovative practices
- ▷ Develop and finance research activities
- ▷ Contribute to the development of a research community on industrial safety
- ▷ Disseminate findings to all interested parties

The work presented in this document is the result of research funded by the FonCSI and by the French Ministry of the environment, in the context of the SAF€RA ERA-NET. Eric Marsden (FonCSI) has worked with the authors to adapt the content to the audience of the *Industrial Safety Cahiers* collection. The opinions presented are those of the authors.

Foundation for an Industrial Safety Culture
A public-interest research foundation

www.FonCSI.org

6 allée Émile Monso — BP 34038
31029 Toulouse cedex 4
France

Twitter: @TheFonCSI
Email: contact@FonCSI.org

Résumé

Titre L'aversion aux catastrophes

Mots-clefs catastrophes, risque, aversion, économie, corrélation

Auteur Christoph Rheinberger et Nicolas Treich

Date de publication juin 2016

L'attention croissante apportée au changement climatique, aux pandémies et aux crises financières met en valeur la question de la prise en compte des risques catastrophiques. Ce document propose une analyse de la manière dont la société aborde, ou devrait aborder, ces événements à faible probabilité mais aux conséquences dramatiques. La question sous-tendant ces travaux est la manière dont la société aborde

- ▷ un accident majeur faisant un grand nombre de victimes ;
- ▷ un grand nombre d'accidents faisant chacun un mort

lorsque les deux situations conduisent au même nombre total de décès.

Nous caractérisons les risques catastrophiques de façon mathématique comme un étalement à moyenne constante de la distribution des pertes, ou comme une distribution « plus risquée » des risques. Nous analysons des études conduites en sciences de la décision, en psychologie et en économie comportementale pour éliciter les attitudes des personnes vis-à-vis de différents risques sociétaux. Cette revue de la littérature trouve davantage de résultats suggérant une aversion aux catastrophes que l'attitude inverse de préférence pour le regroupement des décès (à nombre de décès constant).

Nous présentons un certain nombre d'explications comportementales possibles pour ces observations, tirées de travaux en psychologie. Ensuite, nous analysons comment différentes fonctions de préférences sociales, utilisées dans la théorie du choix social (l'analyse économique de l'agrégation des préférences individuelles en une préférence collective), représentent l'aversion aux catastrophes. Nous expliquons pourquoi l'aversion aux catastrophes peut rentrer en conflit avec des préférences pour l'équité. Le document termine par une discussion des approches utilisées pour évaluer et réguler les risques catastrophiques dans différents secteurs, avec une analyse de la manière dont ces préférences pourraient être intégrées à l'analyse coût-bénéfices.

À propos des auteurs

Nicolas Treich est économiste. Il dirige le Laboratoire d'économie des ressources naturelles (Lerna) à la Toulouse School of Economics. Ses travaux de recherche concernent l'analyse coût-bénéfice, la théorie du risque, l'économie environnementale et l'économie comportementale.

Christoph Rheinberger était chercheur post-doctorant à la Toulouse School of Economics. Il travaille aujourd'hui à l'agence européenne des produits chimiques (ECHA) à Helsinki.

Courriel : `ntreich@toulouse.inra.fr`

Pour citer ce document

Rheinberger et Treich (2016). *L'aversion aux catastrophes*. Numéro 2016-02 des *Cahiers de la Sécurité Industrielle*, Fondation pour une Culture de Sécurité Industrielle, Toulouse, France (ISSN 2100-3874). DOI : 10.57071/882rpq. Disponible à `foncsi.org/fr`.

Abstract

Title Catastrophe aversion: social attitudes towards common fates

Keywords catastrophes, risk, social choices, aversion, economics

Author Christoph Rheinberger and Nicolas Treich

Publication date June 2016

In light of climate change and other existential threats, policy commentators sometimes suggest that society should be more concerned about catastrophes. This document reflects on what is, or should be, society's attitude toward such low-probability, high-impact events. The question underlying this analysis is how society considers

- a major accident that leads to a large number of deaths;
- a large number of small accidents that each kill one person

where the two situations lead to the same total number of deaths.

We first explain how catastrophic risk can be conceived of as a spread in the distribution of losses, or a "more risky" distribution of risks. We then review studies from decision sciences, psychology, and behavioral economics that elicit people's attitudes toward various social risks. This literature review finds more evidence against than in favor of catastrophe aversion.

We address a number of possible behavioral explanations for these observations, then turn to social choice theory to examine how various social welfare functions handle catastrophic risk. We explain why catastrophe aversion may be in conflict with equity concerns and other-regarding preferences. Finally, we discuss current approaches to evaluate and regulate catastrophic risk, with a discussion of how it could be integrated into a benefit-cost analysis framework.

About the authors

Nicolas Treich leads the thematic group on Environmental Economics and Natural Resources at the Toulouse School of Economics. His research concerns benefit-cost analysis, risk theory, environmental economics and behavioral economics.

Christoph Rheinberger was a post-doctoral researcher at the Toulouse School of Economics. He now works at the European Chemicals Agency ECHA, in Helsinki.

Contact: ntreich@toulouse.inra.fr

To cite this document

Rheinberger and Treich (2016). *Catastrophe aversion: social attitudes towards common fates.* Number 2016-02 of the *Industrial Safety Cahiers*, Foundation for an Industrial Safety Culture, Toulouse, France (ISSN 2100-3874). DOI: 10.57071/882rpq. Available at www.foncsi.org/en.

Contents

Introduction

Over the last decade, catastrophic risks such as those posed by natural disasters, financial collapse, health epidemics, nuclear war, and bioterrorism have met with growing interest of scholars in the social sciences [Posner 2004; Bostrom et Ćirković 2008; Sunstein 2009; Taleb 2007; Barro et Ursua 2012]. A straightforward way to gauge the rising interest in the field of economics is to count papers published in leading journals that contain keywords related to catastrophes and disasters. Figure 1 presents the result of such a bibliometric analysis[1], demonstrating that the level of interest has indeed rallied since the early 2000s.

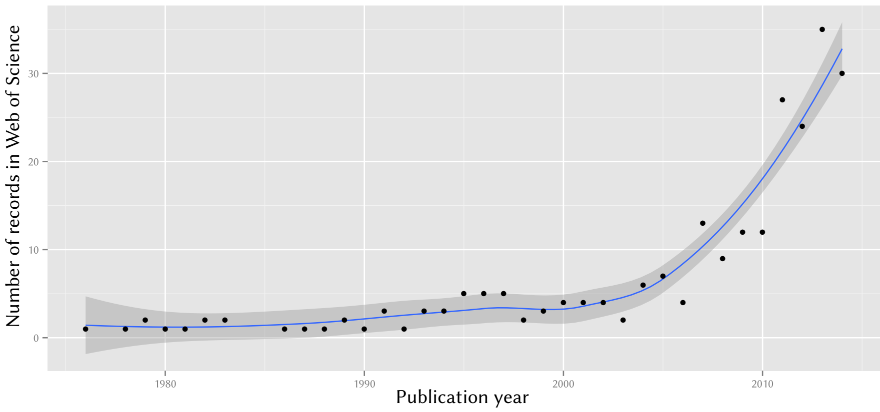


Figure 1 – *The growing interest of economists in catastrophes and disasters, as measured by a bibliometric search of the Web of science™ with the keywords* `catastrophe`, `catastrophes`, `catastrophic`, `disaster`, `disasters`, *and* `disastrous`. *The graph shows the number of articles on these topics published in selected economics journals between 1974 and 2014.*

Recently, economists have devoted special attention to the modeling of climate catastrophes. The following are some important contributions. In his dismal theorem, [Weitzman 2009] is concerned with the implications of structural uncertainty for the economics of low-probability, high-impact catastrophes. He provocatively argues that cost-benefit analysis (CBA) cannot be used to assess "fat-tailed" risks[2]. [Millner 2013] scrutinizes the implications of the dismal theorem and proposes an alternative framework in which the contribution of catastrophes to welfare loss is dampened. [Pindyck et Wang 2013] use a general equilibrium model of production, capital accumulation, and household preferences to assess how much society should spend on reducing the probability and impact of a climate catastrophe. [Barro 2015] shows that optimal investment in climate change mitigation increases with both social risk aversion and the probability and size of catastrophes. Most recently, [Martin et Pindyck 2015] analyze the limits of cost-benefit analysis (CBA) to deal with multiple catastrophic risks[3].

[1] The set of keywords and the list of journals included in the bibliometric analysis appear in the Appendix.

[2] In economics, a risk is a probability distribution over future outcomes. A fat-tailed risk is one in which extreme events (located in the tails of its probability distribution) occur more frequently than for risks with a more normal (Gaussian, for example) probability distribution.

[3] The recent debate about climate catastrophes is reminiscent of that about "Doomsday models" in the early days of resource economics (see [Solow 1972]).

Definition

Weitzman's dismal theorem

Harvard economist Martin Weitzman argues that while standard economic models of behavior under uncertainty are applicable to situations where the distribution of outcomes can be approximated by a Gaussian-type probability distribution, they are poorly suited to modelling situations with "fat-tailed" probability distributions, where catastrophic outcomes are possible (with very small probabilities). In his view, the possible consequences for an insurance company of clients' car accidents (a large number of outcomes with relatively low compensation payouts) are not of the same nature as the possible consequences of climate change (small probability of very adverse events).

Weitzman argues that in the latter situation, standard CBA breaks down. When some of the potential outcomes involve the deaths of hundreds of millions of people, not to mention the destruction of the world economy, he argues that the multiplication of huge damages by tiny probabilities, as used in CBA, is not a useful decision-support tool. The argument revolves around properties of the extreme tail values of "fat-tailed" functional forms, which are commonly used in economic models of the distribution of damages from some risks.

What do these contributions have in common? Starting from the premise of a representative agent who seeks to maximize expected utility over uncertain consumption paths, they model catastrophic risk in essentially the same fashion. As [Weitzman 2009] puts it:

> *The basic idea is that a society trading off a decreased probability of its own catastrophic demise against the cost of lowering the probability of that catastrophe is facing a decision problem conceptually analogous to how a person might make a tradeoff between decreased consumption as against a lower probability of that person's own individually catastrophic end.*

Since economists deal with catastrophic — and even existential — risks just like they deal with ordinary consumption risks, they also presume that society *should* be catastrophe averse in the very same way the representative agent is risk averse with regard to aggregated consumption. The sanctity of the representative agent in climate-economic models is highlighted by the discussion of Weitzman's dismal theorem, which has focused on the limitations of the proposed constant relative risk aversion (CRRA) utility function [Nordhaus 2011; Pindyck 2011; Horowitz et Lange 2014; Millner 2013; Arrow et Priebsch 2014].

In this document, we introduce an alternative framework which conceptualizes catastrophes as social risks that bear a small chance of many people dying (or, more generally, facing a loss) together. The key aspect here is the *coincidence* of many deaths, because the social planner in our framework cares about the total number of fatalities in each state of the world. In other words, our focus is on **population risk**. We characterize the catastrophic potential of a risk by the **spread in the distribution of fatalities within the population at threat**. Our main objective then is to explore what are conceivable attitudes toward catastrophic risks. We focus on threats to human life and limb mostly because the expected number of lives saved is the prevailing benefit measure of policies designed to avoid climate catastrophes, or nuclear accidents, or terrorist attacks. The expectancy criterion fails, however, to account for society's desire to avoid catastrophes. Let us illustrate this claim by the following example.

The U.S. President and Vice President do not travel together in order to avoid a "decapitation strike", *i.e.* the possibility of the simultaneous deaths of the commander-in-chief and his second. On the other hand, it seems reasonable to suggest that, when traveling with his family, the President should take the same plane to reduce the prospect of **bereavement**. In a seminal paper on the value of life, [Schelling 1968] proposed that:

> *If a family of four* must *fly, and has a choice among four aircraft, of which it is known that one is defective but not known which one, it should be possible to persuade them to fly together. The prospects for each individual's survival are the same, no matter how they divide themselves among the aircraft, but the prospects for bereavement are nearly eliminated through the correlation of their prospects. Society's interest, in support of the family's interest, should be to see that they are permitted and encouraged to take the same plane together.*[4]

[4] The problem can be scaled up and re-phrased to represent the ongoing debate about climate catastrophes. Assume that with some probability a tipping point in global warming will be reached at the end of the 21st century. Passing the tipping point might lead to the extinction of humanity with probability p, while the human species will adapt to more extreme climatic conditions with probability $(1-p)$. Another conceivable scenario is that passing the tipping point will cause a portion p of humanity to promptly die and the others to adapt. Both scenarios bear the same social risk in terms of the expected number of deaths, but few people would dispute that they are fundamentally different.

The above example illustrates that, depending on the decision context, one may or may not prefer the situation that gives rise to the possibility of simultaneous deaths. This prompts two important questions:

1. How do we behave in the face of a looming catastrophe?
2. How should we behave in order optimally to protect ourselves against catastrophes?

Structure of the document

In the remainder of the document, we collect insights from decision theory, behavioral economics, psychology, social choice and risk management studies to reflect upon these questions. We proceed as follows:

- ▷ In chapter 1, we provide a **formal definition** of what we mean by catastrophic risk. We use this definition to introduce different attitudes toward catastrophic risk that characterize preferences over the frequency and size of adverse outcomes.
- ▷ In chapter 2 we present a review and discussion of **empirical studies** that explore how people make decisions in face of catastrophic risks. These include survey studies and laboratory experiments that examine subjects' reactions to risky choices involving monetary outcomes. We also analyze a number of behavioural factors that could explain the results obtained in the empirical studies, including the way in which emotions affect people's perception of risk, issues of inequity aversion, salience, and psychophysical numbing and framing effects.
- ▷ Chapter 3 provides a summary of the **social choice** approach to evaluating risky social situations. This consists of analyzing the normative implications of different ways in which a social planner can aggregate individial welfare into a mathematical expression of a society's collective welfare. The approaches analyzed include utilitarianism, *ex ante* prioritarianism and *ex post* prioritarianism. We show that it is difficult to develop a plausible model of preferences for a decision-maker who is averse to both *ex ante* risk inequity and to catastrophic outcomes.
- ▷ Chapter 4 analyzes what regulators actually do when they manage potentially catastrophic risks. Approaches discussed include cost-benefit analysis (which is catastrophe neutral in its standard form) and F-N societal risk criteria (which may represent either catastrophe aversion or acceptance, depending on the tolerability criteria chosen).

1

Definitions

The term *catastrophe* has a different meaning for different people. This is true both in general, and within the field of economics. For the sake of clarity, we start with a precise statistical definition of what we mean by catastrophic risk. This definition is consequentialist, focusing solely on the number of fatalities. It characterizes catastrophes as a “bunching of fatalities”, without specifying whether the bunching arises because of a common cause, or among one identified group of people, or at the same time, or in the same place **[Lathrop 1982]**. In our framework, individuals are identical, and the identity of victims and survivors does not matter. Moreover, we will compare social risks solely in terms of the spread in the distribution of fatalities. That is, we will keep the expected number of fatalities constant across the compared situations.

Key issue

> In this chapter, we assume that the reader is familiar with standard definitions of risk and uncertainty used in economics. For an overview of these fundamental definitions, we suggest the interested reader consult the document *Risk attitude & economics* in the companion *Viewpoints on industrial safety* collection, freely available online at `foncsi.org`.

1.1 What does it mean for a risk to be catastrophic?

Intuition. How to value the possibility of a big accident as opposed to many small accidents?

Let us compare two situations where two individuals, Holmes and Moriarty, are exposed to a risk of death. In situation A, represented to the left of figure 1.1, either Holmes will die, or Moriarty will die, with equal probability. In situation B, either Holmes and Moriarty die together, or both stay alive, with equal probability. In situation C, both Holmes and Moriarty always die.

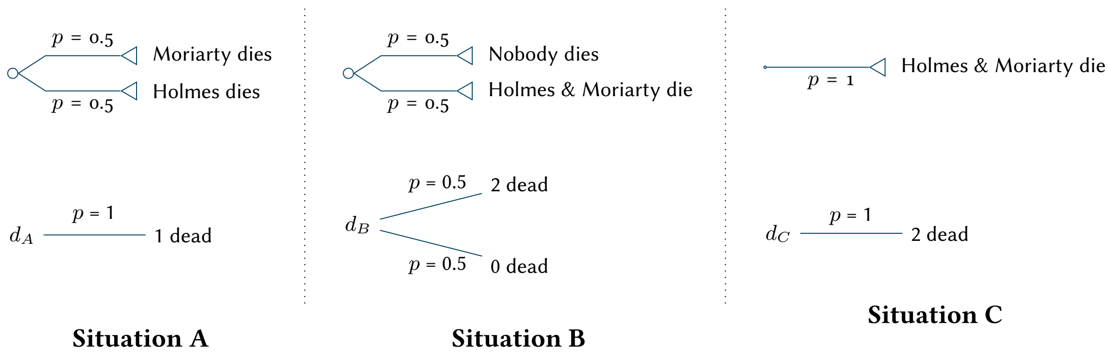


Figure 1.1 – *Probability trees (top) and the corresponding distribution of fatalities (bottom) for three risk situations.*

We can represent the probability distributions associated with the number of deaths in situations A, B and C as illustrated in figure 1.2. For example, distribution d_A represents the distribution of the number of deaths in situation A (1 dead with probability 1), and CDF_A represents the associated cumulative distribution function[1].

More formally, consider a population of $i = 1 \ldots N$ individuals, each of whom faces the probability of dying $p_i \in [0,1]$ due to a particular cause. The risk of death is modeled as a Bernoulli random variable $\tilde{x}_i$ which takes on the value 1 if individual i dies and 0 otherwise. We are interested in the **distribution of fatalities across the population at risk**:

$$d \stackrel{\text{def}}{=} \sum_{i=1}^{N} \tilde{x}_i$$

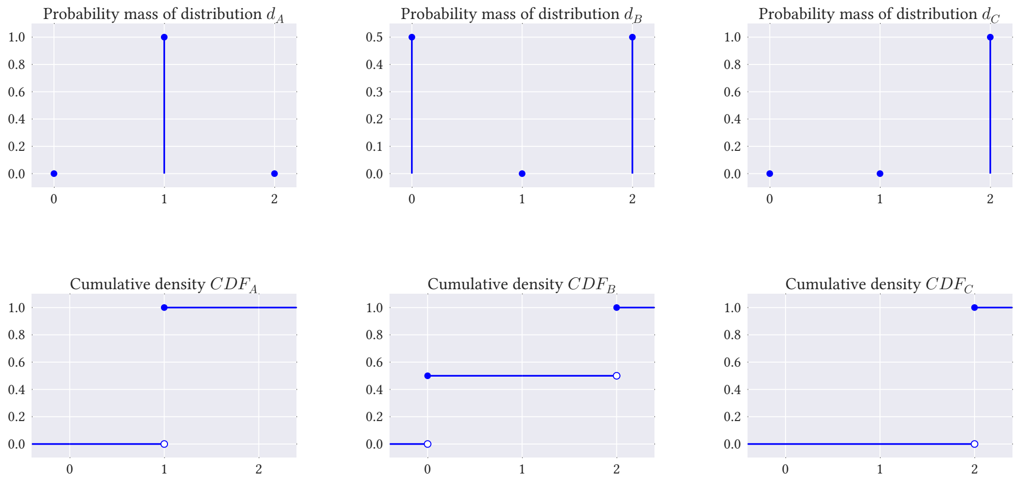


Figure 1.2 – *Probability distributions of the number of deaths in situations A, B and C described above*

Which of situations A, B and C do we prefer? Clearly, situation C is the most unfavourable, since it always leads to more deaths than the other situations. This idea is formalized in economics by the notion of *first-order stochastic dominance*.

Definition

First-order stochastic dominance

In economics, a gamble[2] G is said to have **first-order stochastic dominance** [Hadar et Russell 1969] over another gamble H if it can be written as

$$G = H + \text{“positive utility”}$$

This means that we can match up the outcomes in G and H so that the outcomes in G are at least as good as those in H (in a pairwise manner), and that some outcomes in gamble G are strictly better (further to the right in the plot of the CDF of utility) than those in H.

More formally, this can condition be written $CDF_G(x) \leq CDF_H(x)$ for all x, with a strict inequality for some x.

Classical decision modelling assumes that decision-makers have an *increasing utility function* (more is preferred to less). If gamble A has first-order stochastic dominance over gamble B, a standard decision-maker will always prefer gamble A. However, in this document we are generally comparing outcomes in terms of numbers of deaths, so are comparing the level of *disutility* of situations (fewer deaths will be prefered to more).

1 Note that while we have used discrete probability distributions in all our definitions, the general concepts of first-order and second-order stochastic dominance and mean-preserving spread presented later in this section are also relevant for continuous probability distributions.

2 Recall that a gamble in economics is a situation where more than one outcome is possible and the probability of each possible outcome is known exactly.

Note that if G first-order stochastically dominates H then G necessarily has a strictly larger expected value: for a population of N individuals, the expected value of the number of deaths

$$\mathbb{E}[G] = \sum_{i=1}^{N} d_G(i) \times i > \mathbb{E}[H] = \sum_{i=1}^{N} d_H(i) \times i$$

The implication doesn't go the other way though. Just because one distribution has a larger expected value than another doesn't mean that the first stochastically dominates the second.

In our example, a classical decision-maker will find situation C the most unfavourable (having the highest disutility), because it has first-order stochastic dominance over situations A and B.

However, the first-order stochastic dominance criterion does not allow us to choose between situations A and B, because they have the same number of expected deaths (1 in each case). To choose between them, we will use the notion of *second-order stochastic dominance*.

Definition

Second-order stochastic dominance [Rothschild et Stiglitz 1970]

Following [Rothschild et Stiglitz 1970], we can define *risk* as a **mean-preserving increase in dispersion** (or variability) of the outcomes[3]. We will say that a gamble G has **second-order stochastic dominance** over another gamble H if it can be written

$$G = H + \text{“risk”}$$

Equivalently, gamble G has at least as high an expected value as gamble H and involves less risk. This concept provides a ranking of equal-mean gambles according to their *degree of risk*.

Classical decision modelling assumes that people are *risk averse*: they prefer situations with a low variability in the possible outcomes (a low level of risk) to situations with more variability. If gamble G has second-order stochastic dominance over gamble H, then G will always be preferred by a risk-averse decision-maker.

We can define second-order stochastic dominance more formally using the notion of *mean-preserving spread*.

Definition

Mean-preserving spread (MPS)

A mean-preserving spread [Rothschild et Stiglitz 1970] is a change from one probability distribution G to another probability distribution H, where H is created by “spreading out” one or more portions of G's probability density function or probability mass function while leaving the mean (the expected value) unchanged.

More formally:

- ▷ probability distributions d_G and d_H have the same mean
- ▷ the graph of CDF_G crosses the graph of CDF_H exactly once
- ▷ CDF_G lies on or above CDF_H to the left of the crossing point and on or below to the right of the crossing point
- ▷ CDF_G and CDF_H are not the same distribution (they differ on some non-empty interval)

Figure 1.3 illustrates this notion for situations A and B. Note that d_A and d_B have the same mean (1 dead), their cumulative distribution functions cross exactly once, and CDF_A is below or equal to CDF_B to the left of 1 and above or equal to it to the right of 1, with a difference between the two distributions between 0 and 2.

[3] For example, a gamble resulting in 4 or 6 deaths based on the result of a coin flip is less risky than a similar gamble leading to 2 or 8 deaths (again depending on the outcome of a coin flip). Note in this definition that the mean outcome (the “expected value”) remains the same for both gambles, as does the probability of each outcome. Thus risk can increase by keeping both the probabilities and the expected value the same.

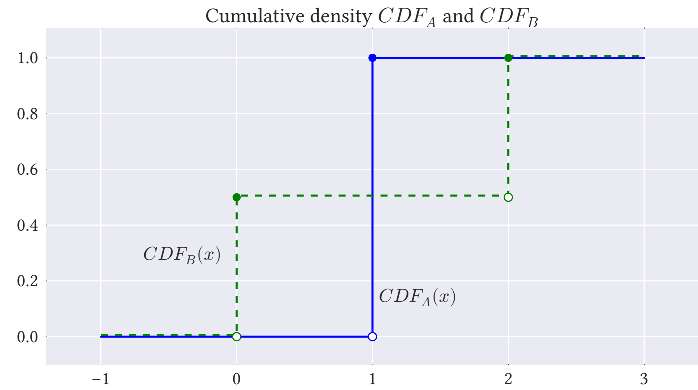


Figure 1.3 – *Comparison of CDF_A and CDF_B for the Moriarty/Holmes example. Situation B is a mean-preserving spread of A; it is more catastrophic than A.*

An important result demonstrated by **[Rothschild et Stiglitz 1970]** is that if B is a mean-preserving spread of A, then A second-order stochastically dominates B. We will say that situation B is "more catastrophic"[4] than situation A.

We define a *more catastrophic* distribution of fatalities — a more catastrophic risk, for short — based on the notion mean-preserving spread defined above.

Catastrophic risk (our definition)

Definition

A distribution of fatalities d_G is **more catastrophic** than another distribution d_H if it is a mean-preserving spread of d_H. Equivalently, d_H is second-order stochastically dominated by d_G.

[Rothschild et Stiglitz 1970] show that this definition is equivalent to the following: for any concave function $f(\cdot)$,

$$\mathbb{E}[f(d_G)] = \sum_{k=0}^{N} d_G(k) f(k) \geq \mathbb{E}[f(d_H)] = \sum_{k=0}^{N} d_H(k) f(k) \tag{1.1}$$

where $d_G(k)$ denotes the probability of observing $k = 0 \ldots N$ deaths.

Note that our definition of catastrophic risk is similar to the definition of *ex post* risk equity used by **[Fishburn et Sarin 1991]**.

Ex ante and ex post inequity

Definition

Ex ante equity concerns the fairness of individuals' marginal probabilities of death, as they exist at the time of decision making. *Ex post* equity concerns the equality of individuals' fates in the outcome of an alternative. Usually, alternatives that are *ex post* equitable are also *ex ante* equitable; but the converse is not true.

To illustrate, if two people play Russian roulette with a single gun, this is *ex ante* equitable, but clearly not *ex post* equitable.

[4] Note that this definition of catastrophic risk differs from the common use of the term (which involves a large number of fatalities due to a single event). Risk associated with a catastrophic event will always be catastrophic in the sense of our economic definition, but the converse is not necessarily true (a risk may be more catastrophic than a commonly experienced risk without necessarily leading to large number of fatalities).

1.2 Attitudes towards catastrophic risk

In the same way as we define attitudes towards risk (risk aversion, risk-seeking and risk neutral preferences), we can define categories of attitudes towards catastrophic risk.

Definition

Attitudes towards catastrophic risk

We define three distinct **attitudes toward catastrophic risk**:

▷ **Catastrophe aversion** is a preference for the less catastrophic distribution, or, equivalently, for a mean-preserving contraction in the distribution of fatalities;

▷ **Catastrophe acceptance** is the reversed preference for the more catastrophic distribution, or, equivalently, in favor of a MPS in the distribution of fatalities;

▷ **Catastrophe neutrality** implies indifference between the two distributions of fatalities.

This definition is consistent with Keeney's seminal work on catastrophe avoidance [Keeney 1980b][5]. In fact, our definition is more general, as Keeney only considered binary distributions in which one outcome is zero fatalities.

Let us consider two examples to illustrate different social attitudes towards catastrophes.

Example

Lifeboat allocation and catastrophe aversion[6]

Four passengers are aboard a sinking boat. There is only one lifeboat, with a design capacity of two people. The captain (whose code of honour requires him to go down with his boat) has the choice between the two following options:

A Put exactly half of the passengers onto the lifeboat. These passengers will survive for sure, whereas the passengers who remain on the sinking boat will die for sure. In order to choose who boards the lifeboat, the skipper designs a fair lottery, so each passenger has a probability of 0.5 of obtaining a place on the lifeboat.

B Let all the passengers board the lifeboat. The overcrowded lifeboat has a 50% chance of sinking, leading to the death of all of the passengers. If the lifeboat does not sink, all of the passengers will survive.

The cumulative density function for the distribution of deaths in situations A and B is shown in figure 1.4. Note that B is a mean-preserving spread of A (the expected number of deaths is the same in the two situations, but the risk in situation B is greater; it is more catastrophic than A).

If one believes that society should have a preference for avoiding catastrophes, then option A is socially preferred to option B. If one instead thinks that it is more equitable if individuals share a common fate, then option B is socially preferred to option A. However, if society is catastrophe neutral (expresses no concern for common fate considerations), then only the expected number of fatalities should matter in this example, with the consequence that options A and B are socially indifferent.

Let us now revisit the presidential travel example.

5 Keeney was not the first to allude to society's fear of catastrophic events. In the 1970s, nuclear engineers started to argue that society's apparent acceptance of daily hazards and its concern for possibly catastrophic hazards suggests catastrophe aversion. The argument was based on the idea that a consistent relationship exists between the severity and frequency of fatal accidents that would somehow reflect social attitudes toward risk [Starr 1969].

6 This example is taken from [Gajdos et al. 2010].

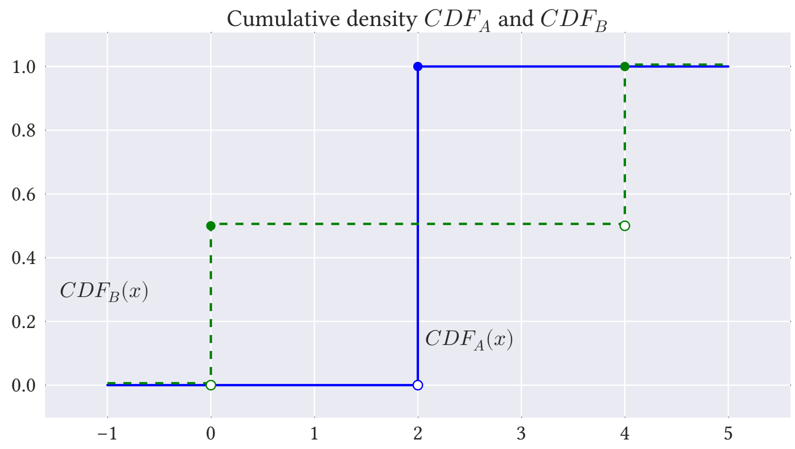


Figure 1.4 – *Cumulative density of the distribution of deaths in lifeboat options A and B*

US presidential travel arrangements

Example

Assume the U.S. President and Vice President need to attend the same ceremony to which they can travel either together (A) or separately (B). Let us assume that the likelihood of a plane crash, and therefore the probability of dying, is p. There are but two admissible distributions of deaths: in case they travel together, either both stay alive or both die (with a chance $d_A(2) = p$); in case they travel separately, there is some chance $d_B(1)' = 2p(1-p)$ that one of them dies in a plane crash, and a very small chance $d_B(2) = p^2$ that both die in two different plane crashes. Moreover, chances are higher that no one dies if they travel together: $d_A(0) = 1 - p > d_B(0) = (1-p)^2$. Figure 1.5 shows the two admissible distributions of deaths as probability trees.

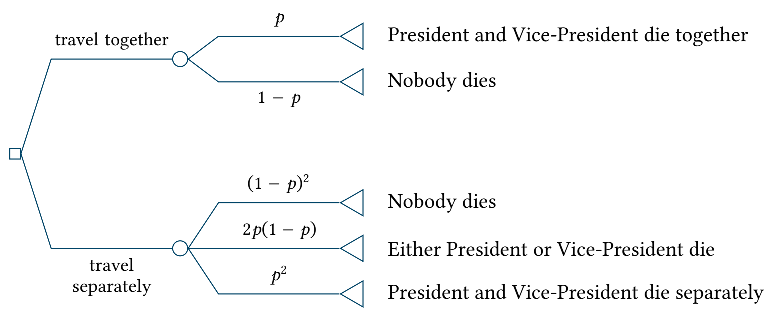


Figure 1.5 – *Probability tree for the presidential travel example*

Catastrophe averse preferences imply that one prefers the two leaders to travel separately, whereas catastrophe accepting preferences imply a preference for the leaders travelling together.

Let us clarify the difference between the MPS definition of catastrophic risk and an alternative definition based on the mean-variance rule of modern portfolio theory. The latter definition implies that catastrophe aversion is equivalent to preferring the less variable of two distributions with identical expected loss in terms of the number of fatalities:

$$B \succ A \text{ if } \mathbb{E}[A] = \mathbb{E}[B] \text{ and } \mathrm{Var}[B] > \mathrm{Var}[A]$$

It is immediate that a more catastrophic distribution in the sense of a MPS must have a greater variance. The converse is not true, however. This is because the variance induces a complete ordering, while a MPS gives only rise to a partial ordering. The MPS criterion is therefore stronger than the increasing variance criterion.

An important point to raise is that catastrophe aversion implies, but is not implied by, preferences for minimizing the maximum probable loss. While at odds with common justifications for catastrophe aversion **[Peterson 2002]**, this non-equivalence reflects an intuitive principle: what matters to a catastrophe averse decision-maker when comparing two distributions of fatalities is not only the likelihood of the worst possible outcome, but also that of the second worst, third worst, *etc.*[7] Our conception of catastrophe aversion can be directly applied to address questions such as: "How should a single accident that takes N lives be weighted relative to N accidents, each of which takes a single life?" **[Slovic et al. 1984, p. 464]**. Indeed, within our framework the question is formally interpreted as a decision between the sum of N i.i.d. Bernoulli random variables, *i.e.* $\sum_{i=1}^{N} d(i)$, versus N times one single Bernoulli random variable, *i.e.* $N \times d(1)$. The latter distribution is more catastrophic than the former, and preferences are determined by the decision-maker's attitude toward catastrophic risk.

[7] In risky situations involving only two individuals, minimax, minimal variance, and catastrophe averse preferences are strictly equivalent (see **[Bernard et al. 2015]**).

2

Catastrophic risk and empirical social choices

We now turn to the available evidence on attitudes toward catastrophes. Below, we provide a summary of empirical studies that put subjects in the shoes of a social planner[1] and ask them to make risky choices affecting other people.

2.1 Risky social choices involving lives

We start with evidence arising from **survey studies** in which subjects evaluate risky social situations involving human lives. Let us stress a number of limitations beforehand. First and foremost, these studies are necessarily hypothetical. Experimental economics involves the observation of choices made by volunteers in a laboratory setting, and typically use monetary incentives (at the end of the experiment, volunteers will receive a monetary payout which depends on the choices they made and on random factors) to align their decisions in a laboratory setting with "real" decisions. Decisions involving possibly many fatalities cannot be incentivized in the same way as economic experiments typically are. Secondly, the perspective of a social planner is unusual for most people and their decisions are thus prone to various heuristics and biases as explored in the psychometric risk literature. Thirdly, subjects are typically asked to consider the problem in isolation and to ignore the social cost of their decisions. In light of these limitations, we recognize that the level of generalizability of these survey studies is limited. Nonetheless, they provide a good starting point. The general picture that emerges from our summary is striking and, perhaps, surprising: most studies summarized in table 2.1 provide **evidence in favor of catastrophe-accepting attitudes**.

We proceed by reviewing various types of risky social choices that have been used to elicit subjects' attitudes toward catastrophic risk. We emphasize that most problems involve "life or death" prospects affecting fewer than 1 000 individuals. The term "catastrophic" is hence used in the technical sense of our definition on page 8, rather than in the colloquial sense of an apocalyptic threat.

Standard gambles. Probably the most famous risky social choice problem is the Asian disease problem [Tversky et Kahneman 1981], which can be illustrated as a standard gamble[2] as in figure 2.1. As the original problem was designed to detect preference reversals, there exists a loss framing and a gain framing (in square brackets):

> *Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows: If Program C [A] is adopted, 400 [200] people will die [be saved]. If Program D [B] is adopted, there is 1/3 probability that nobody [600 people] will die [be saved], and 2/3 probability that 600 people [no people] will die [be saved]. Which of the two programs would you favor?*

[1] In economics, a *social planner* is a decision-making authority which aims to achieve the best result for all parties. This role could for example be played by the government, though in practice many factors lead governments to make decisions which are far from ideal for their constituents.

[2] A standard gamble is a decision problem in which a "sure thing" prospect is compared to a risky prospect.

Author(s)	Year	Method	People at risk	Evidence
Tversky & Kahneman	1981	Standard gamble	200-600 deaths	mixed results, framing effect
Hammerton et al.	1982	Standard gamble	100 deaths	catastrophe aversion
Fischhoff	1983	Standard gamble	100-100 000 deaths	catastrophe acceptance
Slovic et al.	1984	Standard gamble	170-300 deaths	catastrophe acceptance
Keller & Sarin	1989	Standard gamble	10-100 deaths	catastrophe acceptance
Keller and Sarin	1989	Paired gamble	10-100 deaths	catastrophe acceptance
Hubert et al.	1991	Paired gamble	10-1 000 deaths	catastrophe aversion
Jones-Lee & Loomes	1995	Ranking choices	25-30 deaths	catastrophe acceptance
Fetherstonehaugh et al.	1997	Ranking choices	5 000-20 000 deaths	catastrophe acceptance
Druckman	2001	Standard gamble	200-600 deaths	mixed results, framing effect
Itaoka et al.	2006	Discrete choice experiment	4 000-20 000 deaths	catastrophe aversion
Abrahamsson & Johansson	2006	Paired gamble	up to 1 000 deaths	catastrophe acceptance
Slovic	2007	Ranking choices	100 deaths	catastrophe acceptance
Olivola & Sagara	2009	Paired gamble	200-600 deaths	catastrophe acceptance
Rheinberger	2010	Paired gamble	up to 100 deaths	catastrophe acceptance
Rheinberger	2010	Discrete choice experiment	1-6 deaths per year	catastrophe acceptance
Carlsson et al.	2012	Ranking choices	up to 200 deaths	catastrophe acceptance among public
Carlsson et al.	2012	Ranking choices	up to 200 deaths	mixed results among experts

Table 2.1 – *Empirical studies of risky social situations involving life and death prospects*

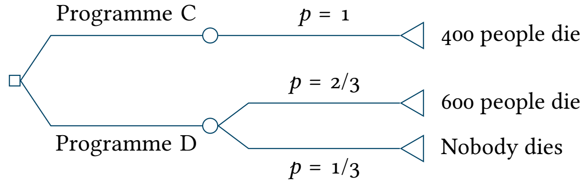


Figure 2.1 – *Risky social decision problem: probability tree for the "asian disease" problem expressed as a standard gamble*

Tversky and Kahneman found strong evidence against catastrophe aversion when the framing emphasized the number of deaths (78% of subjects preferred program D to C) and, conversely, strong evidence for catastrophe aversion when the framing emphasized the number of lives saved (72% of subjects preferred program A to B).

Dozens of replication studies have confirmed that people are consistently more risk seeking in the loss domain, and more risk averse in the gain domain **[Kühberger 1998]**, a phenomenon known as *prospect theory*.

In a notable study, **[Olivola et Sagara 2009]** manipulated responses to the Asian disease problem by showing subjects differently shaped distributions of mortality-causing events beforehand. The results suggest that distributions of event-related death tolls which people observe in their daily lives govern the evaluation of life-saving interventions and associated risk preferences .

Another important study on risky social choices was devised by Fischhoff **[Fischhoff 1983]**. Assuming that the natural way of framing such decision problems is one that emphasizes the loss of life, he explored preferences over catastrophic risk by varying both the loss of lives and the corresponding probability in a systematic way. A large majority of subjects (70% and more) chose the probabilistic option across all of Fischhoff's treatments, providing strong evidence for catastrophe-accepting preferences. In a similar vein, **[Slovic et al. 1984]** asked several hundred students whether they would rather reduce the frequency of single-fatality accidents or of large-scale accidents involving 300 deaths. More than 70% of their subjects preferred reducing the frequency of single-fatality accidents, even though reducing the frequency of large-scale accidents implied a 10% reduction in the expected number of fatalities.

The only standard gamble study that finds some support for catastrophe-averse preferences asked subjects whether society should rather sacrifice one (anonymous) individual in exchange for the survival of a group of 99 individuals, or accept that each of the 100 individuals faces a 1/100 probability of dying **[Hammerton et al. 1982]**[3]. 100 out of 118 subjects approved the sacrifice. Later replications by **[Keller et Sarin 1988]** did not, however, confirm the original result (42 out of 53 subjects preferred the gamble).

Paired gambles. In this type of decision problem, two risky options are confronted. Keller and Sarin **[Keller et Sarin 1988]** framed paired-gamble problems in the following way:

> *There are 100 islanders who are susceptible to a specific fatal disease which has recently appeared on the mainland. [...] All susceptible people must be injected with the serum within 24 hours, or each will have a 15% chance of contracting the disease and eventually dying. There is no time to acquire more serum. There are only 3 000 milligrams of the serum available. As the public health officer, it is your job to choose between the following options:*
>
> *A) Give the same low dose of 30 milligrams of serum to all 100 susceptible islanders. 50 of those susceptible are northerners, 50 are southerners. Each susceptible person will have an independent 10% chance of dying. The expected number of deaths is 10.*
>
> *B) Divide up the available serum among the 50 northerners who are susceptible to the disease. Thus, these people will receive a higher 60 milligram dose. Each of the 50 will now have an independent 5% chance of dying. Since the 50 susceptible southerners will receive none of the serum, each will still have a 15% chance of dying by contracting this disease. The expected number of deaths is 10.*

All but one of the 53 subjects chose option A, which is a MPS of option B. In figure 2.2 we divide the population at risk by 50 to illustrate that the problem is one of splitting a scarce resource among two individuals so that common fates become more likely (option A), or one individual is favored over the other so that it is more likely to observe one fatality (option B). Keller and Sarin used several paired gambles of this type, all pointing to catastrophe-accepting preferences. **[Hubert et al. 1991]** used a different paired-gamble method. In their study, a sample of 22 high-ranked French public administrators indicated the probabilities that would make them indifferent between 100-fatality and 1 000-fatality accidents and a 1/10 chance of a 10-fatality accident. The results indicate strong catastrophe aversion. However, the findings have to be taken with a grain of salt as they rely on a sample of subjects who are responsible for local safety and may therefore consider additional decision constraints.

[3] Individuals' risk of dying was described as independent so that there was only a tiny probability (10^{-20}) of 100 deaths, but a significant probability (0.264) of observing two or more deaths.

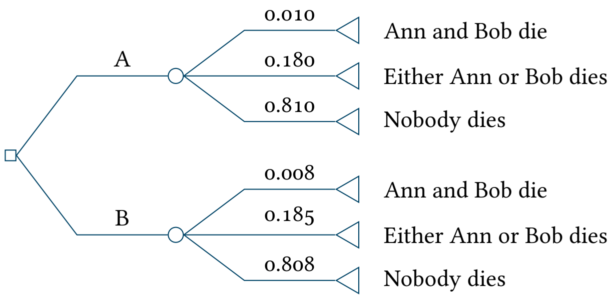


Figure 2.2 – *Risky social decision problem: paired gamble*

Two recent studies by [Abrahamsson et Johansson 2006] and by Rheinberger [Rheinberger 2010] extend the paired-gamble approach, using the gamble-tradeoff method [Wakker et Deneffe 1996]. The gamble-tradeoff method is a chained utility elicitation protocol, which requires subjects to equalize two risky gambles by stating an indifference value. The stated value is then used as reference value in the next elicitation step so that one ends up with a sequence of outcomes that are equally spaced in utility units. Based on the assumption of a CRRA-disutility function over the number of lives lost, $f(d) = d^{\gamma}$, the two studies estimate the subject-specific coefficient of relative risk aversion $\hat{\gamma}$. In both studies, the majority of subjects made choices implying catastrophe-accepting preferences; *i.e.*, $\hat{\gamma} < 1$.

Which of these two roads would you protect from avalanches?

	Road A	Road B
Expected accidents during 20 years	4	2
Expected number of fatalities per accident	3	6
I would protect:	❑ Road A	❑ Road B

Figure 2.3 – *An example of the choices presented to participants in the study* [Rheinberger 2010]

Constant relative risk aversion (CRRA)

Definition

A constant relative risk aversion utility function models the fact that absolute risk aversion decreases with wealth (richer people take more economic risks). It is a frequently used utility function when modelling investment decisions made by individuals.

In its common use in finance, CRRA represents an individual's level of aversion to risk. In this document it is used to analyze society's level of aversion to different numbers of deaths.

Discrete choices. A straightforward extension of paired gambles is to vary the frequency and severity of accidents systematically in a series of different combinations of scenarios from which subjects choose the preferred one. The result is a discrete choice experiment[4] with options that may be phrased as alternative safety policies. For instance, [Itaoka et al. 2006] examine the influence of risk characteristics involved with fossil fuel and nuclear power generation on willingness to pay for the reduction of mortality risks. The attributes used to describe nuclear risks included the probability of a catastrophic accident and the corresponding expected loss of lives. The authors find evidence of catastrophe aversion, meaning that

4 A discrete choice model attempts to explain choices between a set of alternatives, examining "which one" decisions as opposed to "how much" decisions over continuous choice variables.

subjects focused on the conditional loss from a nuclear accident, not its probability. In contrast, **[Rheinberger 2010]** did not find any indication for catastrophe aversion in a discrete choice experiment specifically designed to examine the tradeoff between the number of road fatalities and the frequency of accidents. In a random utility framework, he estimates a structural model of preferences over catastrophic risk assuming CRRA-type risk preferences. The estimated coefficient of relative risk aversion is well below unity ($\hat{\gamma} = 0.84$), suggesting that subjects put distinctly more weight on the frequency of accidents than on the number of deaths.

Ordinal choices. The functional choice of the CRRA-disutility function can be related to a fundamental psychophysical principle known as Weber's law. This principle states that people react with diminishing sensitivity to steady increments in a stimulus (*cf.* definition on page 22). In an influential study, **[Fetherstonhaugh et al. 1997]** applied the principle to life-saving interventions and found that subjects are largely insensitive to the number of lives saved against a background of increasing numbers of lives at risk. Fetherstonehaugh and co-authors suggest a form of "psychophysical numbing" to explain the inability to appreciate losses of life as they become more catastrophic. In economic terms this insensitivity implies that the marginal utility of saving an additional life is decreasing and may even become negative **[Västfjäll et al. 2015]**.

In a different setup, **[Jones-Lee et Loomes 1995]** studied the perceived safety of transportation in London's Underground railway system. 222 subjects assessed the following statement on an ordinal ranking scale ranging from `<strongly disagree>` to `<strongly agree>`:

> *25-30 deaths in a single Underground accident is worse than 25-30 deaths in separate Underground accidents.*

Only 20% of the subjects agreed that the single accident was worse, while 67% thought separate accidents were worse. Having expressed their preferences over different transportation risks, subjects were then asked to assume that a specified safety budget had to be allocated to reducing fatalities from either large-scale or small-scale accidents. A minority of 23% preferred investing in the reduction of catastrophic accidents. **[Carlsson et al. 2012]** undertook a similar study on traffic safety in Sweden, surveying both members of the general public and public administrators, and found that the expressed risk preferences were more nuanced. While members of the general public preferred the avoidance of frequent single-fatality accidents over the avoidance of infrequent large-scale accidents, there was no clear preference among public administrators. Roughly 40% of them preferred avoiding either large-scale or small-scale accidents, while the remaining 20% stated indifference between the prospects.

Key issue

We conclude that the findings from empirical social choice studies involving multiple-fatality risks are somewhat blurry. Nevertheless, they suggest that the majority of study participants — notably most of them Americans or Europeans — do not want regulators to adopt catastrophe-averse policies for social risks to life and limb.

2.2 Risky social choices involving money stakes

So far, we have looked at studies that elicit preferences over risks involving human lives. Next, we briefly summarize the key findings of lab experiments that explore **other-regarding preferences** in risky choices over **money gambles**. Of course, preferences revealed in such gambles are different from those involving human lives, as is the decision context. So while these studies are not perfectly informative about attitudes towards catastrophic risk in the narrow sense, they are properly incentivized and may therefore reveal some dimensions of social risk attitudes more accurately than hypothetical choices. Existing studies focus on three issues:

1. **equality of opportunity**, or *ex ante* fairness (equality of expected value before the event);
2. **equality of outcome**, or equality of *ex post* payoffs (after the event);
3. attitudes toward social risks and in particular toward the correlation between one subject's prospects and the prospects of the other subjects.

Evidence on correlation attitude is relevant because it tells us whether or not subjects prefer to win/lose simultaneously.

Lotteries. One way to study correlation attitude in risky choices is by means of lottery tasks in which social context plays a role. Several such studies have recently been conducted. [Rohde et Rohde 2011] analyze how individuals' decisions under risk depend on risks that others face in the same environment. Their results suggest that people prefer risks in the *gain* domain to be independent rather than correlated. [Friedl et al. 2014] look at the take-up of insurance for correlated and uncorrelated risks and find that the average willingness-to-pay for insurance is significantly higher for latter type of risk. This suggests that people prefer risks in the *loss* domain to be correlated rather than independent. [Linde et Sonnemans 2012] and Lahno and Serra-Garcia [Lahno et Serra-Garcia 2015] zoom in on peer effects in risky choices and find that risk attitudes are affected by concerns for relative payoff: subjects are more risk averse in situations in which they can earn at most as much as their peers than they are in situations in which they will earn at least as much. [Harrison et al. 2013] explore yet another aspect of social risk. They elicit individual and group risk attitudes using multiple price lists and find them to be highly correlated.

Dictator games. [Konow 2000], [Krawczyk et Le Lec 2010], [Brock et al. 2013], and [López-Vargas 2014] present evidence from probabilistic Dictator games.

Dictator games

Definition

Dictator games [Kahneman et al. 1986], studies in experimental economics, involve choices between options with uncertain consequences for both the decision-maker and some other individuals. They allow the analysis of other-regarding preference in the presence of risk.

In the **standard dictator game**, one player (dictator) is asked to allocate a certain amount between himself and another player (the recipient). A dictator who seeks only to maximize his utility will allocate all of the money to himself, but in practice many players choose to allocate some of the money to the recipient. The most commonly observed split is a half-half split, illustrating a preference for fairness.

In a **probabilistic dictator game**, players are asked to allocate probabilities to win an indivisible good. Two players face a coin-flip risk between prospects $\mathcal{A}$ and $\mathcal{B}$. For example, both prospects may pay either 5€ or 50€ but outcomes are positively or negatively correlated. Using obvious notations, we have: $\mathcal{A}: \frac{1}{2}(5,5) \oplus \frac{1}{2}(50,50)$ vs $\mathcal{B}: \frac{1}{2}(5,50) \oplus \frac{1}{2}(50,5)$. If player 1 dislikes both unequal outcomes and risk, she will prefer prospect $\mathcal{A}$ over prospect $\mathcal{B}$, although she faces the same personal risk under both prospects.

A significant share of subjects in all of these studies exhibit preferences for *ex ante* equal chances, highlighting that fairness and equity preferences matter in risky choices.

[Brock et al. 2013] note, however, that preferences based exclusively on *ex ante* or *ex post* comparisons cannot generate the empirical patterns they observe. Even with regard to attitudes toward social risks there is no clear evidence that individuals avoid negatively correlated risks more often than positively correlated risks. *Ex post* fairness in outcomes would predict that people strictly prefer positively correlated gambles. [López-Vargas 2014] finds that subjects take substantially more risk when outcomes are *ex post* fairly distributed, but experiments by [Brennan et al. 2008] and [Bolton et Ockenfels 2010] do not support this result.

Public goods game

Definition

In a standard public goods game [Olson 1965], participants secretly choose how many of their private tokens to put into a public pot. The tokens in this pot are multiplied by a factor (greater than one and less than the number of players, N) and this "public good"[5] payoff is evenly divided among players. Each participant also keeps the tokens they do not contribute.

The group's total payoff is maximized when everyone contributes all of their tokens to the public pool. However, if players make strategic decisions on how much to contribute (applying principles developed in game theory), each will tend to "free-ride" and contribute nothing to the common pot.

Public goods games. [Milinski et al. 2008] initiated a wave of experimental papers that explore under which conditions people cooperate to avoid catastrophic outcomes. In their experiment, six players are endowed with 40€. The game stretches over ten periods, in each of which players decide how much they want to contribute to the avoidance of a climate catastrophe. Players have no means to communicate and, hence, cannot coordinate. The "catastrophe" is averted with certainty if at least 120€ are raised at the end of the game. In that case, each player receives the amount of money he or she has left. Otherwise, players face the collective risk $p = 0.9$ of losing everything, implying perfectly correlated financial risks. There exist two symmetric pure strategy equilibria: players contribute 0€ in every period, yielding an expected payoff of 4€ to each player; or players contribute 2€ in every period, yielding everyone a certain payoff of 20€. Only half of the ten participating groups reached the efficient equilibrium. Replications by [Tavoni et al. 2011] and [Barrett et Dannenberg 2012] demonstrate that communication dramatically increases the success in raising the target amount.

In terms of attitudes toward catastrophic risks, the results of public goods games are more difficult to interpret than those of lotteries or dictator games. Indeed, the measurement of risk preferences is confounded by the strategic nature of the game: the avoidance of the catastrophe is also the efficient equilibrium.

2.3 Behavioral explanations

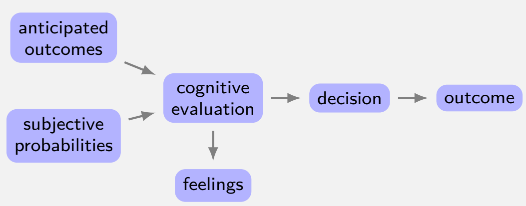
Several behavioral factors might explain the patterns emerging from the empirical studies reviewed above. We briefly visit some of them. Importantly, we neither suggest that the presented explanations are exhaustive nor that they are mutually exclusive.

Risk perception and emotions. Subjects in the above studies evaluated risky social situations in their context. It is well known that this context is shaped by the attention that society gives to various threats [Kasperson et al. 1988] and that catastrophes are produced and consumed by society in the sense that some risks receive much more attention than others [Zeckhauser 1996]. Dread, outrage and other negative feelings determine how badly a risk is perceived [Slovic 1987; Slovic et al. 2000]. It is therefore of little surprise that [Itaoka et al. 2006] found strong evidence for catastrophe aversion in choices that involved a catastrophic nuclear accident, a rare hazard over which the individual has little control, while [Rheinberger 2010] found catastrophe acceptance in choices over road accidents, an everyday hazard over which the individual has significant control. These findings coincide with the risk-as-feelings hypothesis of [Loewenstein et al. 2001], described below.

Definition

The risk-as-feelings hypothesis

Most theories of how risk affects people's decision-making take a cognitive approach, assuming that people assess the desirability and likelihood of possible outcomes and base their decisions on the mathematical notion of (utility-weighted) expectation. This is also called a *consequentialist* perspective, meaning that people make decisions based on their appreciation of the consequences of choice alternatives, as illustrated below (adapted from [Loewenstein et al. 2001]).

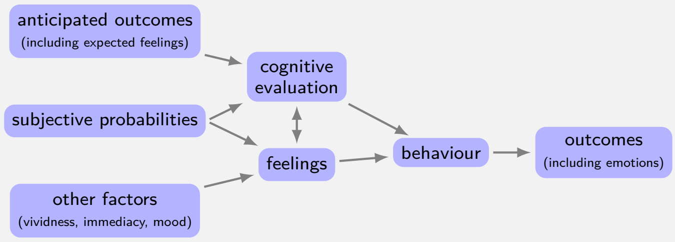


Researchers in psychology have long analyzed the role of an additional factor, emotions, in risk perception and risk-related behaviour. Some researchers have integrated **anticipated emotions** in models of decision-making, assuming that people anticipate how they will feel about the possible outcomes and integrate this in their cognitive decision-making process. The research shows that emotional reactions to risks can — and often do — diverge from cognitive evaluations of the same risks, as their determinants differ. Another strand of research on the **psychometric**

5 In economics, the term *public good* refers to costly, "non-excludable" projects that everyone can benefit from, regardless of how much they contribute to its creation (for instance a public green, or air quality).

paradigm suggests that people's perception of hazards are influenced by *dread* (perceived lack of control, perceived catastrophic potential) and by the *unknown nature* of the risk (hazard is unobservable, unknown, new).

The **risk as feelings** hypothesis integrates these strands in noting that people's responses to risky situations result both from emotional responses to the cognitive processing of the risks, such as anxiety and fear, and from anticipated emotions. According to this theory, people's decisions are determined by the interaction of these two influencing factors, as illustrated below.



Inequity aversion. One such reaction that pervades the evaluation of life and death prospects is that people dislike sending individuals to their "doom" (they are *ex post* inequity averse). An insightful study [Keller et Sarin 1988] asked subjects not only about their preferred option but also about the *fairer* option. The answers to the two questions were almost perfectly correlated, indicating that people chose what they deemed to be morally defensible. Their preferences seem to be linked to anticipated feelings such as guilt, remorse, and regret that might diminish the decision-maker's utility after the realization of the risk [Ellingsen et al. 2010; Battigalli et Dufwenberg 2007].

Salience and immediacy effects. Subjects' evaluations of social risks are likely to be affected by salience (prominence) and immediacy effects. Risks that subjects have never experienced before (as is likely for low-probability, large-loss events) are assessed differently from those just experienced or from those witnessed frequently. For instance, [Huang et al. 2013] found that people living close to a nuclear power plant in China perceived the risk of a nuclear accident to be much larger and the acceptability of the risk to be much lower after the Fukushima-Daïichi accident.

The salience hypothesis is consistent with a process-level model of magnitude evaluation based on memory sampling and relative judgment that does not rely on stable, underlying value representations (such as utility functions) to explain attitudes toward catastrophic risks [Olivola et Sagara 2009].

Definition

A process-level model of risk perception

Research in psychology suggests a process-based model for people's assessment of magnitudes (for example, the death toll from events):

1. think about comparable events present in their short-term memory then in their long-term memory;
2. compare the death toll for the event of interest with each of these comparable events;
3. the level of "shock" associated with a given death toll is a function not of the raw number of deaths, but rather of the proportion of recalled events of lower magnitude.

This model is supported by evidence that people are much better at making relative judgments than absolute ones. The model suggests that our assessment of an event's death toll will be a function of our accumulated experience and of recent events and their context. Our diminishing marginal sensitivity to human fatalities could be related to the fact that death tolls are distributed in such a way that most deadly events involve few deaths, and very few events involve large numbers of human fatalities.

The model provides a possible explanation for the differences in cross-cultural sensitivity observed in empirical studies, which may reflect cross-national variations in distributions of number of fatalities (for instance, in Western countries, events which cause large numbers of fatalities are rare compared with certain countries in Asia).

Alternatively, the salience hypothesis is also consistent with context-dependent choice models that assume that decision-makers are risk seeking when a prospect's upside is more salient and risk averse when its downside is more salient (*e.g.* [Bordalo et al. 2012]).

Salience theory of choice under risk

Definition

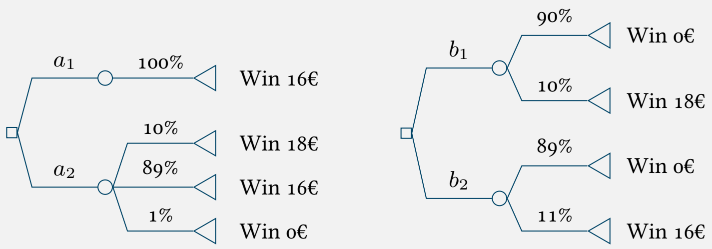
In classical economics, a decision-maker is assumed to process all available information (on probabilities, on severities) according to standard axioms. However, our mind has a tendency to focus on odd, different or unusual items[6]. When our attention is focused on specific aspects of the environment, the information in those aspects will receive disproportionate weighting in subsequent judgments, a phenomenon that cognitive psychologists call *salience*.

This concept can explain a number of well-known deviations from rational decision-making, including the shift from risk aversion to risk seeking as gains are reflected into losses (which motivated the development of prospect theory[7]). These shifts in risk attitudes can be predicted solely based on the salience of payoffs.

Probability weighting. It is well documented that people do not treat probabilities linearly (as required by expected utility theory) when faced with gambles over money stakes: they tend to overemphasize small probabilities and underemphasize large probabilities. For example, [Bruhin et al. 2010] analyzed almost 18 000 decisions over gambles involving real monetary gains and losses and found that roughly 80% of their 448 subjects exhibited significant deviations from linear probability weighting. One way to model such distortions in decision-making under risk is to apply a probability weighting function.

The Allais paradox

Definition



Economist Maurice Allais developed an experiment examining people's choices between two sets of gambles, described in the two probability trees above. Most people choose a_1 and b_1, which means that their preferences are not linear in probabilities. Indeed, for a decision-maker whose utility function is noted u and who chooses a_1 and b_1, the a lotteries imply that

$$1.0u(16€) > 0.89u(16€) + 0.1u(18€) + 0.01u(0€) \tag{2.1}$$

whereas the b lotteries imply that

$$0.9u(0€) + 0.1u(18€) > 0.89u(0€) + 0.11u(16€) \tag{2.2}$$

Equation 2.2 can be rewritten as

$$0.01u(0€) + 0.1u(18€) > 0.11u(16€) \tag{2.3}$$

$$0.01u(0€) + 0.1u(18€) > 1u(16€) - 0.89u(16€) \tag{2.4}$$

$$0.01u(0€) + 0.1u(18€) + 0.89u(16€) > 1u(16€) \tag{2.5}$$

which contradicts equation 2.1.

6 This is a useful ability for individuals who have to survive in a dangerous environment.

7 For an approachable introduction to prospect theory, see *Risk attitude & economics*, number 2014-01 of *Viewpoints on industrial safety*, freely available from foncsi.org.

One way of interpreting the Allais paradox is that people "weigh" probabilities in a manner that is not quite proportional. For example, when asked to explain their choices, many people indicate that they dislike the 1% chance of winning nothing in lottery b_2, and this small probability had a significant effect on their decision (the difference between a 0% probability of winning 0€ in a_1 and a 1% chance of winning 0€ in a_2 informs their decision more than the difference between the 89% probability of winning 0€ in b_2 and the 90% probability of winning 0€ in b_1).

Definition

Probability weighting function

Classical definitions of expected utility in economics require the existence of a function u mapping from the set of possible outcomes X to $\mathbb{R}$ such that the function $U(p) = \sum_{x \in X} p(x)u(x)$ represents a preference relationship over X (u is called the decision-maker's *utility function*).

Theories based on a probability weighting function require the existence of a function $\pi : [0, 1] \rightarrow [0, 1]$ such that

$$V(p) = \sum_{x \in X} \pi(p(x))u(x)$$

represents a preference relationship. The probability weighting function π takes the true (objective) probabilities and "warps" them into what are sometimes called *decision weights*.

Note that a probability weighting function that increases the weighting of low probabilities is sufficient to explain the Allais paradox described above.

Although probabilities were commonly known to subjects of the reviewed empirical choice studies, there is little reason to believe that probability weighting would not affect the choices made. Firstly, subjects may distort given probabilities when making choices due to decision weighting (as in prospect theory). Secondly, they may not believe the announced probabilities and may replace them by subjective probabilities. The commonly observed pattern of probability weighting is best modeled by an inverse S-shaped function **[Prelec 1998]**, and there is some evidence suggesting that the same pattern also applies to life and death prospects **[Rheinberger 2010]**. If that were the case, the empirical evidence should tend to favor catastrophe aversion, since small probabilities are typically inflated. This suggests that probability weighting either does not apply in choices over life and death prospects, or — more probable to us — is dominated by other psychological factors.

Psychophysical numbing. The psychophysical numbing hypothesis **[Fetherstonhaugh et al. 1997]** is a direct explanation for why people do not exhibit catastrophe averse preferences in life and death prospects. This hypothesis is related to a phenomenon known as "Weber's law".

Definition

Weber's law

19th century psychologists E. Weber and G. Fechner discovered a psychophysical principle that describes how we perceive changes in our environment. They found that people's ability to detect changes in a physical stimulus rapidly decreases as the size of the stimulus increases. This phenomenon, known today as "Weber's law", implies that for a change in a stimulus to become *just* noticeable, a fixed percentage must be added. Perceived difference in stimuli is a relative matter: for a small stimulus, a change is detected for a small increment, whereas for a large stimulus, only large increments are detected. Empirical studies have demonstrated that the growth of sensory magnitude is best fit by a power function of the stimulus magnitude[8].

Applying this principle to the value of averted fatalities suggests that a form of psychophysical numbing may result from our inability to appreciate losses of life as they become more catastrophic.

Reasons invoked for the diminished marginal sensitivity to the value of life include

- ▷ people's inability emotionally to appreciate the loss of lives in absolute numbers **[Slovic 2007]**;

[8] This is why certain scales for measuring perception of intensities, such as the decibel scale for sound, are logarithmic.

- ▷ a form of **motivation crowding-out** in which knowing the number of lives one cannot save reduces the utility of helping [Västfjäll et al. 2015];
- ▷ the disparity in the valuation of identified *vs* statistical victims [Small et Loewenstein 2003; Jenni et Loewenstein 1997].

All of these reasons may actually explain why the social disutility derived from an accident in which 1 000 people die might be less than 10 times the social disutility from another accident claiming 100 lives (catastrophe accepting preferences).

Framing effects. The empirical evidence on attitudes toward catastrophic risk is subject to various framing effects [Kühberger et al. 1999]. The Asian disease problem illustrates that one important effect relates to whether consequences are framed as lives *lost* or *saved.* Interestingly, when [Druckman 2001] modified the original problem (*cf.* page 13) to:

> “ *If Program C is adopted, 400 people will die and 200 people will be saved. If Program D is adopted, there is 1/3 probability that nobody will die and 600 people will be saved, and 2/3 probability that 600 people will die and no people will be saved. Which of the two programs would you favor?*

43% of subjects chose the catastrophe averse option C, which is approximately mid-way between the results he found for the gain framing (68.1%) and the loss framing (22.8%). Other framing effects are known to play a role in decisions under risk, and perhaps even more so in risky social situations. [Stewart et al. 2003] show that the set of available options has a large effect on choices under risk. This suggests that many people value prospects relative to one another, which gets support from the psychophysical numbing hypothesis. Issues of **accountability** have also been shown to affect preferences over life and death prospects. For example, subjects in [Rheinberger 2010] declared in debriefing questions that the social planner should be held accountable for multiple small-scale accidents, while rare large-scale accidents were rationalized as random acts.

Key issue

To summarize, the results of empirical social choice studies are consistent with various effects found in psychometric risk research. However, it is not clear which of the effects are predominant in explaining the observed attitudes toward catastrophic risks.

3

Catastrophic risk and social choice theory

The summary of empirical social choice studies provided in the previous chapter suggests a number of behavioral motives that might affect society's attitudes toward catastrophe. In this chapter, we turn our attention to **normative models** and explain how various risk attitudes can be grounded in social choice theory.

Definition

Social choice theory

Social choice theory[1] is the study of collective decision processes and procedures. It analyzes the aggregation of individual inputs (such as individual preferences or welfare) into collective outputs (such as collective preferences, group decisions and social welfare). It addresses questions such as *How can we rank the social desirability of different social alternatives?*

The social choice approach is most popular among philosophers and economists. It typically puts forward a set of axioms and explores their logical implications. This immediately raises a thorny question: Who decides about the axioms that society ought to adopt? The intuition of philosophers and economists might be poor guidance for collective decision-making. Relying instead on real choices, such as those made in ballots, does not seem very useful either, since existing policies and political institutions do not tell us much about the desirability of specific axioms. Experimental economics could be useful here, but experimental outcomes are often plagued by self-serving bias, and it is difficult properly to measure social fairness concerns in the lab [Johansson-Stenman et Konow 2010]. One could rely on stated preferences instead, by asking people which scenarios they would prefer to see implemented. However, applying ethical theories is often complex, and it is unclear how one should elicit lay people's opinions about the desirability of technical axioms in any informative way. Balancing these issues, we decided to simply present the most important axioms pertaining to catastrophic risks and leave it to the reader to judge their appeal for policymaking.

Social choice theory and cognate disciplines have a long-standing interest in attitudes toward risk. In his aggregation theorem, Harsanyi proposed that the social planner should maximize the sum of individual utilities as if individuals would maximize expected utility under a veil of ignorance (*i.e.*, without knowing their identity) [Harsanyi 1955]. This approach, know as **utilitarianism**, has however been widely contested for it is indifferent to both *ex ante* and *ex post* inequalities in the distribution of outcomes[2].

1 For further information, see the Stanford Encyclopedia of Philosophy entry at `https://plato.stanford.edu/entries/social-choice/`.

2 Prominent criticism includes that of Diamond [Diamond 1967], Sen [Sen 1970], Myerson [Myerson 1981], Broome [Broome 1984], Epstein and Segal [Epstein et Segal 1992], Ben-Porath et al. [Ben-Porath et al. 1997], and Fleurbaey [Fleurbaey 2010].

Social welfare functions

Definition

Welfare economics uses microeconomic techniques to evaluate well-being (or welfare) at the level of a country. A core concept is the *social welfare function* (SWF) [Bergson 1938], a coherent and consistent ranking of social states in terms of their desirability. The SWF can be thought of as a mechanism for assigning a "social value" to public policies (for example concerning distribution of income, or of risks, in a population).

The SWF concerns collective choices, which apply to all, but different individuals in a population may prefer different SWFS. For example, some may prefer a "social state" with more total income even if it is unevenly distributed, whereas others may be willing to give up some total income in order to have a more equitable distribution of income.

A number of alternative social welfare functions have been proposed to capture concerns for the worse-off in the evaluation of risky social situations [Bovens et Fleurbaey 2012]. In the context of mortality risk, [Adler et al. 2014] examined the most popular SWFS with respect to sensitivity to wealth and baseline risk, equal value of risk reduction, preference for risk equity and catastrophe aversion, and compared them to welfare assessments based on the standard CBA framework. Below, we illustrate some of these results using the canonical example of [Diamond 1967].

Example 2: three risky social situations

Example

Consider three risky social situations $\mathcal{S} = \mathcal{A}, \mathcal{B}, \mathcal{C}$, in which individuals $i = 1, 2$ face the risk of dying. They face two equiprobable states $j = \{\text{head}, \text{tail}\}$ (where j is determined by a coin toss). The state-dependent individual utilities u_{ij} associated with each situation are represented by the following outcome matrices:

$$\mathcal{A}: \begin{array}{c c} & \begin{array}{cc} i=1 & i=2 \end{array} \\ \begin{array}{c} \text{head} \\ \text{tail} \end{array} & \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \end{array} \quad \text{vs.} \quad \mathcal{B}: \begin{array}{c c} & \begin{array}{cc} i=1 & i=2 \end{array} \\ \begin{array}{c} \text{head} \\ \text{tail} \end{array} & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{array} \quad \text{vs.} \quad \mathcal{C}: \begin{array}{c c} & \begin{array}{cc} i=1 & i=2 \end{array} \\ \begin{array}{c} \text{head} \\ \text{tail} \end{array} & \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \end{array}$$

Situation $\mathcal{A}$ implies the most catastrophic risk, while situations $\mathcal{B}$ and $\mathcal{C}$ induce the same distribution of fatalities. Situation $\mathcal{B}$ is *ex post* unequal as in each state one individual dies while the other one stays alive. Situation $\mathcal{C}$ is *ex ante* and *ex post* unequal, as individual 2 is going to die and individual 1 stays alive irrespective of whether the coin lands on head or tail.

This example highlights the apparent conflicts between fairness considerations and social risk assessment in general [Fishburn et Sarin 1991, 1994; Gajdos et al. 2010] and between catastrophe avoidance and risk equity concerns [Keeney 1980b,a; Fishburn 1984; Bommier et Zuber 2008] in particular. [Lichtenstein et al. 1990] conclude that "one can achieve only one of the apparently laudable but inconsistent goals of fatality minimization, risk equity, and catastrophe avoidance". How do common SWFs handle these conflicting goals?

Prioritarianism

Definition

In ethics and political philosophy, the concept of prioritarianism states that the desirability of an outcome is a function of overall well-being across all individuals, with **extra weight given to worse-off individuals**. Whereas a utilitarian social planner maximizes total utility, and is therefore indifferent to the way in which utilities are distributed within the population, a prioritarian social planner gives greater priority to welfare changes affecting worse-off individuals[3].

Concerning choices under risk, the literature in economics distinguishes between *ex ante* and *ex post* prioritarianism. The *ex ante* prioritarian social planner maximizes the sum of transformed expected utilities, while the *ex post* prioritarian social planner maximizes the expectation of the sum of transformed utilities. As a result, the *ex post* prioritarian social planner cares about the difference in realized utilities *ex post*, once the risk is resolved. In contrast, the *ex ante* prioritarian social planner cares about the difference in expected utilities *ex ante*, before the risk is resolved.

Expressed mathematically, the corresponding social welfare functions are:

3 Mathematically, this means maximizing the sum of a strictly concave transformation of utilities.

▷ Utilitarian: $W_U(\mathcal{S}) = \sum_{i=1}^{n}\sum_{j=1}^{m} p_j u_{ij}$

▷ *Ex ante* prioritarian: $W_{EAP}(\mathcal{S}) = \sum_{i=1}^{n} \varphi\left(\sum_{j=1}^{m} p_j u_{ij}\right)$

▷ *Ex post* prioritarian: $W_{EPP}(\mathcal{S}) = \sum_{j=1}^{m} p_j \sum_{i=1}^{n} \varphi(u_{ij})$

where $i = 1...n$ represents individuals, $j = 1...m$ are states and p_j their respective probabilities. Statewise utility is noted u_{uj} and $\varphi(\cdot)$ is a strictly concave function.

Applied to the three social situations described in Example 2 above, we have:

▷ Utilitarian SWF: $W_U(\mathcal{A}) = W_U(\mathcal{B}) = W_U(\mathcal{C}) = 1$, demonstrating that the utilitarian SWF is catastrophe and inequity neutral.

▷ *Ex ante* prioritarian SWFs: $W_{EAP}(\mathcal{A}) = W_{EAP}(\mathcal{B}) = 2\varphi(1/2) > W_{EAP}(\mathcal{C}) = \varphi(1) + \varphi(0)$, demonstrating that *ex ante* prioritarian SWFs are catastrophe neutral, but exhibit *ex ante* inequity aversion.

▷ *Ex post* prioritarian SWFs: $W_{EPP}(\mathcal{A}) = W_{EPP}(\mathcal{B}) = W_{EPP}(\mathcal{C}) = \varphi(1) + \varphi(0)$, showing that *ex post* prioritarian SWFs result in the same preferential order as utilitarianism.

The above results demonstrate that common SWFs are catastrophe neutral; *i.e.* they do not distinguish between situations $\mathcal{A}$ and $\mathcal{B}$. Environmental economists should raise an eyebrow here, because policy evaluation is typically based on a utilitarian SWF, implying that popular integrated assessments models of climate change (*e.g.* DICE, RICE, FUND, PAGE) are catastrophe neutral! Let us, hence, turn to recent SWF approaches that do allow social attitudes toward catastrophe to be captured.

▷ *Ex post* transformed SWFs: $W_{EPT}(\mathcal{S}) = \sum_{j=1}^{m} p_j \phi\left(\left(\sum_{i=1}^{n} \varphi(u_{ij})\right)/n\right)$ for individuals $i = 1...n$, states $j = 1...m$, statewise utility u_{ij}, and a strictly concave function φ. The transformation function ϕ may be linear, concave, or convex. Applied to Example 2 we obtain $W_{EPT}(\mathcal{A}) = 0.5\phi(2\varphi(1)) + 0.5\phi(2\varphi(0))$ and $W_{EPT}(\mathcal{B}) = W_{EPT}(\mathcal{C}) = \phi(\varphi(1) + \varphi(0))$, so that preferences are catastrophe accepting (respectively averse) if and only if ϕ is convex (resp. concave). One noteworthy transformation is the equally distributed equivalent (EDE[4]) proposed by **[Fleurbaey 2010]**, which is obtained by setting $\phi = \varphi^{-1}$ (where ϕ is convex by the concavity of φ). When applied to example 2, the EDE results in catastrophe accepting attitudes toward social risk and neutrality toward *ex ante* inequity.

▷ Catastrophe-sensitive (CS) SWFs: $W_{CS}(\mathcal{S}) \sum_{j=1}^{m} \frac{1}{\varepsilon}(1 - \prod_{i=1}^{n}(1 - \varepsilon u_{ij}))$ for individuals $i = 1...n$, states $j = 1...m$, statewise utility u_{ij}, and the constraint $\varepsilon u_{ij} < 1$. When applying the CS-SWF proposed by **[Bommier et Zuber 2008]** to example 2, we obtain $W_{CS}(\mathcal{A}) = 1 - 0.5\varepsilon$ and $W_{CS}(\mathcal{B}) = W_{CS}(\mathcal{C}) = 1$. Consequently, attitudes are averse toward catastrophic risk if and only if ε is positive; they are catastrophe accepting iff ε is negative, and they are always neutral toward *ex ante* inequity.

The above examples allow us to draw several conclusions. First, *ex ante* approaches are always catastrophe neutral. Indeed, all that matters under *ex ante* approaches is individual expected utility. Therefore, the comparison of the realized levels of individual utility is irrelevant for welfare assessments. Consequently, *ex ante* approaches do not care about whether people reach simultaneously high or low utility levels in a specific state. In contrast, *ex post* approaches can be catastrophe-sensitive. Typically, a social planner who dislikes *ex post* inequity will prefer situation $\mathcal{A}$ over situations $\mathcal{B}$ and $\mathcal{C}$, as only in the first situation are utility levels equal across states. For that reason, the planner is sometimes said to exhibit **common-fate preferences** **[Schelling 1968]**.

Not all *ex post* approaches are catastrophe-sensitive, however. As W_{EPT} and W_{CS} illustrate, the key property that induces sensitivity to the possibility of catastrophe is **non-separability**.

[4] An *equally distributed equivalent* (or EDE) is a notion used in welfare economics. In our context, it is the level of individual risk exposure which, if equally distributed across the population, would lead to the same level of social welfare as the existing distribution of risk exposure.

Separable utility functions

Definition

Consider an individual who has a preference ordering $\prec$ over outcomes which consist of bundles of some finite number of commodities. Suppose these commodities can be divided into two commodity groups X and Y. The possible outcomes are described by $X \times Y$, with typical element (x, y). Preferences are said to be separable on X if $(x, y) \succ (x', y)$ for some $y \in Y$ implies $(x, y') \succ (x', y')$ for all $y' \in Y$.

For example, consider a preference relation over baskets of apples and oranges. Apples are separable from oranges if the preference for baskets containing the same number of oranges is determined only by the number of apples in the basket.

Example[5] of non-separable preferences

Example

Suppose that you have to choose between chicken and fish for today's and tomorrow's dinners. Possible outcomes are described by:

$$\{\text{chicken today, fish today}\} \times \{\text{chicken tomorrow, fish tomorrow}\}$$

If you always have to choose between these two meals, one day you will get bored of eating the same meal, and you will have a preference relation much like the following:

$$\begin{aligned}(\text{fish today, chicken tomorrow}) &\succ (\text{chicken today, fish tomorrow}) \\ &\succ (\text{fish today, fish tomorrow}) \\ &\succ (\text{chicken today, chicken tomorrow})\end{aligned}$$

In this situation, the choice attributes are not independent; what you prefer eating tomorrow is conditioned by what you eat today.

The separability of individuals' preferences (across goods, time periods, states of the world, *etc.*) is a frequently held assumption in economics. Concerning social choice, many scholars deem non-separability a problematic property because it implies that the level of utilities of individuals who are unaffected by a particular policy might still matter for the evaluation of that policy.

Finally, we point out that *ex post* approaches are not Pareto efficient and have the additional drawback of neutrality toward *ex ante* inequity[6]. We conclude that allowing for catastrophe-sensitive preferences comes at the cost of abandoning other normatively desirable goals. This conclusion traces all the way back to the seminal result of Keeney [Keeney 1980b], which states that there is always a conflict between *ex ante* and *ex post* risk equity[7]. Of course, the result tells us nothing about the attitude of the decision-maker toward catastrophic risk. It simply highlights the trade-offs that the decision-maker might face if he is averse to both *ex ante* risk inequity and to catastrophic outcomes.

5 This example is taken from [Fishburn 1970, p. 42].

6 See [Fleurbaey 2010, p. 658] and [Bovens et Fleurbaey 2012, p. 242-243] for a response to these objections.

7 One important limitation of Keeney's theorem is that it assumes individual risks are statistically independent. As we show in a companion paper [Bernard et al. 2015], it is possible to relax the independence assumption provided that a risk transfer from the more exposed to the less exposed person does not decrease the statistical correlation between their risks.

4

Regulatory approaches to catastrophic risk

The question of how potentially catastrophic risks should be regulated is, of course, not new. Already in 1975, the U.S. Nuclear Regulatory Commission stated **[USNRC 1975]**:

> *The public appears to accept more readily a much greater social impact from many small accidents than it does from the more severe, less frequent occurrences that have a smaller societal impact.*

In this chapter, we explore the apparent link between regulatory practice and people's beliefs and attitudes. We start our exposition by summarizing the use of cost-benefit analysis (CBA) for assessing health and safety policies. We then discuss alternative risk assessment metrics that are currently used to manage large-scale industrial risks. Finally, we will examine the political economy of risk regulation and the impact looming catastrophes have on it.

4.1 Standard Cost-Benefit Analysis

The standard economic approach to evaluate health and safety policies is CBA. So, how does CBA handle catastrophic risk? Standard CBA quantifies mortality risks in monetary units by multiplying the expected number of fatalities by the value of statistical life[1] (VSL), *i.e.* $\mathbb{E}[d] \times \text{VSL}$. Accordingly, policy-induced changes in risk have a monetary value that is proportional to a VSL figure (*e.g.*, the U.S. Environmental Protection Agency currently uses a value of about 8 million USD per statistical life). Since our definition of catastrophe aversion keeps $\mathbb{E}[d]$ constant, it is straightforward to see that CBA is catastrophe neutral. As an illustration, consider the monetized benefit of eliminating mortality risk in situations $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ of Example 2 (*cf.* page 26). Remember that exactly one individual will die in each of the three situations and, hence, the value of a risk-eliminating policy is in each situation just equal to the VSL. There is no premium whatsoever for eliminating the most catastrophic risk, which is associated with situation $\mathcal{A}$.

It has been argued, however, that we should employ context-specific VSL values. For example, **[Sunstein 2004]** suggests that the VSL for mortality risks that produce unusual fear or involve special suffering could be larger than standard VSL, that VSL for young people might be higher than for the elderly, and that VSL for risks affecting predominantly wealthy individuals could be higher than that for poorer individuals. Taken to the extreme, this idea would imply the collection of individuals' willingness to pay (WTP)[2] for each specific mortality risk reduction under scrutiny. The outcome of the CBA exercise would hinge on individual probability changes induced by each project.

To illustrate, consider again Example 2. When comparing situations $\mathcal{A}$ and $\mathcal{B}$, *ex ante* individual probabilities of dying are equal to 1/2 in both situations. The risks faced by the two individuals are identical, so there is no reason to justify different values for a risk elimination project.

---

[1] The term "value of statistical life" is somewhat controversial **[Cameron 2010]** and U.K. regulators therefore use the term "value of prevented fatality" (VPF) instead. In any case, the metric refers to an individual's marginal rate of substitution between income and mortality risk.

[2] The *Willingness To Pay* for a specified project or outcome measures how much an individual would be prepared to pay to benefit from the project or outcome. In the context of cost-benefit analysis of safety investments and the examples discussed in this document, we are generally concerned with WTP for a marginal change in the level of mortality risk to which the individual is exposed.

However, when we compare situations A and C, we find that *ex ante* individual probabilities differ because in situation C one individual is sure to die while the other one bears no risk at all. One can easily imagine that the former individual's WTP to eliminate the certainty of his or her own death is infinite.

More generally, policy assessments based on CBA are sensitive to how heterogeneity in individual risks affects the aggregate WTP for implementing a risk-reducing project [Pratt et Zeckhauser 1996]. Previous work has shown that this effect depends, among other things, on whether information concerns heterogeneity in baseline risk or in changes in risk [Armantier et Treich 2004] and whether the valuation is based on compensating or equivalent variation [Hammitt et Treich 2007].

In practice, there are at least two other motives that may be invoked to justify catastrophe attitudes in CBA. The first relates to the **perception of risk**. It is quite conceivable that people are willing to pay a premium to avoid or reduce catastrophic risks, either because of the small probabilities and large uncertainties associated with such events [Viscusi 1998; Treich 2010], or because of the horror of widespread deaths as invoked in popular movies. In his book *Catastrophe: Risk and Response*, [Posner 2004] attempts to estimate the monetary benefit of preventing human extinction in the year 2100. Presuming a VSL of $28 million and a future population of 12 billion people, he arrives at a tentative figure of $336 quadrillion. Posner maintains an inflated VSL value precisely to reflect the horror of extinction. Others have argued that even this 27-digit figure is much too small to capture the economic value of human extinction (see [Bostrom 2013]).

The second motive for including catastrophe attitudes into CBA relates to **other-regarding preferences**. Recall Schelling's quote in the introduction section (*cf.* 2), which mentions the prospect of bereavement associated with the possibility of surviving relatives. People may anticipate to tremendously suffer from an inequality that arises after the realization of the risk. As we have seen, there is experimental evidence supporting the notion of *ex post* inequity aversion. This is not to be confounded with altruistic motives as studied in the VSL literature [Bergstrom 1982; Jones-Lee 1992]. The latter are *ex ante* concerns and hence cannot be sensitive to the possibility of a catastrophe. Nonetheless, there is evidence that other-regarding preferences play a significant role in empirical VSL studies [Hammitt et Haninger 2010; Gerking et al. 2014], and this leaves open the question of the specific effect of catastrophic risk.

4.2 Alternative approaches

Aside from CBA, two other quantitative criteria are frequently used in real-world risk assessments: individual risk criteria (IRC) and societal risk criteria (SRC). As their labels suggest, the two criteria relate to different dimensions of risk. IRC are concerned with the distribution of risk across individuals. They place an upper limit on the "tolerable" or "acceptable" risk of death to individuals and, therefore, promote equity in the distribution of risk. SRC, on the other hand, are concerned with aggregate risk over groups of people. They place limits on the "tolerable" frequency of adverse outcomes (*e.g.* fatal accidents) of specified sizes. Depending on how the tolerability criterion is defined, SRC may promote catastrophe aversion or acceptance in decisions over risky social situations. In the following, we briefly review two ways of implementing SRC in a regulatory framework.

FN criterion. The most common way to implement SRC is by means of so-called Frequency-Number (FN) diagrams that plot the number of fatalities d against the cumulative frequency of events $F(d)$ that cause d or more deaths, and by a FN-criterion line $C(d)$ that specifies which risks are judged as tolerable[3]. When plotted on double-logarithmic scales as in figure 4.1, risk tolerability is determined by the (negative of the) slope factor γ of a straight FN-criterion line. A slope factor $\gamma > 1$ implies catastrophe averse preferences on behalf of the regulator[4]. [Evans et Verlander 1997] demonstrate that adhering to the FN-criterion is equivalent to adopting a minimax decision rule in which the regulator chooses the policy option j that fulfills:

3 FN-criteria are used in the Netherlands, Norway, Switzerland, and the U.K., among others [Rheinberger 2010; Bedford 2013].

4 A catastrophe-neutral regulator would, of course, adopt a slope criterion $\gamma = 1$, and a catastrophe-accepting regulator a slope $\gamma < 1$.

$$\min_j \max_d \left\{ \frac{F_j(d) d^\gamma}{C(d)} \right\}$$

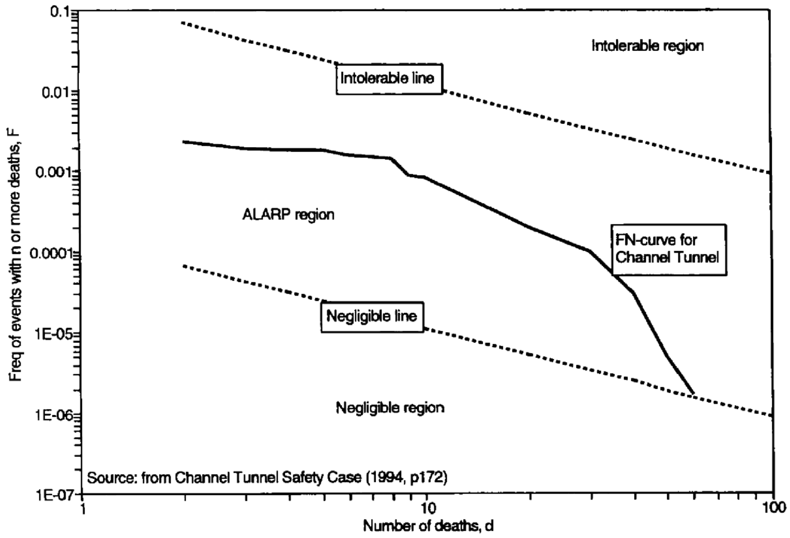


Figure 4.1 – *Example of an F-N diagram, from* [Evans et Verlander 1997]

Expected disutility criterion. An alternative SRC can be established based on the tenets of expected utility theory. Consider a catastrophe averse society and adopt the CRRA-disutility function[5] over the number of lives lost, $f(d) = d^\gamma$, to model catastrophe aversion. An accident of uncertain size $\tilde{d}$ can then be evaluated by the expected disutility

$$\mathbb{E}\left[f(\tilde{d})\right] = \sum_{k=0}^{N} \pi_k d^\gamma$$

where π_k is the probability that the event causes exactly k deaths. While *a priori* unknown, this probability can be approximated by the normalized frequency of size-k events; *i.e.* $\pi_k \approx f(k)/f$, where f is the overall frequency of events that cause deaths [Evans et Verlander 1997][6]. $\mathbb{E}[f(\tilde{d})]$ is compared against some tolerability threshold Ψ set by the regulator. Risk is deemed intolerable if $\mathbb{E}[f(\tilde{d})] > \Psi$ [7]. If the disutility function is assumed to be quadratic ($\gamma = 2$), a risk-as-variance definition can be invoked [Stallen et al. 1996]. In that case, uncertainty over the distribution of fatalities is regarded as a major cause of catastrophe aversion and this concern can be directly incorporated as a function of the spread of the distribution; *i.e.* the social risk is deemed intolerable if $\mathbb{E}[\tilde{d}] + \kappa\sigma[\tilde{d}] > \Psi$, where $\kappa > 1$ is an aversion factor that accounts for the larger spread in the distribution of fatalities as measured by its standard deviation $\sigma[\tilde{d}]$.

Regulatory frameworks that have adopted one of the above SRC follow a two-step approach. They first assess the tolerability of risk before and after the implementation of one or more policies. In the second step, they discard policy options that do not satisfy the tolerability criterion and perform the cost-benefit assessment only for the remaining policy options [Paté-Cornell 2002]. The major problem we see with intolerability criteria and similar decision standards is that they are conceptually fluffy. In the absence of a theoretical foundation, who is going to set such standards? We will touch on this and related questions of responsibility and accountability in policymaking related to catastrophic risk in the next section.

5 CRRA: see definition on page 16.

6 [Bedford 2013] argues that the approximation is mistaken for rare events (involving large n) and therefore difficult to defend in settings such as climate catastrophes. Instead, he proposes a two-parameter disutility function that extends the above approach to account for uncertainty in π_k.

7 [Bohnenblust et Slovic 1998] propose to marry the expected disutility criterion with standard CBA by multiplying for each policy option the expectation over the weighted number of deaths $\mathbb{E}[f(\tilde{d})]$ rather than the expected number of deaths $\mathbb{E}[\tilde{d}]$ by the VSL, obtaining a money measure of the disutility of lives lost.

4.3 The political economy of catastrophic risk regulation

Why is it that some risks receive a lot of attention both from policymakers and the general public, while others go almost undetected despite their potentially catastrophic nature? Consistent with the Social Amplification of Risk perspective [Kasperson et al. 1988], we suggest that several politico-economic factors may help to explain why and how society decides to emphasize one risk, but not another. Due to space limits, we can but discuss a few factors that are of particular relevance for understanding social attitudes toward catastrophic risk.

Perception. In *The Law of Fear: Beyond the Precautionary Principle*, [Sunstein 2005] forcefully argues that developed countries invest too much in the prevention of low-probability, high-impact risks. Sunstein's core argument combines insights from psychology and political economy: the public tends to overestimate the probability of dread events, and policymakers respond to the public's fears and anxieties in a populist manner[8]. Breyer made a similar point before when he alluded to the excessive congressional reaction to popularized risks as the "vicious circle of risk regulation" [Breyer 1993].

[Wiener 2016] counter-argues that catastrophic risks are prone to what he calls the "unavailability" heuristic: uncommon risks are not available to human perception since they are, by definition, rare events[9]. It is, therefore, not contradictory that people overestimate small risks like plane crashes or tornadoes [Slovic et al. 2000], but underestimate or ignore rare risks altogether [Klinke et Renn 2002]. This might explain Posner's observation that policymakers respond to catastrophic risks far less than to other, more common risks [Posner 2004]. The summarized arguments suggest a complex pattern of policymakers' responses to catastrophic risk, which might well depend on a number of other factors directly linked to the public's perception.

Responsibility. Another crucial factor for understanding the regulation of risks is the question of responsibility. In a world that encounters ever more man-made risks, we forge our own destiny [Giddens 1999]. Two questions come to the fore: Who is to determine which risks society accepts and which ones it does not? And who is to blame in the event of a catastrophe? It seems natural that government has a particular responsibility in overseeing activities that involve the risk of possibly many fatalities. It does so by enforcing safety standards and by intervening if those standards are not complied with. Yet it often does so in a seemingly arbitrary way, so that society ends up spending hundreds of millions of euros on avoiding a statistical fatality from one cause, and barely fifty-thousand euros on avoiding a statistical fatality from another cause [Tengs et al. 1995; Ramsberg et Sjöberg 1997]. While this might be puzzling at first sight, it demonstrates that government does not exercise its regulatory responsibility in a social vacuum.

Firms engaged in hazardous activities are another social entity involved in the regulation of risks. Their responsibility is to operate in the safest way possible or, at least, so that the residual risk is *as low as reasonably practicable* [Paté-Cornell 2002]. While firms are generally liable for their activities, and may be required to outlay large sums to compensate victims of their activities, they may not be sufficiently capitalized to pay off the litigation cost associated with a big accident. A very big accident may even destroy the legal system and its institutions [Wiener 2016]. Hence, there is a rationale in favor of regulatory actions that intend to prevent catastrophic events. Yet there are also reverse effects. Think of climate change: impacts are on a global scale and span wide into the future. Thus, there is little incentive for a local regulator to act unilaterally.

In democratic countries, citizens themselves take a share of responsibility in that they either tolerate a risky activity in exchange for its social benefits or oppose it exactly because of the risk involved: think of nuclear power in France vs Germany. The difficulty is, of course, that public preferences are often erratic and change over time. Moreover, as safety preferences are often catalyzed by big accidents [Kasperson et al. 1988], citizens tend to emphasize immediate risks that have a clear cause. Consequently, they are often more concerned with ubiquitous or

[8] The issue of whose perceptions should ultimately count in public policy is analyzed in [Portney 1992], [Pollak 1998], [Johansson-Stenman 2008], and [Salanié et Treich 2009] among others.

[9] Wiener notes that blockbuster movies like *Deep Impact* or *The Day After Tomorrow* do little to alleviate the unavailability heuristic since the public tends to perceive such movies as what they are: good entertainment.

timely threats than with rare but disastrous events that may happen sometime in the future. This highlights the close connection between responsibility and perception of risks.

Group interest. We have already alluded to various interest groups that have a stake in risk regulation and may now ask what role they play in regulating catastrophic risks? The answer seems to largely depend on the characteristics of the risk. One key criterion for regulators' response to catastrophic threats is how regulatory costs and benefits are shared. Consider a policy intervention that seeks to make road tunnels safer. The corresponding cost is borne by all taxpayers, while the benefits go primarily to road users in regions where there are a lot of tunnels. The possibility of a catastrophic accident affects the number of road users. It is not clear, however, whether their group size affects policymaking: if the user group is larger, the policy enjoys more political support; at the same time, there are stronger incentives to free-ride **[Olson 1971]**.

The regulatory cost might be borne by a specific industry that contributes to generating the potentially catastrophic risk in the first place, as is the case with greenhouse gas emissions. It is well known that the polluter-pays principle triggers strong political opposition and the influence and visibility of various lobby groups matters a lot for "pricing" catastrophe avoidance. Regulatory capture — the fact that a firm or industry may gain control over the agency meant to regulate it — can be a problem in government decision-making **[Laffont et Tirole 1991]**, which has been identified as a contributing factor in a number of major industrial accidents[10]. Interest groups exert pressure on politicians and/or policymakers who are concerned with their reputation **[Kuran et Sunstein 1999]**. They typically do so with the help of the media, which have an important role in shaping the public's perception of risks that may (or may not) have the potential for a catastrophe. This explains why not all catastrophes receive the same level of public attention.

[10] In the aftermath of the Fukushima-Daiïchi accident, von Hippel **[von Hippel 2011]** commented in a *New York Times* op-ed: *Nuclear power is a textbook example of the problem of "regulatory capture" [...which] can be countered only by vigorous public scrutiny and Congressional oversight, but in the 32 years since Three Mile Island, interest in nuclear regulation has declined precipitously.*

5

Conclusions

We have presented a framework to think systematically about catastrophes: the big ones that are going to increase both in frequency and size as ever more people are at risk, and the *very* big ones that threaten our existence on Earth. We have argued that standard economic approaches to assess health and safety risks do not account for important dimensions such as the sheer size of a catastrophe or the disutility of bereavement. Other approaches are conceivable, though. We have formally defined the notions of catastrophe aversion and acceptance as preferences for a mean-preserving contraction and spread in the distribution of fatalities, respectively. Evaluations based on these notions capture different *ex post* attitudes toward social risk. Our review of existing empirical choice studies suggests that, **in many contexts, people tend to be catastrophe-accepting**. This means that they prefer a gamble that offers the chance to escape the catastrophe over a certain loss, even if the expected loss is equal or smaller in the latter situation.

It is hard to tell which motives drive catastrophe-accepting attitudes. We have discussed several behavioral phenomena that may explain the observed empirical patterns. Perhaps the most important one is **inequity aversion**. Since a more catastrophic situation is generally also more equitable *ex post*, there is a clear link between catastrophe and equity. In the limit, maximal *ex post* equity is attained: either nobody or everybody dies. Framing is another factor that matters in choices over life or death prospects. People care about whether lives can be saved or must be sacrificed. We speculate that this emphasis on saving lives is hardwired to our moral instinct.

We end with a somewhat unsatisfactory reply to the two questions posed in the introduction. It is neither clear whether we are, nor whether we should be catastrophe averse. In future research, it may be interesting to study how the context systematically shapes our attitudes toward catastrophic risk. It may also be useful to identify the political forces (*e.g.* populism, responsibility, lobbying) that may induce or hinder regulatory action in face of a looming catastrophe. On the normative side, we see a need to characterize optimal risk policies in economic models under catastrophe-sensitive preferences. As our discussion has revealed, different disciplines including decision theory, behavioral economics, psychology, social choice and risk management have different perspectives on those issues. This calls for more integrated research on the management of catastrophic risks.

5

Appendix

The bibliometric analysis depicted in figure 1 is based on a query of the Web of Science™ database. The query words `catastrophe`, `catastrophes`, `catastrophic`, `disaster`, `disasters` and `disastrous` were used to search through abstracts, titles and key words of articles published between 1974 and 2014 in the following list of economics journals: American Economic Journal: Applied Economics, American Economic Journal: Economic Policy, American Economic Journal: Macroeconomics, American Economic Journal: Microeconomics, American Economic Review, Econometrica, Economic Journal, European Economic Review, Games and Economic Behavior, International Economic Review, Journal of Econometrics, Journal of Economic Theory, Journal of Finance, Journal of Financial Economics, Journal of International Economics, Journal of Labor Economics, Journal of Monetary Economics, Journal of Political Economy, Journal of Public Economics, Journal of the European Economic Association, Rand Journal of Economics, Review of Economics and Statistics, Review of Economic Studies, Quantitative Economics, Quarterly Journal of Economics, Theoretical Economics, American Journal of Agricultural Economics, Ecological Economics, Energy Economics, Environmental and Development Economics, Environmental and Resource Economics, Journal of Agricultural and Resource Economics, Journal of Environmental Management and Economics, Land Economics, Review of Environmental Economics and Policy.

Open science

The data associated with this analysis has been archived on Figshare with DOI `10.6084/m9.figshare.1564853`. The R code used to generate the figure is available on Figshare with DOI `10.6084/m9.figshare.1564855`.

Bibliography

Abrahamsson, M. and Johansson, H. (2006). *Risk preferences regarding multiple fatalities and some implications for societal risk decision making: An empirical study.* Journal of Risk Research, 9(7):703–715. DOI: 10.1080/13669870600799853.

Adler, M. D., Hammitt, J. K., and Treich, N. (2014). *The social value of mortality risk reduction: VSL versus the social welfare function approach.* Journal of Health Economics, 35:82–93. DOI: 10.1016/j.jhealeco.2014.02.001.

Armantier, O. and Treich, N. (2004). *Social willingness to pay, mortality risks and contingent valuation.* Journal of Risk and Uncertainty, 29(1):7–19. DOI: 10.1023/B:RISK.0000031442.60436.ea.

Arrow, K. J. and Priebsch, M. (2014). *Bliss, catastrophe, and rational policy.* Environmental and Resource Economics, 58(4):491–509. DOI: 10.1007/s10640-014-9788-6.

Barrett, S. and Dannenberg, A. (2012). *Climate negotiations under scientific uncertainty.* Proceedings of the National Academy of Sciences, 109(43):17372–17376.

Barro, R. J. (2015). *Environmental protection, rare disasters and discount rates.* Economica, 82(325):1–23. DOI: 10.1111/ecca.12117.

Barro, R. J. and Ursua, J. F. (2012). *Rare macroeconomic disasters.* Annual Review of Economics, 4:83–109. DOI: 10.1146/annurev-economics-080511-110932.

Battigalli, P. and Dufwenberg, M. (2007). *Guilt in games.* American Economic Review, 97(2):170–176. DOI: 10.1257/aer.97.2.170.

Bedford, T. (2013). *Decision making for group risk reduction: Dealing with epistemic uncertainty.* Risk Analysis, 33(10):1884–1898. DOI: 10.1111/risa.12047.

Ben-Porath, E., Gilboa, I., and Schmeidler, D. (1997). *On the measurement of inequality under uncertainty.* Journal of Economic Theory, 75(1):194–204. DOI: 10.1006/jeth.1997.2280.

Bergson, A. (1938). *A reformulation of certain aspects of welfare economics.* The Quarterly Journal of Economics, 52(2):310–334. DOI: 10.2307/1881737.

Bergstrom, T. C. (1982). *When is a man's life worth more than his human capital?* In Jones-Lee, M., Ed., *The Value of Life and Safety*, pages 3–26. North-Holland.

Bernard, C., Rheinberger, C. M., and Treich, N. (2015). *Catastrophe aversion and risk equity under dependent risks.* mimeo, Toulouse School of Economics.

Bohnenblust, H. and Slovic, P. (1998). *Integrating technical analysis and public values in risk-based decision making.* Reliability Engineering & System Safety, 59(1):151–159. DOI: 10.1016/S0951-8320(97)00136-1.

Bolton, G. E. and Ockenfels, A. (2010). *Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States: Comment.* American Economic Review, 100(1):628–633. DOI: 10.1257/aer.100.1.628.

Bommier, A. and Zuber, S. (2008). *Can preferences for catastrophe avoidance reconcile social discounting with intergenerational equity?* Social Choice and Welfare, 31(3):415–434. DOI: 10.1007/s00355-007-0292-6.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2012). *Salience theory of choice under risk.* The Quarterly Journal of Economics, 127(3):1243–1285. DOI: 10.1093/qje/qjs018.

Bostrom, N. (2013). *Existential risk prevention as global priority.* Global Policy, 4(1):15–31. DOI: 10.1111/1758-5899.12002.

Bostrom, N. and Ćirković, M. M. (2008). *Global catastrophic risks.* Oxford University Press. ISBN: 978-0199606504.

Bovens, L. and Fleurbaey, M. (2012). *Evaluating life or death prospects.* Economics and Philosophy, 28(2):217–249. DOI: 10.1017/S0266267112000235.

Brennan, G., González, L. G., Güth, W., and Levati, M. V. (2008). *Attitudes toward private and collective risk in individual and strategic choice situations.* Journal of Economic Behavior & Organization, 67(1):253–262. DOI: 10.1016/j.jebo.2007.02.001.

Breyer, S. (1993). *Breaking the Vicious Circle: Toward Effective Risk Regulation.* Harvard University Press. ISBN: 978-0674081147, 144 pages.

Brock, J. M., Lange, A., and Ozbay, E. Y. (2013). *Dictating the risk: Experimental evidence on giving in risky environments.* American Economic Review, 103(1):415–437. DOI: 10.1257/aer.103.1.415.

Broome, J. (1984). *Uncertainty and fairness.* The Economic Journal, 94(375):624–632. DOI: 10.2307/2232707.

Bruhin, A., Fehr-Duda, H., and Epper, T. (2010). *Risk and rationality: Uncovering heterogeneity in probability distortion.* Econometrica, 78(4):1375–1412. DOI: 10.3982/ECTA7139.

Cameron, T. A. (2010). *Euthanizing the value of a statistical life.* Review of Environmental Economics and Policy, 4(2):161–178. DOI: 10.1093/reep/req010.

Carlsson, F., Daruvala, D., and Jaldell, H. (2012). *Do administrators have the same priorities for risk reductions as the general public?* Journal of Risk and Uncertainty, 45(1):79–95. DOI: 10.1007/s11166-012-9147-3.

Diamond, P. A. (1967). *Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: Comment.* Journal of Political Economy, 75(5):765–766. DOI: 10.1086/259353.

Druckman, J. N. (2001). *Evaluating framing effects.* Journal of Economic Psychology, 22(1):91–101. DOI: 10.1016/S0167-4870(00)00032-5.

Ellingsen, T., Johannesson, M., Tjøtta, S., and Torsvik, G. (2010). *Testing guilt aversion.* Games and Economic Behavior, 68(1):95–107. DOI: 10.1016/j.geb.2009.04.021.

Epstein, L. G. and Segal, U. (1992). *Quadratic social welfare functions.* Journal of Political Economy, 100(4):691–712. DOI: 10.1086/261836.

Evans, A. W. and Verlander, N. Q. (1997). *What is wrong with criterion FN-lines for judging the tolerability of risk?* Risk Analysis, 17(2):157–168. DOI: 10.1111/j.1539-6924.1997.tb00855.x.

Fetherstonhaugh, D., Slovic, P., Johnson, S. M., and Friedrich, J. (1997). *Insensitivity to the value of human life: A study of psychophysical numbing.* Journal of Risk and Uncertainty, 14(3):283–300. DOI: 10.1023/A:1007744326393.

Fischhoff, B. (1983). *Predicting frames.* Journal of Experimental Psychology: Learning, Memory, and Cognition, 9(1):103–116.

Fishburn, P. C. (1970). *Utility Theory for Decision Making.* John Wiley and Sons. Publications in Operations Research (Book 15), ISBN: 978-0471260608, 234 pages.

Fishburn, P. C. (1984). *Equity axioms for public risks.* Operations Research, 32(4):901–908. DOI: 10.1287/opre.32.4.901.

Fishburn, P. C. and Sarin, R. K. (1991). *Dispersive equity and social risk.* Management Science, 37(7):751–769. DOI: 10.1287/mnsc.37.7.751.

Fishburn, P. C. and Sarin, R. K. (1994). *Fairness and social risk I: unaggregated analyses.* Management Science, 40(9):1174–1188. DOI: 10.1287/mnsc.40.9.1174.

Fleurbaey, M. (2010). *Assessing risky social situations.* Journal of Political Economy, 118(4):649–680. DOI: 10.1086/656513.

Friedl, A., De Miranda, K. L., and Schmidt, U. (2014). *Insurance demand and social comparison: An experimental analysis.* Journal of Risk and Uncertainty, 48(2):97–109. DOI: 10.1007/s11166-014-9189-9.

Gajdos, T., Weymark, J. A., and Zoli, C. (2010). *Shared destinies and the measurement of social risk equity.* Annals of Operations Research, 176(1):409–424. DOI: 10.1007/s10479-009-0588-2.

Gerking, S., Dickie, M., and Veronesi, M. (2014). *Valuation of human health: An integrated model of willingness to pay for mortality and morbidity risk reductions.* Journal of Environmental Economics and Management, 68(1):20–45. DOI: 10.1016/j.jeem.2013.11.010.

Giddens, A. (1999). *Risk and responsibility.* The Modern Law Review, 62(1):1–10. DOI: 10.1111/1468-2230.00188.

Hadar, J. and Russell, W. R. (1969). *Rules for ordering uncertain prospects.* American Economic Review, 59(1):25–34.

Hammerton, M., Jones-Lee, M. W., and Abbott, V. (1982). *Equity and public risk: Some empirical results.* Operations Research, 30(1):203–207. DOI: 10.1287/opre.30.1.203.

Hammitt, J. K. and Haninger, K. (2010). *Valuing fatal risks to children and adults: Effects of disease, latency, and risk aversion.* Journal of Risk and Uncertainty, 40(1):57–83. DOI: 10.1007/s11166-009-9086-9.

Hammitt, J. K. and Treich, N. (2007). *Statistical vs. identified lives in benefit-cost analysis.* Journal of Risk and Uncertainty, 35(1):45–66. DOI: 10.1007/s11166-007-9015-8.

Harrison, G. W., Lau, M. I., Rutström, E. E., and Tarazona-Gómez, M. (2013). *Preferences over social risk.* Oxford Economic Papers, 65:25–64. DOI: 10.1093/oep/gps021.

Harsanyi, J. C. (1955). *Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility.* Journal of Political Economy, 63(4):309–321. ideas.repec.org/a/ucp/jpolec/v63y1955p309.html, DOI: 10.1007/978-94-010-9327-9_2.

Horowitz, J. and Lange, A. (2014). *Cost-benefit analysis under uncertainty — a note on Weitzman's dismal theorem.* Energy Economics, 42:201–203. DOI: 10.1016/j.eneco.2013.12.013.

Huang, L., Zhou, Y., Han, Y., Hammitt, J. K., Bi, J., and Liu, Y. (2013). *Effect of the Fukushima nuclear accident on the risk perception of residents near a nuclear power plant in China.* Proceedings of the National Academy of Sciences, 110(49):19742–19747. DOI: 10.1073/pnas.1313825110.

Hubert, P., Barny, M. H., and Moatti, J. P. (1991). *Elicitation of decision-makers' preferences for management of major hazards.* Risk Analysis, 11(2):199–206. DOI: 10.1111/j.1539-6924.1991.tb00596.x.

Itaoka, K., Saito, A., Krupnick, A., Adamowicz, W., and Taniguchi, T. (2006). *The effect of risk characteristics on the willingness to pay for mortality risk reductions from electric power generation.* Environmental and Resource Economics, 33(3):371–398. DOI: 10.1007/s10640-005-3605-1.

Jenni, K. and Loewenstein, G. (1997). *Explaining the identifiable victim effect.* Journal of Risk and Uncertainty, 14(3):235–257. DOI: 10.1023/A:1007740225484.

Johansson-Stenman, O. (2008). *Mad cows, terrorism and junk food: Should public policy reflect perceived or objective risks?* Journal of Health Economics, 27(2):234–248. DOI: 10.1016/j.jhealeco.2007.04.004.

Johansson-Stenman, O. and Konow, J. (2010). *Fair air: distributive justice and environmental economics.* Environmental and Resource Economics, 46(2):147–166. DOI: 10.1007/s10640-010-9356-7.

Jones-Lee, M. W. (1992). *Paternalistic altruism and the value of statistical life.* Economic Journal, 102(410):80–90.

Jones-Lee, M. W. and Loomes, G. (1995). *Scale and context effects in the valuation of transport safety.* Journal of Risk and Uncertainty, 11(3):183–203. DOI: 10.1007/BF01207785.

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1986). *Fairness and the assumptions of economics.* The Journal of Business, 59(4). DOI: 10.1086/296367.

Kasperson, R. E., Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., Kasperson, J. X., and Ratick, S. (1988). *The social amplification of risk: a conceptual framework.* Risk Analysis, 8(2):177–187. DOI: 10.1111/j.1539-6924.1988.tb01168.x.

Keeney, R. L. (1980a). *Equity and public risk.* Operations Research, 28(3):527–534. DOI: 10.1287/opre.28.3.527.

Keeney, R. L. (1980b). *Utility functions for equity and public risk.* Management Science, 26(4):345–353. DOI: 10.1287/mnsc.26.4.345.

Keller, L. R. and Sarin, R. K. (1988). *Equity in social risk: Some empirical observations.* Risk Analysis, 8(1):135–146. DOI: 10.1111/j.1539-6924.1988.tb01160.x.

Klinke, A. and Renn, O. (2002). *A new approach to risk evaluation and management: Risk-based, precaution-based, and discourse-based strategies.* Risk Analysis, 22(6):1071–1094. DOI: 10.1111/1539-6924.00274.

Konow, J. (2000). *Fair shares: Accountability and cognitive dissonance in allocation decisions.* American Economic Review, 90(4):1072–1091. DOI: 10.1257/aer.90.4.1072.

Krawczyk, M. and Le Lec, F. (2010). *'Give me a chance!' An experiment in social decision under risk.* Experimental Economics, 13(4):500–511. DOI: 10.1007/s10683-010-9253-5.

Kuran, T. and Sunstein, C. R. (1999). *Availability cascades and risk regulation.* Stanford Law Review, 51(4):683–768.

Kühberger, A. (1998). *The influence of framing on risky decisions: A meta-analysis.* Organizational Behavior and Human Decision Processes, 75(1):23–55. DOI: 10.1006/obhd.1999.2830.

Kühberger, A., Schulte-Mecklenbeck, M., and Perner, J. (1999). *The effects of framing, reflection, probability, and payoff on risk preference in choice tasks.* Organizational Behavior and Human Decision Processes, 78(3):204–231. DOI: 10.1006/obhd.1999.2830.

Laffont, J.-J. and Tirole, J. (1991). *The politics of government decision-making: A theory of regulatory capture.* The Quarterly Journal of Economics, 106(4):1089–1127. DOI: 10.2307/2937958.

Lahno, A. M. and Serra-Garcia, M. (2015). *Peer effects in risk taking: Envy or conformity?* Journal of Risk and Uncertainty, 50(1):73–95. DOI: 10.1007/s11166-015-9209-4.

Lathrop, J. W. (1982). *Evaluating technological risk: prescriptive and descriptive perspectives.* In Kunreuther, H. C. and Ley, E. V., Ed., *The Risk Analysis Controversy: An Institutional Perspective.* Springer.

Lichtenstein, S., Gregory, R., Slovic, P., and Wagenaar, W. A. (1990). *When lives are in your hands: Dilemmas of the societal decision maker.* In Hogarth, R. M., Ed., *Insights in Decision Making: A Tribute to Hillel J. Einhorn*, pages 91–106. University of Chicago Press.

Linde, J. and Sonnemans, J. (2012). *Social comparison and risky choices.* Journal of Risk and Uncertainty, 44(1):45–72. DOI: 10.1007/s11166-011-9135-z.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., and Welch, N. (2001). *Risk as feelings.* Psychological Bulletin, 127(2):267–286. DOI: 10.1037/0033-2909.127.2.267.

López-Vargas, K. (2014). *Risk attitudes and fairness: Theory and experiment.* Working paper. econweb.umd.edu/ lopez-vargas/LopezVargas_JMP_2014.pdf.

Martin, I. W. R. and Pindyck, R. S. (2015). *Averting catastrophes: the strange economics of Scylla and Charybdis.* American Economic Review, 105(10):2947–2985. DOI: 10.1257/aer.20140806.

Milinski, M., Sommerfeld, R. D., Krambeck, H.-J., Reed, F. A., and Marotzke, J. (2008). *The collective-risk social dilemma and the prevention of simulated dangerous climate change.* Proceedings of the National Academy of Sciences, 105(7):2291–2294. DOI: 10.1073/pnas.0709546105.

Millner, A. (2013). *On welfare frameworks and catastrophic climate risks.* Journal of Environmental Economics and Management, 65(2):310–325. DOI: 10.1016/j.jeem.2012.09.006.

Myerson, R. B. (1981). *Utilitarianism, egalitarianism, and the timing effect in social choice problems.* Econometrica, 49(4):883–897. DOI: 10.2307/1912508.

Nordhaus, W. D. (2011). *The economics of tail events with an application to climate change.* Review of Environmental Economics and Policy, 5(2):240–257. DOI: 10.1093/reep/rer004.

Olivola, C. Y. and Sagara, N. (2009). *Distributions of observed death tolls govern sensitivity to human fatalities.* Proceedings of the National Academy of Sciences, 106(52):22151–22156. DOI: 10.1073/pnas.0908980106.

Olson, M. (1965). *The Logic of Collective Action: public goods and the theory of groups.* Harvard University Press. 176 pages.

Olson, M. (1971). *The Logic of collective action: Public goods and the theory of groups.* Harvard University Press, 2nd edition. ISBN: 978-0674537514.

Paté-Cornell, E. (2002). *Risk and uncertainty analysis in government safety decisions.* Risk Analysis, 22(3):633–646. DOI: 10.1111/0272-4332.00043.

Peterson, M. (2002). *The limits of catastrophe aversion.* Risk Analysis, 22(3):527–538. DOI: 10.1111/0272-4332.00036.

Pindyck, R. S. (2011). *Fat tails, thin tails, and climate change policy.* Review of Environmental Economics and Policy, 5(2):258–274. DOI: 10.1093/reep/rer005.

Pindyck, R. S. and Wang, N. (2013). *The economic and policy consequences of catastrophes.* American Economic Journal: Economic Policy, 5(4):306–339. DOI: 10.1257/pol.5.4.306.

Pollak, R. A. (1998). *Imagined risks and cost-benefit analysis.* American Economic Review, 88(2):376–380. www.jstor.org/stable/116951.

Portney, P. R. (1992). *Trouble in Happyville.* Journal of Policy Analysis and Management, 11(1):131–132. DOI: 10.2307/3325137.

Posner, R. A. (2004). *Catastrophe: risk and response.* Oxford University Press. ISBN: 978-0195306477.

Pratt, J. W. and Zeckhauser, R. J. (1996). *Willingness to pay and the distribution of risk and wealth.* Journal of Political Economy, 104(4):747–763. DOI: 10.1086/262041.

Prelec, D. (1998). *The probability weighting function.* Econometrica, 66(3):497–527. DOI: 10.2307/2998573.

Ramsberg, J. A. L. and Sjöberg, L. (1997). *The cost-effectiveness of lifesaving interventions in Sweden.* Risk Analysis, 17(4):467–478. DOI: 10.1111/j.1539-6924.1997.tb00887.x.

Rheinberger, C. M. (2010). *Experimental evidence against the paradigm of mortality risk aversion.* Risk Analysis, 30(4):590–604. DOI: 10.1111/j.1539-6924.2009.01353.x.

Rohde, I. M. T. and Rohde, K. I. M. (2011). *Risk attitudes in a social context.* Journal of Risk and Uncertainty, 43(3):205–225. DOI: 10.1007/s11166-011-9127-z.

Rothschild, M. and Stiglitz, J. E. (1970). *Increasing risk: I. a definition.* Journal of Economic Theory, 2(3):225–243. DOI: 10.1016/0022-0531(70)90038-4.

Salanié, F. and Treich, N. (2009). *Regulation in Happyville.* The Economic Journal, 119(537):665–679. DOI: 10.1111/j.1468-0297.2009.02221.x.

Schelling, T. C. (1968). *The life you save may be your own.* In Chase, S. B., Ed., *Problems in Public Expenditure Analysis*, pages 143–144. Brookings.

Sen, A. (1970). *Interpersonal aggregation and partial comparability.* Econometrica, 38(3):393–409. DOI: 10.2307/1909546.

Slovic, P. (1987). *Perception of risk.* Science, 236(4799):280–285. DOI: 10.1126/science.3563507.

Slovic, P. (2007). *"If I look at the mass I will never act": Psychic numbing and genocide.* Judgment and Decision Making, 2(2):79–95.

Slovic, P., Fischhoff, B., and Lichtenstein, S. (2000). *Facts and fears: Understanding perceived risk.* In Slovic, P., Ed., *The Perception of Risk*, pages 137–154. Earthscan.

Slovic, P., Lichtenstein, S., and Fischhoff, B. (1984). *Modeling the societal impact of fatal accidents.* Management Science, 30(4):464–474. DOI: 10.1287/mnsc.30.4.464.

Small, D. A. and Loewenstein, G. (2003). *Helping a victim or helping the victim: Altruism and identifiability.* Journal of Risk and Uncertainty, 26(1):5–16. DOI: 10.1023/A:1022299422219.

Solow, R. M. (1972). *Notes on "doomsday models".* Proceedings of the National Academy of Sciences of the United States of America, 69(12):3832–3833. DOI: 10.1073/pnas.69.12.3832.

Stallen, P. J. M., Geerts, R., and Vrijling, H. K. (1996). *Three conceptions of quantified societal risk.* Risk Analysis, 16(5):635–644. DOI: 10.1111/j.1539-6924.1996.tb00813.x.

Starr, C. (1969). *Social benefit versus technological risk: what is our society willing to pay for safety?* Science, 165(3899):1232–1238. DOI: 10.1126/science.165.3899.1232.

Stewart, N., Chater, N., Stott, H. P., and Reimers, S. (2003). *Prospect relativity: How choice options influence decision under risk.* Journal of Experimental Psychology: General, 132(1):23–46. DOI: 10.1037/0096-3445.132.1.23.

Sunstein, C. R. (2004). *Valuing life: A plea for disaggregation.* Duke Law Journal, 54(2):385–445.

Sunstein, C. R. (2005). *Laws of fear: Beyond the precautionary principle.* Cambridge University Press. ISBN: 978-0521615129, 248 pages.

Sunstein, C. R. (2009). *Worst-case scenarios.* Harvard University Press. ISBN: 978-0674032514, 352 pages.

Taleb, N. N. (2007). *The Black Swan: The Impact of the Highly Improbable.* Random House. ISBN: 978-1400063512.

Tavoni, A., Dannenberg, A., Kallis, G., and Löschel, A. (2011). *Inequality, communication, and the avoidance of disastrous climate change in a public goods game.* Proceedings of the National Academy of Sciences, 108(29):11825–11829. DOI: 10.1073/pnas.1102493108.

Tengs, T. O., Adams, M. E., Pliskin, J. S., Safran, D. G., Siegel, J. E., Weinstein, M. C., and Graham, J. D. (1995). *Five-hundred lifesaving interventions and their cost-effectiveness.* Risk Analysis, 15(3):369–390. DOI: 10.1111/j.1539-6924.1995.tb00330.x.

Treich, N. (2010). *The value of a statistical life under ambiguity aversion.* Journal of Environmental Economics and Management, 59(1):15–26. DOI: 10.1016/j.jeem.2009.05.001.

Tversky, A. and Kahneman, D. (1981). *The framing of decisions and the psychology of choice.* Science, 211(4481):453–458. DOI: 10.1126/science.7455683.

USNRC (1975). *NUREG-75/014 (WASH-1400) Reactor safety study, an assessment of accident risks.* Technical report, US Nuclear Regulatory Commission. a.k.a. "the Rasmussen Report" (superseded by NUREG-1150). DOI: 10.2172/7134131.

Viscusi, W. K. (1998). *Rational Risk Policy — The 1996 Arne Ryde Memorial Lectures.* Oxford University Press. ISBN: 978-0198293637, 160 pages.

von Hippel, F. N. (2011). *It could happen here.* New York Times, March 24:A31.

Västfjäll, D., Slovic, P., and Mayorga, M. (2015). *Pseudoinefficacy: Negative feelings from children who cannot be helped reduce warm glow for children who can be helped.* Frontiers in Psychology, 6:616. DOI: 10.3389/fpsyg.2015.00616.

Wakker, P. and Deneffe, D. (1996). *Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown.* Management Science, 42(8):1131–1150. DOI: 10.1287/mnsc.42.8.1131.

Weitzman, M. L. (2009). *On modeling and interpreting the economics of catastrophic climate change.* The Review of Economics and Statistics, 91(1):1–19. DOI: 10.1162/rest.91.1.1.

Wiener, J. B. (2016). *The tragedy of the uncommons: On the politics of apocalypse.* Global Policy, 7(S1):67–80. DOI: 10.1111/1758-5899.12319.

Zeckhauser, R. (1996). *The economics of catastrophes.* Journal of Risk and Uncertainty, 12(2-3):113–140. DOI: 10.1007/BF00055789.

You can extract these bibliographic entries in BibTeX format by clicking on the paperclip icon.

You can download this document, and others in the *Cahiers de la Sécurité Industrielle* collection, from FonCSI's web site.

FONCSI
Fondation pour une culture
de sécurité industrielle

Foundation for an Industrial Safety Culture
a public interest research foundation

www.FonCSI.org

6 allée Émile Monso – BP 34038
31029 Toulouse cedex 4
France

Twitter: @TheFonCSI
Email: contact@FonCSI.org

ISSN 2100-3874

FONCSI
Fondation pour une culture de sécurité industrielle

6 allée Émile Monso
ZAC du Palays - BP 34038
31029 Toulouse cedex 4

www.foncsi.org