

## Response to the Reviewers of “The Coming Breakthrough in Risk Research”

Carlo Jaeger

First of all let me thank the reviewers for their encouraging comments – the time and effort they put into them are particularly appreciated in view of the many difficulties one is faced with when thinking about risk and rationality in view of today’s global problems.

### **Reviewer 1**

*“1) When dealing with uncertainty, (as far as I understood) the only source of uncertainty for a rational agent is the stochasticity of the phenomena observed. Yet there is also another source that is that of incomplete information. Could the author comment on that?”*

The revised abstract now states that the source of uncertainty for a rational agent may lie both in the randomness of the relevant phenomena and in the incompleteness of information about given states of affairs. This point is then elaborated in section 2, where the mathematical notion of stochasticity is distinguished from the notion of objective probabilities.

*“2) The example on page 9, can also be explained by making use of conditioned probability, in any case it is worth noticing that the basic mechanism of extraction with replacement is another way to produce multiplicative noise, that in various forms (De Solla Price, AB models) is at the basis of some of the log-normal/power-laws distributions observed.”*

The point is well-taken, I pick it up at the end of section 4.

*3) Does the form of the random shock ( $\xi_i$ ) play any role in the derivation of solution for iterated games?*

I had not seen the importance of this question before reading the comment by the reviewer. I address it in the discussion of equation (3).

*4) in the application to climate change a third source of uncertainty is given by the error in measurement of the quantity of interest. How to deal to compute correctly the risk with this third source of uncertainty?*

In the revised version, this point is addressed in section 7 in the paragraph on “The emphasis on the subjective aspects of both probability and utility”.

### **Reviewer 2**

Special thanks for the exceptionally detailed and careful reading of the manuscript!

*“First, I think the paper discusses in a novel way the relation between decision theory and uncertainty in the context of policy design, but I am not sure this objective comes clear to the reader from the beginning. The authors should better clarify in the introduction and*

*conclusions what are the contributions of the paper: should the reader expect a review or a novel conceptual framework? "*

The revised version addresses this issue right at the beginning and takes it up again in the conclusion.

*Page 9 "The two cases yield not only the same expected utility, " : Shouldn't one add "under the assumption that the second urn is unbiased"?*

*"(as usual in these problems, one considers an ideal setting without any biases)" added "Consider the case of picking a ball twice with replacement" : please explain in one more sentence how the replacement works.*

The setting is spelled out in detail.

*Page 11 "If a society or social network is unable to maintain such equilibria for sufficient amounts of time, it is bound to disintegrate, as patterns of communication and interaction will break down" : please rephrase, I suppose this means to maintain at least one those equilibria. Also maybe the subsequent sentence "Metastable equilibria offer a possible representation of social conventions, rules, norms and the like." seems to be the assumption under which the previous sentence holds true.*

Both sentences have been rephrased accordingly.

*Page 12 "And fourth, it is crucial to distinguish between marginal measures that work under the assumption that the system one is part of remains in the same basin of attraction and inframarginal measures where transition from one such basin to another one are essential." : please explain, at least in a footnote what is meant by marginal and inframarginal measures. Does this mean here statistics over a subset of variables in the system or over the time periods when the system visit a metastable equilibrium?*

Rephrased accordingly.

*Page 12 "All four implications can be spelled out with the specific format of transition functions considered in (3). This format, however, presupposes that for each agent the decision problem (2) has a unique solution, that action spaces and utility functions don't change and that the overall network changes only in a random way." : This statement is not obvious to me. If it has been proven in previous work, it would help to cite those works and say in one sentence the intuition why it holds.*

The paragraph has been rewritten so as to be more precise and to offer warrants for the claims made.

*Page 13 "Moreover, it leads to an additional, critical insight for risk management and governance. This is the acknowledgement that in order to become practically useful, (4) needs to be specified with the help of major additional assumptions about the particular problem at hand." : what are the characteristics of Eq. 4 that leads to such an insight?*

Explained in an additional paragraph.

*Page 15. "They proved a famous theorem about quantum mechanics, here we are interested in the generalization of classical logic that they introduced, because it is relevant for tackling uncertainty and ambiguity in decision-making." : I suggest to add a note or more precise reference to the specific result the author want to refer to here, so that the interested reader can look at it.*

More precise references have been added, together with a sentence about the point of the Kochen-Specker theorem.

*Page 20 "What they cannot do is to show the one best strategy for tackling climate change.": I think one- two more sentences could help the reader in this logical step that is very crucial for the paper. If I understand right the argument: first, we are not able to assign ex ante probability distributions to the outcomes of two given different future actions; second once we have chosen one of the two action, we are excluding from observation the other one and we will not be able to know more, even ex post, about the probability distributions for the action we did not take. So we will never know for sure which one was better. I am trying to understand what is the fundamental hypothesis that makes this statement true. Is it simply that we are not able to assign ex ante probability distributions to the outcomes of actions? But this is the case for basically all policy context? DO I understand right?*

This comment is closely related to the next one, so I discuss them together below.

*Page 21 "The first practical implication then is that German climate policy cannot and need not be justified by a cost-benefit calculus based on a comparison between short term losses in GDP from climate policy and long-term damages to GDP from climate change. It cannot be so justified because the only way to find out actual GDP in, say, 2100 without climate policy, would be by not implementing climate policy and seeing what would happen. And the only way to find out GDP in 2100 under conditions of stringent climate policy would be to realize such a policy. These two possibilities correspond to incompatible propositions, they cannot be known together."*

*I would suggest to (re)explain why estimating the GDP value in 2100 under the two scenarios cannot be done today in a meaningful sense. I am trying to understand what is precise argument made by the authors here.*

*Consider the following example. In principle, if we know enough of the dynamics of a system and the uncertainty associated to the factors that we do not control is small enough, then we can reasonably predict the trajectories resulting from two initial conditions that are sufficiently different. This holds both if we simulate a system with a computer and if we devise a physical experiment. Now, suppose we would be allowed to run only one trajectory in a given scenario. We could still reasonably predict what would have happened in another scenario and thus decide the scenario that is best according to some criterion. So what is the fundamental difference, between the above experimental set up and the GDP of Germany under different climate policies?*

*For instance, in the example above the prediction is not possible anymore if the system exhibits deterministic chaos, i.e. if the dynamics, even in the absence of stochasticity, is such that differences in initial conditions are amplified exponentially during the evolution. It would also not be possible if there is an unknown feedback between the position of the system in the trajectory and the future trend in the trajectory, or if the expectations of the agents on the future of the trajectory make unknown contributions to the dynamics of the trajectory today. In the examples I made, it is a matter of either unknown dynamics or chaotic dynamics (in the technical meaning) or both, that makes impossible to predict the outcome of the trajectory we do not choose to run.*

*I would suggest the authors to better explain if the reason why "These two possibilities correspond to incompatible propositions, they cannot be known together." Comes from the uncertainty I mentioned or from something else.*

*My concern here is the following. The analogy to quantum logic and the notion of complementary variables that is suggested by the authors is very interesting. However, one*

*could argue that quantum logic is a way to extend classic logic in a way to encompass the facts of quantum mechanics, and in particular the constraints of complementarity among certain variables such as position and momentum of a particle. However, we cannot really say that we fully understand why we do not know what we don't know. Quantum mechanics is very counterintuitive and challenges our cognitive abilities, maybe because our brain has evolved dealing with phenomenon at the scale of meters, while the physics at the scale of atoms is different. If we tell policy makers that the future impact of climate policies is like quantum mechanics and we cannot even know why we do not know, I am not sure how this will make the debate evolve.*

This comment and the previous one were extremely helpful to me. Understanding complementary variables is hard in itself, finding out how to use them in risk research may be even harder. The two comments led me to rewrite both the end of section 6 and parts of section 7. Much more work is warranted in this area: it offers opportunities for other papers, and for other authors, too.

Typos highlighted by reviewer 2

Corrected as suggested, with the exception of “complementarity logic” – this is as it should be (“complementary logic” is not the intended meaning).