Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Model Uncertainty in Risk Analysis and Decision Theory: A Preliminary Investigation

Emanuele Borgonovo, Veronica Cappelli, Fabio Maccheroni, Massimo Marinacci

Working Paper n. 592

This Version: November, 2016

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Model Uncertainty in Risk Analysis and Decision Theory: A Preliminary Investigation*

Emanuele Borgonovo[#] Veronica Cappelli[§] Fabio Maccheroni[#] Massimo Marinacci[#]

[#]Department of Decision Sciences and IGIER, Università Bocconi
[§]Department of Economics and Decision Sciences, HEC Paris

November 2016

Abstract

The purpose of this note is to discuss the relation between model uncertainty in risk analysis and decision theory.

1 Introduction

Having different focuses, the fields of risk analysis and decision theory have been evolving in parallel for the last decades. While the former specialized in risk assessment, risk characterization, and risk communication toward the end of understanding and managing risk, the latter specialized in the study of decision criteria for choice under uncertainty and of their different descriptive and prescriptive implications. Recently, it emerged in risk analysis the demand for decision criteria that go beyond expected loss. At the same time, decision theory faces the problem of implementing a more sophisticated quantification of uncertainties involved in the newly developed criteria.

Our analysis aims at connecting the two fields thus allowing to share their new developments and leading to a modern decision analysis where a risk informed risk analysis becomes the input of applied decision theory.

While a first step in this direction was taken by Borgonovo et al. (2016), this chapter focuses on the issue of model uncertainty. In particular, the purpose of this note is to discuss the relation between model uncertainty in the risk analysis setup and in modern decision theory. We will then sketch how novel criteria in decision theory can be meaningfully applied in decision analysis contexts when different specifications of model uncertainty are involved.

2 Model uncertainty in risk analysis and decision theory

Decision theory aims at modeling decision situations, both from a descriptive and prescriptive viewpoint. Some of its recent developments focused on the issue of model uncertainty

*The authors gratefully acknowledge the financial support of the European Research Council (INDIMACRO) and of the AXA Research Fund.

(see Marinacci, 2015, for a review). The typical decision situation features a decision maker who has to choose an action, among those available, whose consequences depend on uncertain factors beyond his control. Canonically, all these ingredients (alternatives, uncertain factors, consequences, and the relation between them) are a datum of the decision problem, together with the information, conditional on which decisions are taken. Uncertain factors are usually called states of the world (or of the environment, or of nature), and the possible mechanisms (or laws) that generate states are represented by a set of probability measures over them. In the language of decision theory, these probability measures are called *probability models* (see Marinacci, 2015, page 1000) and represent the posited intrinsic variability of states. When the latter set is not a singleton, there is *probabilistic model uncertainty*, or model uncertainty from a decision theory viewpoint (sometimes called model ambiguity). In this sense, model uncertainty refers to uncertainty about the stochastic nature of states' realizations. Moreover, often, the degree of confidence of the decision maker about these models is represented by a *prior probability* over them.

As de Finetti (1971) writes:

> A scientific theory, in the sense of a law, is not a statement whose truth or falsity is objectively decidable. It seems therefore reasonable to analyse its validity, from this point of view, with probabilistic arguments, necessarily subjective.

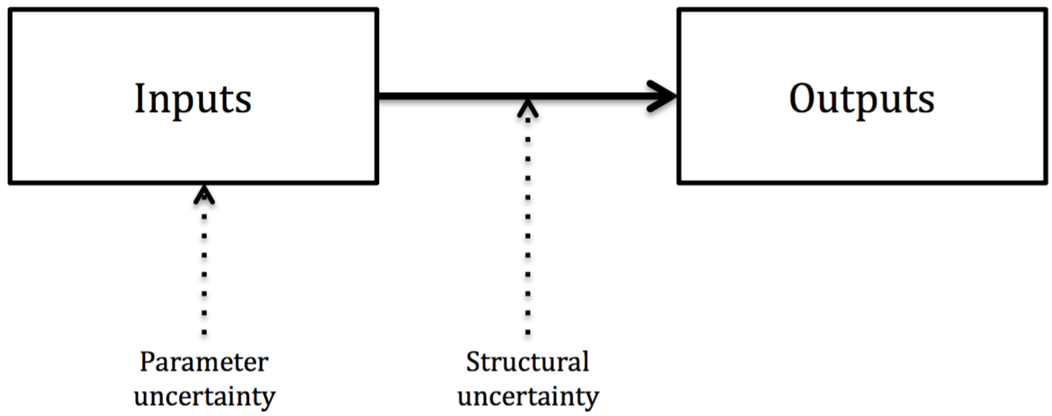
Summing up, model uncertainty in decision theory is typically the object of two probabilistic layers of analysis: a first layer featuring probability models on states that quantify inherent randomness, and a second layer characterized by a prior probability quantifying model uncertainty on probability models.

In risk analysis, a model is a (often mathematical) representation of reality which captures some of its aspects relevant to a specific objective, for instance a decision or a prediction. It is usually in the form of a mapping from input variables to output variables, and it is intended to represent relationships between quantities in a system.[1] Model uncertainty, from a risk analysis viewpoint, is uncertainty about this representation of reality. See, e.g., Apostolakis (1994), Zio and Apostolakis (1996), Apostolakis (1999), Nielsen and Aven (2003). There is some consensus about the fact that there are two qualitatively different kinds of model uncertainty: uncertainty in the output caused by uncertainty in the functional form of the model is referred to as *model structure uncertainty*; uncertainty in the output caused by uncertainty in the input variables is referred to as *parameter uncertainty.*[2]

Another distinction about which there is consensus is the one between stochastic and deterministic models. In a *deterministic model*, each input vector is associated to a unique value of the output, while in a *stochastic model* an input vector is mapped into a distribution of output values. Thus deterministic models can be seen as special cases of stochastic models. This classification is independent of the previous one and they might well coexist in the analysis. In turn, all the models used in risk analysis usually consist of a multiple submodels, describing the system at different levels, and uncertainty may affect any of these levels.

[1] The input and output variables correspond to the various quantities in the system itself.

[2] However, as observed by Apostolakis and Wu (1993), and Buslik (1994), in some cases model structure uncertainty can be reframed as as parameter uncertainty.



As we will see, the relation between model uncertainty in decision theory and in risk analysis is not completely obvious.[3] We will expand on this point in the following sections. Specifically, we will illustrate how model uncertainty in risk analysis can be accommodated into a decision theoretic framework. This will be done by putting into correspondence the constituents of risk analysis and those of decision theory. Then we will briefly show how this allows for the treatment of model uncertainty in a decision analysis. Furthermore, we will illustrate our techniques with an example, and we will conclude with a discussion.

3 Distributions

A *(discrete probability) distribution* on a set Z is a function

$$\begin{array}{rccl} p: & Z & \to & [0,1] \\ & z & \mapsto & p(z) \end{array}$$

such that $p(z) \neq 0$ for finitely many elements of Z and $\sum_{z \in Z} p(z) = 1$. The value $p(z)$ is the probability of the singleton set $\{z\}$. Denoting by $\{z_i\}_{i \in I}$ a finite subset of Z such that $\sum_{i \in I} p(z_i) = 1$, we often write $p = \langle p_i, z_i \rangle_{i \in I}$. For every subset Y of Z, $p(Y) = \sum_{y \in Y} p(y)$ is the probability of the set Y.

When Z is an interval of real numbers a *decumulative probability distribution* is a left continuous decreasing function p with supremum value 1 and infimum value 0. The value $p(z)$ is the probability of the set $\{y \in Z : y \geq z\}$. For every sub-interval $Y = [a, b)$ of Z, $p(Y) = p(a) - p(b)$ is the probability of the set Y.[4]

4 Risk triplets in a risk analysis setup

We adopt the description of risk proposed by Kaplan and Garrick (1981). Among those proposed in the SRA Glossary (2015) this one represents a good compromise between simplicity and flexibility, and it is widely adopted in applications of risk analysis (e.g. by the US NRC).

[3] In particular, it depends on the definition of state of the world or of the environment and on the requirements about states' observability.

[4] The probability of any sub-interval can then be computed because the probability of the singleton set $\{z\}$ is $p(z) - p(z^+)$.

Kaplan and Garrick informally define *hazard* as "a source of danger" and *risk* as the "possibility of loss or injury" and "the degree of probability of such loss". They then carry out a risk analysis in terms of triplets

$$\langle S_i, \ell_i, x_i \rangle \tag{1}$$

each of the components consisting of an answer to the following questions:

1. "What can happen?" The answer identifies a *scenario* S_i.

2. "How likely is it that it will happen?" The answer indicates the *likelihood* $\ell_i = \ell(S_i)$ of scenario S_i.

3. "If it does happen, what are the consequences?" The answer is the *consequence* of scenario S_i.

Hazard is then formally defined as the set of doublets

$$H = \langle S_i, x_i \rangle_{i \in I}$$

while risk is defined as the set of triplets

$$R = \langle S_i, \ell_i, x_i \rangle_{i \in I}.$$

They consider three formats that capture and quantify, the concept of "likelihood."

Format 1. (Frequency)

This applies when we have a repetitive situation, and we ask, "How frequently does scenario i occur?" In this case the likelihood is expressed as a frequency $\ell_i = \phi(S_i)$, abbreviated ϕ_i, and risk becomes $R = \langle S_i, \phi_i, x_i \rangle_{i \in I}$.

Format 2. (Probability)

When the situation is "one shot", like a mission to Mars, we want to quantify our degree of confidence that the mission will succeed. In this case likelihood is expressed as a probability $\ell_i = p(S_i)$, abbreviated p_i, and the triplets become $R = \langle S_i, p_i, x_i \rangle_{i \in I}$.

Format 3. (Probability of Frequency)

The third format applies when we have a repetitive situation, or can imagine one as a thought experiment, so that the frequency exists, but since we haven't done the experiment we are uncertain about what that frequency would be. We therefore express our state of knowledge about that frequency with a distribution $\langle \rho_\theta, \phi_\theta \rangle_{\theta \in \Theta}$. We call this the "Probability of Frequency" format, here $\ell_i = \langle \rho_\theta, \phi_\theta(S_i) \rangle$, abbreviated $\langle \rho_\theta, (\phi_\theta)_i \rangle_{\theta \in \Theta}$, and

$$R = \left\langle S_i, \langle \rho_\theta, (\phi_\theta)_i \rangle_{\theta \in \Theta}, x_i \right\rangle_{i \in I}. \tag{2}$$

Notice that to any such risk corresponds to a family of risks $R_\theta = \langle S_i, (\phi_\theta)_i, x_i \rangle_{i \in I}$ in frequency format paired with a distribution ρ on Θ, that is, $R = \langle \rho_\theta, R_\theta \rangle_{\theta \in \Theta}$.

We hereby consider the generalized version of Format 3 proposed by Kaplan (1997) in which $x_i = x(S_i)$ is a random outcome, formally a conditional distribution over final outcomes in a set Z (where conditioning is made on scenarios). The model uncertainty analysis of these risks is better understood by decomposing each risk in two components

$$\langle S_i, \ell_i, x_i \rangle_{i\in I} \equiv \langle S_i, \ell_i \rangle_{i\in I} \wedge \langle S_i, x_i \rangle_{i\in I} .$$

The first component $\langle S_i, \ell_i \rangle_{i\in I}$ is a *model of scenario realization* saying that scenario S_i occurs with likelihood $\ell_i = \langle \rho_\theta, \phi_\theta (S_i) \rangle$; the second is $\langle S_i, x_i \rangle_{i\in I}$ is a *conditional model of final outcomes realization* saying that conditional on the realization of scenario S_i, the final outcome z occurs with likelihood $x_i (z) = x (z \mid S_i)$.

In this perspective, the types of model uncertainty in risk analysis that we discussed in the introduction emerge in a simple and compelling way. For every θ, ϕ_θ is an aleatory model of the realizations of $S_1, ..., S_N$, and epistemic uncertainty about this stochastic model is represented by ρ. Thus, according to the previous classification, ρ captures state of knowledge (= epistemic) parameter uncertainty while each ϕ_θ captures stochastic (= aleatory) parameter uncertainty.

For every i, given S_i, the distribution x_i of final outcomes represents the uncertainty about the relation between the inputs and the outputs. This uncertainty may result from two sources. Either, the model of final outcome realization is aleatory and thus the input-output relation becomes probabilistic, or the same model is deterministic, but its specific structure is only partially known and this again induces a probabilistic relation between inputs and outputs, or both. In other words, the probabilistic nature of x_i captures model structure uncertainty.

If one wants to distinguish the aleatory and the epistemic components of model structure uncertainty, it is necessary to augment the risk triplet format by considering triplets of the form

$$\left\langle S_i, \langle \rho_\theta, (\phi_\theta)_i \rangle_{\theta\in\Theta}, \langle \pi_\gamma, (x_\gamma)_i \rangle_{\gamma\in\Gamma} \right\rangle_{i\in I} = \left\langle S_i, \langle \rho_\theta, \phi_\theta (S_i) \rangle_{\theta\in\Theta} \right\rangle_{i\in I} \wedge \left\langle S_i, \langle \pi_\gamma, x_\gamma (S_i) \rangle_{\gamma\in\Gamma} \right\rangle_{i\in I} \quad (3)$$

where each x_γ is a scenario-conditional aleatory model of the outcome,[5] and π is an epistemic probability model representing the degree of confidence in x_γ. Following the distinction we introduced above, the first case (aleatory structural model uncertainty without epistemic structural model uncertainty) corresponds to a degenerate π,[6] the second (deterministic structural model with epistemic structural model uncertainty) to degenerate x_γ's, and the third (aleatory structural model uncertainty without epistemic structural model uncertainty) to nondegenerate π and x_γ's.

Also in this case a family of (aleatory) triplets

$$R_{(\theta,\gamma)} = \langle S_i, \phi_\theta (S_i), x_\gamma (S_i) \rangle_{i\in I}$$

emerges and it is paired with a (joint epistemic) distribution (ρ, π) on $\Theta \times \Gamma$. The final triplet can thus be written as, $R = \left\langle \rho_\theta \times \pi_\gamma, R_{(\theta,\gamma)} \right\rangle_{\theta\in\Theta,\gamma\in\Gamma}$.

Clearly, both the models of scenario generation and those of scenario-outcome propagation, can consist of several submodels each entailing different levels of uncertainty, as we will exemplify below. At this level of abstraction the distinction fades.

5 Lottery acts in a decision theory setup

In this section we relate the Anscombe and Aumann (1963) framework to the Kaplan and Garrick setup presented above. Following Anscombe and Aumann, a decision problem under

[5] For each S_i, the final outcome z occurs with likelihood $x_\gamma (z \mid S_i)$.

[6] And we are back to the formulation (2) of Kaplan (1997).

uncertainty features a decision maker who has to choose among a set of actions whose final outcomes depend stochastically on uncertain factors beyond his control, called states of the environment.[7] Formally, for every action d, $f^d(s)$ is the distribution of final outcomes in a set Z resulting from the choice of d and conditional on the realization of state s in a measurable space S and, the (simple and measurable) function

$$\begin{array}{rccc} f^d: & S & \rightarrow & X \\ & s & \mapsto & f^d(s) \end{array}$$

is called Anscombe and Aumann act. This definition implies that every act can be written as

$$f^d(s) = \begin{cases} x_1 & s \in S_1 \\ x_2 & s \in S_2 \\ \ldots & \ldots \\ x_N & s \in S_N \end{cases}$$

where $\{x_1, x_2, ..., x_N\}$ is a finite set of random outcomes and $\{S_1, S_2, ..., S_N\}$ is a finite measurable partition of S. This expression of acts immediately delivers a one to one correspondence with hazards given by

$$f^d \equiv \langle S_i, x_i^d \rangle_{i \in I} = H^d.$$

In the perspective of risk analysis, H^d is the hazard corresponding to implementing action d. From the discussion of the previous section, it follows that f^d can be seen as representing the structural model of action d conditional on the realization of states. As we discussed above it can be desirable to disentangle the epistemic and aleatory parts of structural model uncertainty. This can be achieved by considering a distribution over acts

$$\langle \pi_\gamma^d, f_\gamma^d \rangle_{\gamma \in \Gamma^d} \equiv \left\langle S_i, \left\langle \pi_\gamma^d, \left(x_\gamma^d\right)_i \right\rangle_{\gamma \in \Gamma^d} \right\rangle_{i \in I} \tag{4}$$

formally, these are the objects of choice considered by Anscombe and Aumann (1963), often called *lottery acts*. They correspond to the output component of (3).

As for the modelling of states' generation, Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2013) enrich the Anscombe and Aumann framework with a family of probability measures $\{\phi_\theta\}_{\theta \in \Theta}$ on S, that describe the intrinsic variability of states: aleatory (state) uncertainty. Which paired with a distribution ρ over Θ gives a full description $\langle \rho_\theta, \phi_\theta \rangle_{\theta \in \Theta}$ of the uncertainty concerning states.[8] The relation with the framework of Kaplan and Garrick is now finalized by noticing that $\langle \rho_\theta, \phi_\theta \rangle_{\theta \in \Theta}$ naturally induces likelihoods $\ell_i = \langle \rho_\theta, \phi_\theta(S_i) \rangle$ on scenarios, therefore yielding the doublets

$$\langle S_i, \langle \rho_\theta, (\phi_\theta)_i \rangle_{\theta \in \Theta} \rangle_{i \in I}. \tag{5}$$

which together with (3) associates with every lottery act a risk in the form (3).

[7] More on states and the implications of their definition on model uncertainty in the discussion below.

[8] Cerreia-Vioglio et al. (2013) use different letters for their objects, e.g., $\{\phi_\theta\}_{\theta \in \Theta}$ is denoted by M and ρ is called μ. We translitterated them in order to ease comparison with the risk triplets setup.

Summing up, we have the following glossary:

KG triplets	AA acts
$H^d = \langle S_i, x_i^d \rangle_{i\in I}$	$f^d(s) = x_i \quad s \in S_i,\ i = 1, ..., N$
$\left\langle S_i, \left\langle \pi_\gamma^d, \left(x_\gamma^d\right)_i \right\rangle_{\gamma\in\Gamma^d} \right\rangle_{i\in I}$	$\left\langle \pi_\gamma^d, f_\gamma^d \right\rangle_{\gamma\in\Gamma^d}$
$R^d = \left\langle S_i, \langle \rho_\theta, (\phi_\theta)_i \rangle_{\theta\in\Theta}, \left\langle \pi_\gamma^d, \left(x_\gamma^d\right)_i \right\rangle_{\gamma\in\Gamma^d} \right\rangle_{i\in I}$	$\left\langle \pi_\gamma^d, f_\gamma^d \right\rangle_{\gamma\in\Gamma^d} \wedge \langle \rho_\theta, \phi_\theta \rangle_{\theta\in\Theta}$

(6)

Notice that ρ can be naturally seen as a prior over the set of probabilistic models $\{\phi_\theta\}_{\theta\in\Theta}$, and in this sense probabilistic model uncertainty can be seen as a special case of parameter uncertainty. But, as we will discuss in the final section, this depends on the adoption of the modelling framework of Anscombe and Aumann. By chosing the setup of Savage with a larger state space and no uncertainty in the description of outcome realization conditional on states, the converse relation can be obtained. In other words, it is always possible to translate one formulation of model uncertainty into the other. This further facilitates the communication between the two fields.

5.1 Decision theoretic risk evaluation

We conclude this section by showing how our analysis can be used to enrich a decision analysis under model uncertainty with state-of-the-art decision theoretic tools. The extra flexibility granted by the use of decision criteria more sophisticated than expected loss, allow to take into account different facets of risk under model uncertainty. Assume a risk analysis, for example, like the one in the example below on the renovation of a nuclear power plant, has produced risk triplets $R^d = \left\langle S_i, \langle \rho_\theta, (\phi_\theta)_i \rangle_{\theta\in\Theta}, \left\langle \pi_\gamma^d, \left(x_\gamma^d\right)_i \right\rangle_{\gamma\in\Gamma^d} \right\rangle_{i\in I}$ corresponding to different renovation projects $d \in D$. To every such risk, it corresponds a lottery act by means of (6), and therefore, when the time of choosing among the different projects comes, the decision maker can consider all the criteria developed in decision theory in order to make a decision (see, e.g., Cerreia-Vioglio et al., 2013). A tractable, yet powerful example is the two-stage decision criterion introduced by Klibanoff, Marinacci, and Mukerji (2005). Without epistemic structural model uncertainty it has the form

$$V\left(R^d\right) = V\left(f^d \wedge \langle \rho_\theta, \phi_\theta \rangle_{\theta\in\Theta}\right) = \sum\nolimits_{\theta\in\Theta} \rho_\theta v\left(\sum\nolimits_{i\in I} \phi_\theta\left(S_i\right) u\left(x_i^d\right)\right) \tag{7}$$

where we can interpret u as capturing attitudes toward aleatory uncertainty and v as capturing risk attitudes towards epistemic uncertainty (as discussed in Marinacci, 2015). If, in addition, there is epistemic structural model uncertainty the previous criterion becomes

$$V\left(R^d\right) = V\left(\left\langle \pi_\gamma^d, f_\gamma^d \right\rangle_{\gamma\in\Gamma^d} \wedge \langle \rho_\theta, \phi_\theta \rangle_{\theta\in\Theta}\right) = \sum\nolimits_{\gamma\in\Gamma^d} \pi_\gamma^d \sum\nolimits_{\theta\in\Theta} \rho_\theta v\left(\sum\nolimits_{i\in I} \phi_\theta\left(S_i\right) u\left(\left(x_\gamma^d\right)_i\right)\right)$$

which is the version of (7) studied by Seo (2009).

Adopting the more pragmatical view of Winkler (1996), these representations allow for the decomposition of the different sources of uncertainty so to distinguish their role and possibly target them in order to reduce vulnerabilities at specific levels. Specifically, the functions u and

v describe how the different sources of uncertainty impact the overall evaluation of alternative d.

As anticipated, this is only but one example. Our translation leads the way to the direct use of the vast majority of recent decision criteria under uncertainty developed in decision theory (see e.g. Gilboa and Marinacci, 2013, for a comprehensive review). Importantly, besides model uncertainty, many more considerations can be embedded in the analysis, ranging from robustness concerns to behavioral traits, as discussed in Borgonovo et al. (2016).

6 An example of seismic risk analysis

In this section we will illustrate the use of our translation in the context of a bona fide risk analysis.

Kaplan, Perla and Bley (1983, henceforth KPB) study the process of a first level probabilistic risk analysis for a nuclear power plant. The analysis is carried on according to the following steps:

1. Seismicity: description of the likelihood of ground motions at the plant location.
2. Fragility of components: description of the likelihood of the failure of the single components conditional on different ground motions.
3. Plant logic: identification of component failures which would lead to failure of the plant (here, core melt).
4. Plant level fragilities: computation of the likelihood of plant failure.

Seismicity is described by means of crustal models that take the form of cumulative distribution functions, that identify the probability of exceedance. Here is the first place where model uncertainty can already be found: it translates into uncertainty about the specific form (or the parameters) of the distribution function.

Fragility of a given component is described by a fragility curve associating to each level of ground acceleration the corresponding failure fraction of the component. Also here model uncertainty can bite and, in particular, may regard the specific form (or the parameters) of the fragility curve of the component.

Plant logic analysis typically involves boolean modeling. Of course there can be uncertainty also at this level, but, as KPB, we will not consider it.

In the final step, the families of fragility curves obtained at step 2 are used to obtain the plant level fragility curve, for a given plant logic. Also here we can find model uncertainty. It can be at the very least inherited by the one about components, but can also descend from uncertainty in the way in which fragility curves are combined (e.g. their correlation structure).

Notice that while the uncertainty regarding ground motion does not depend on the design of the plant, i.e. its components and their combination, the one about the plant fragility curves obviously does, because as components change also the families of their fragility curves do.

States and eartquake models As commonly done, KPB describe earthquakes intensities as peak ground accelerations $s \in \mathbb{R}_+$. The state space is thus $S = \mathbb{R}_+$. A crustal dynamics model produces a decumulative probability distribution $\phi : S \to [0,1]$ called *seismicity curve.* At any point $s \in S$, the value $\phi(s)$ represents the frequency of exceedance, that is, the frequency with which earthquakes of acceleration s or greater occur at the plant location per year.

Yet, our knowledge of the crustal dynamics that determine peak ground accelerations is still incomplete. For example, the information available to the risk analyst, say recorded seismic history and geologic data, prevents him from knowing the curve ϕ with great accuracy, and this is expressed by considering a family $\{\phi_\theta : \theta \in \Theta\}$ of possible curves, the set θ may be seen as the set of alternative geophysical models, say different crustal dynamics equations. These, in turn, may determine a family $\{\phi_\theta\}_{\theta\in\Theta}$ of Pareto distributions.

The risk analyst's degree of confidence in the various seismicity curves can thus be described by another probability distribution ρ on Θ.

Anscombe-Aumann acts and consequences The design decision d_k about a specific component k (for example, the choice of a given kind of service water pump or of shear wall) determines the *component's fragility curve* $f^{d_k} : S \to [0,1]$, describing for every peak ground acceleration s the probability $f^{d_k}(s)$ of failure of that specific component. The probability of meltdown in each state s is then determined, by plant logic analysis, by the vector $d = (d_1, ..., d_K)$ of all design decisions that constitutes the *plant design* d.

Let's consider a toy case in which the only design decisions are $d_1 =$"service water pumps" and $d_2 =$"shear wall", the *frequency of plant failure*, that is, the *plant's fragility curve*, in state s is given by

$$f^d(s) = f^{(d_1,d_2)}(s) = 1 - \left(1 - f^{d_1}(s)\right)\left(1 - f^{d_2}(s)\right) = f^{d_1}(s) + f^{d_2}(s) - f^{d_1}(s) f^{d_2}(s).$$

In fact, core damage can be caused either by failure of component 1 or failure of component 2 here assumed to be independent. At this point, note that measurement of (the continuum of) peak ground accelerations is discrete and therefore both f^{d_1} and f^{d_2} are increasing staircase functions, and so it is $f^d(s)$. In other words, $f^d(s)$ is the *act (à la Anscombe-Aumann)* that corresponds to nuclear plant design d, whose (first level) *consequences* are failure fractions (meltdown probabilities) and depend on both the design decisions and the intensities of earthquakes in the plant location. Such intensities are thus the relevant *states* of the environment.

In general, for each plant design d, there exist $0 = s_1 \leq s_2 \leq ... \leq s_N < \infty$ such that, setting $x_n^d = f^d(s_n)$ for all $n = 1, ..., N$, we have

$$f^d(s) = \begin{cases} x_1^d & s \in [s_1, s_2) = S_1 \\ x_2^d & s \in [s_2, s_3) = S_2 \\ \dots & \dots \\ x_N^d & s \in [s_N, \infty) = S_N \end{cases} \tag{8}$$

where S_i is the *scenario* in which an earthquake of intensity $s \in [s_i, s_{i+1})$ occurs and x_i^d is the consequent failure probability. As anticipated, *consequences* are failure probabilities, hence the consequence set can be modelled as $X = [0,1]$.

This shows how the plant design determines and Anscombe and Aumann act $f^d : S \to [0,1]$ or equivalently by the hazard

$$H^d = \langle S_i, x_i^d \rangle_{i\in I}.$$

As it happens for seismicity curves also the determination of frequency of seismic core melt conditional on the design d and the peak ground acceleration is difficult. Two possible reasons are the some degree of variability of the fragility estimation of the single components and the complexity of the system under consideration. In the toy example above, it may be the case that component failure are not independent, but the depence structure is unknown. Notice that KPB also consider the fact that the fragility curves of the single component might be themselves subject to uncertainty.

In any case this leads to the fact that to each plant design d is associated a family $\left\{f_\gamma^d : \gamma \in \Gamma^d\right\}$ of possible plant fragility curves (and correspondingly a family of acts/hazards). Again, the risk analyst's degree of confidence in the various curves can be described by a probability distribution π^d on Γ^d.

Notice that both π_γ^d, the degree of confidence in the plant fragility curve being f_γ^d, and the plant fragility curves f_γ^d, which determine the probability of failure of the system, depend on d. In fact, different designs can lead to different probabilities of failures of the system under different plant logic hypotheses. Indeed, since d describes a vector of components of the plant, whenever we modify a component this has an impact, for a given logic model, upon the degree of confidence in the plant fragility curve being f_γ^d, that is π_γ^d.

For example, if Γ^d is finite and $\pi^d = \left\{\langle \pi_\gamma^d, \gamma \rangle\right\}$, then a lottery

$$\left(f_\gamma^d, \pi_\gamma^d\right)_{\gamma \in \Gamma^d}$$

of Anscombe-Aumann acts is obtained, the corresponding risk analysis object is a stochastic hazard

$$\langle \pi_\gamma^d, H_\gamma^d \rangle_{\gamma \in \Gamma^d} = \left\langle S_n, \left\langle \pi_\gamma^d, \left(x_\gamma^d\right)_i \right\rangle_{\gamma \in \Gamma^d} \right\rangle_{i \in I}.$$

Finally, passing from hazards to risks we obtain

$$\left\langle S_i, \langle \rho_\theta, (\phi_\theta)_i \rangle_{\theta \in \Theta}, \left\langle \pi_\gamma^d, \left(x_\gamma^d\right)_i \right\rangle_{\gamma \in \Gamma^d} \right\rangle_{i \in I}.$$

The analysis performed above pertains to a first level probabilistic risk assessment. The estension to the subsequent levels is of course doable. It is both conceptually and mathematically straightforward but notationally heavy. It basically consists in enriching the consequence space with additional layers of uncertainty regarding the relation between core damage and radioactive release (level two) and between radioactive release and effects on the population (level three).

7 Discussion

In this section we discuss some issues raised in the literature about scenarios and models.

Scenarios

Each scenario in the risk analysis setup is an event in (that is, a subset of) the state space S of the decision theoretic setup (see Anscombe and Aumann, 1963 and Savage, 1954). A matter which may be worth considering is whether, in general, scenarios form an exclusive and exhaustive list of the events that are consequence relevant. This question has two relevant implications: one conceptual and that goes at the hearth of the theory and one practical. The brief answer, in the risk analysis setup, is yes: for any given risk, their disjoint union covers the full state space, that is the set of "all that can happen." In particular, for any given risk, scenarios are mutually exclusive (see footnote 4 in Kaplan and Garrick, 1981, p. 13).

With respect to exhaustiveness, among others, Kaplan and Garrick point out (p. 14) that the list of scenarios might, in principle, be infinite. Hence, it is unlikely that a scenario analysis or a risk analysis be exhaustive in terms of the identified scenarios.

The pragmatic solution proposed by Kaplan and Garrick (1981) is to introduce a scenario called *other* that encompasses all scenarios that cannot be envisioned, making the list virtually exhaustive. At the same time, the problem of attaching a consequence to the residual scenario remains. Adding an *other* consequence, although coherent, is clearly problematic in that its evaluation is hardly determinable. A possibility is considering set valued consequences. The analysis we performed can be easily adapted to this situation, by adopting the extension of Anscombe and Aumann (1963) due to Jeleva and Jaffray (2011) and Viero (2009). This approach allows in general to consider misspecified consequences.

Another approach considers the possibility of not being able to specify a residual scenario: this is the case of the so called unknown unknowns. A caveat: when unknown unknowns enter the picture, it can be the case that the analyst may not know he is ignoring a scenario or, for a given and known scenario, he may not know he is ignoring consequences attached to it. Also in this case there are decision theoretic frameworks we can rely on, for example the one of Karni and Viero (2014), by suitably extending the translation we performed.

Models

Winkler (1994) and Bier (1993), among others, discuss some of the conceptual issues that may arise whenever models considered are not mutually exclusive or collectively exhaustive. In particular this poses a challenge to assigning a probability distribution over the considered models to express the (absolute) degree of confidence that the analyst assigns to their validity.

This issue can be addressed from different perspectives. One entails taking a conditional confidence approach, so that the probabilistic weight assigned to a model expresses the analyst's confidence on the model relative to the family of the considered ones, which in turn depends on the state of knowledge, that in our analysis is taken as given for any decision problem. Another adapts the line of reasoning we followed above for the case of completing scenarios. This case can be framed as considering a mixed model weighted according to the residual confidence. Another possibility involves the use of capacities in place of probabilities to quantify the various degree of confidence in the conceived models.

States of the world vs. states of the environment

The stochasticity of final outcomes, conditional on states realization, suggests that the space S of *states of the environment* of the Anscombe and Aumann framework does not reduce all the uncertainty in the system. This is not the case in another fundamental decision theoretic setup: the one of Savage (1954), in which states are complete descriptions of (all the outcome relevant aspects of) the world, in fact called *states of the world*. Conditional upon them the final outcome is deterministic. As discussed by Marinacci (2015), this determinism comes at the cost of a larger state space with respect to the one of Anscombe and Aumann. Relative to the latter framework a Savagean state of the world determines not only the state of the environment, but also describes how selected actions and realized states of the environment jointly determine final outcomes. Denote by D the set of available actions. A Savage state is a pair $(s, r) \in S \times R$ where R is the set of all *result functions* $r : D \times S \to Z$. Any function r determines the final outcome $z' = r(d', s')$ of every action $d' \in D$ in every state of the environment $s' \in S$. Therefore, for every action d, $\zeta^d(s, r) = r(d, s)$ is the final outcome resulting from the choice of d and conditional on the realization of state (s, r). The function

$$\begin{array}{cccc} \zeta^d : & S \times R & \to & Z \\ & (s, r) & \mapsto & \zeta^d(s, r) \end{array}$$

is called *Savage act*. The analysis of Marinacci (2015) can be adapted to show that probabilistic model uncertainty in the Savage framework we just sketched, accounts simultaneously for parametric model uncertainty and structural model uncertainty in the Anscombe and Aumann framework we adopted here. At the same time, when the grand state space $S \times R$ is considered, the elements of R can be called hidden states because they have a different observability standing relative to the elements of S. While the component s of (s, r) is tacitly assumed to be an observable that will realize in the implicit time horizon (say, a peak ground acceleration in the seismic case),[9] r is an hypothetical causality relation between (action, state of the environment) pairs and outcomes,[10] typically not observable and that (by definition) cannot realize because only one action d can be taken.

[9] A fact in the words of de Finetti (1971).

[10] A theory in the words of de Finetti (1971).

References

[1] Anscombe, F. J., Aumann, R. J., 1963. A definition of subjective probability. The annals of mathematical statistics, 34, 199-205.

[2] Apostolakis, G., 1990. The concept of probability in safety assessments of technological systems. Science, 250, 1359–1364.

[3] Apostolakis, G., 1994. A commentary on model uncertainty (No. NUREG/CP–0138; CONF-9310377–). Nuclear Regulatory Commission, Washington, DC (United States). Div. of Safety Issue Resolution; Maryland Univ., College Park, MD (United States); EG and G Idaho, Inc., Idaho Falls, ID (United States).

[4] Apostolakis, G. E.,1999. The distinction betweeen aleatory and epistemic uncertainties is important; An example from the inclusion of aging effects in the PSA. PSA99, Washington DC.

[5] Nilsen, T., Aven, T., 2003. Models and model uncertainty in the context of risk analysis. Reliability Engineering & System Safety, 79, 309–317.

[6] Bier, V. M., 1994. Some illustrative examples of model uncertainty (No. NUREG/CP–0138; CONF-9310377–). Nuclear Regulatory Commission, Washington, DC (United States). Div. of Safety Issue Resolution; Maryland Univ., College Park, MD (United States); EG and G Idaho, Inc., Idaho Falls, ID (United States).

[7] Borgonovo, E., Cappelli, V., Maccheroni, F., Marinacci, M., 2015. Risk Analysis and Decision Theory: Foundations. IGIER working paper 556, Universitá Bocconi, Milano.

[8] Buslik, A., 1994. A Bayesian approach to model uncertainty (No. NUREG/CP–0138; CONF-9310377–). Nuclear Regulatory Commission, Washington, DC (United States). Div. of Safety Issue Resolution; Maryland Univ., College Park, MD (United States); EG and G Idaho, Inc., Idaho Falls, ID (United States).

[9] Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., Montrucchio, L., 2013. Ambiguity and robust statistics. Journal of Economic Theory, 148, 974–1049.

[10] de Finetti, B. 1971. Probabilitá di una teoria e probabilitá dei fatti. In Studi in onore di G. Pompilj. Oderisi Editore, Gubbio.

[11] Gilboa, I., Marinacci, M., 2013. Ambiguity and the Bayesian Paradigm, in: Acemoglu, D., Arellano, M., Dekel, E. (Eds.), Advances in Economics and Econometrics: Theory and Applications, Cambridge University Press.

[12] Jaffray, J. Y., Jeleva, M., 2011. How to deal with partially analyzable acts?. Theory and decision, 71, 129-149.

[13] Kaplan, S., Garrick, B. J., 1981. On the quantitative definition of risk. Risk analysis, 1, 11-27.

[14] Kaplan, S., 1997. The Words of Risk Analysis. Risk Analysis, 17, 407–417.

[15] Kaplan, S., Perla, H. F., Bley, D. C., 1983. A methodology for seismic risk analysis of nuclear power plants. Risk Analysis, 3, 169-180.

[16] Karni, E., Vierø, M. L., 2014. Awareness of unawareness: a theory of decision making in the face of ignorance, mimeo.

[17] Klibanoff, P., Marinacci, M., Mukerji, S., 2005. A smooth model of decision making under ambiguity. Econometrica, 73, 1849–1892.

[18] Paté-Cornell, M., Fischbeck, P., 1992. Aversion to epistemic uncertainties in rational decision making: effects on engineering risk management, in: Risk-Based Decision Making in Water Resources V. New York: ASCE, pp. 200–218.

[19] Savage, L., 1954. The foundations of Statistics. Wiley and Sons.

[20] Seo, K., 2009. Ambiguity and Second-Order Belief. Econometrica, 77, 1575-1605.

[21] Vierø, M. L., 2009. Exactly what happens after the Anscombe–Aumann race?. Economic Theory, 41, 175-212.

[22] Winkler, R. L., 1994. Model uncertainty: probabilities for models? (No. NUREG/CP–0138; CONF-9310377–). Nuclear Regulatory Commission, Washington, DC (United States). Div. of Safety Issue Resolution; Maryland Univ., College Park, MD (United States); EG and G Idaho, Inc., Idaho Falls, ID (United States).

[23] Winkler, R., 1996. Uncertainty in probabilistic risk assessment. Reliability Engineering & System Safety, 54, 127–132.

[24] Zio, E., Apostolakis, G. E., 1996. Two methods for the structured assessment of model uncertainty by experts in performance assessments of radioactive waste repositories. Reliability Engineering & System Safety, 54, 225-241.