*Behavioural Public Policy* (2025), **9**, 614–629
doi:10.1017/bpp.2024.59

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

# Knightian uncertainty in the regulatory context

Cass R. Sunstein

Robert Walmsley University Professor, Harvard University, Cambridge, MA, USA
Email: csunstei@law.harvard.edu

This essay contains a number of declarative sentences, but the concluding paragraphs might be considered a (mildly desperate) request for help, especially from decision theorists. Some instructive efforts have of course been made (for example, Cagliarini and Heath, 2000). I have been working on these issues for many years, and with permission, I draw here on my prior work, including an essay in the *Yale Journal on Regulation* a few years ago (Sunstein, 2020a). I am grateful to the editors of this issue and an anonymous reviewer for exceedingly valuable comments.

(Received 16 October 2024; accepted 22 October 2024)

**Abstract**

In 1921, John Maynard Keynes and Frank Knight independently insisted on the importance of making a distinction between uncertainty and risk. Keynes referred to matters about which 'there is no scientific basis on which to form any calculable probability whatever'. Knight claimed that 'Uncertainty must be taken in a sense radically distinct from the familiar notion of Risk, from which it has never been properly separated'. Knightian uncertainty exists when people cannot assign probabilities to imaginable outcomes. People might know that a course of action might produce bad outcomes A, B, C, D and E, without knowing much or anything about the probability of each. Contrary to a standard view in economics, Knightian uncertainty is real, and it poses challenging and unresolved issues for decision theory and regulatory practice. It bears on many problems, potentially including those raised by artificial intelligence. It is tempting to seek to eliminate the worst-case scenario, and thus to adopt the maximin rule, which might seem to be the appropriate approach under Knightian uncertainty. But serious problems arise if eliminating the worst-case scenario would (1) impose high risks and costs, (2) eliminate large benefits or potential 'miracles' or (3) create uncertain risks.

**Keywords:** cost–benefit analysis; Knightian uncertainty; precautions; risk; worst-case scenarios

## 'We simply do not know'

In some contexts, risk-related problems involve hazards of ascertainable probability. It might be possible to say that the risk of death from a certain activity is 1/100,000. Or it might be possible to say that the risk of catastrophic harm from some activity is 2%. Alternatively, we might be dealing with a problem where the probability of harm cannot be specified but has a known range – say, from 1/20,000 to 1/40,000, with an exposed

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

population of 10 million. Or we might be able to say that the risk of catastrophic harm is under 10% but above 1%.

At the same time, it is possible to imagine situations in which analysts, agents or observers cannot assign probabilities to potential outcomes – a topic that has received considerable attention (Knight, 1921; Caballero and Krishnamurthy, 2008; Nishimura and Ozaki, 2017; Kay and King, 2020). They might think or know that outcomes A, B, C, D and E are possible, but they might not know the probability of each. They might not be able to identify a specific probability or even a known range of probabilities. They might be frequentists who do not know and cannot find relevant frequencies. They might be Bayesians who lack necessary information.

In 1921, Frank Knight wrote:

> Uncertainty must be taken in a sense radically distinct from the familiar notion of Risk, from which it has never been properly separated.. The essential fact is that 'risk' means in some cases a quantity susceptible of measurement, while at other times it is something distinctly not of this character; and there are far-reaching and crucial differences in the bearings of the phenomena depending on which of the two is really present and operating. (Knight, 1921, p. 19–20)

Knight's own account, influenced by William James, was quite radical (Rizzo and Dold, 2021). Knight believed that uncertainty is real, including in markets, and that the problem could not be overcome or brushed aside by modeling subjective probabilities (id.). Jon Elster offers a vivid example: 'One could certainly elicit from a political scientist the subjective probability that he attaches to the prediction that Norway in the year 3000 will be a democracy rather than a dictatorship, but would anyone even contemplate *acting* on the basis of this numerical magnitude?' (Elster, 1983).

Regulators, entrepreneurs (Rizzo and Dold, 2021) and ordinary people are sometimes acting in situations of Knightian uncertainty, which I will understand here as those in which outcomes can be identified but no probabilities can be assigned. So understood, Knightian uncertainty is easy to distinguish from risk, where it is possible to identify outcome and to assign probabilities to them (Elster, 1983; Bewley, 1988; Davidson, 1991). We can also imagine cases like those above, involving what I shall call bounded uncertainty, or Knightian uncertainty within some kind of band: we know that the probability of certain outcomes is above (say) 10% but below (say) 20%, but we do not know anything else. If we keep in mind the idea of bounded uncertainty, we need to distinguish it from unbounded uncertainty; let us call the latter 'pure' Knightian uncertainty. There is a continuum here: If uncertainty is narrowly bounded (say, the probability is at most 2% and at least 1.99%), we are close to a situation of risk; if uncertainty exists within large bounds (say, the probability is at most 99% and at least 1%), we are close to a situation of pure Knightian uncertainty.

Knight emphasized the difficulty or impossibility of assigning probabilities to outcomes, but he also signaled the problem of *ignorance*, in which we are unable to specify either the probability of bad outcomes or their nature – where we do not even know the kinds or magnitudes of the harms that we are facing (Arrow, 1984; Smithson, 1989; Harremoes, 2003; Aven and Steen, 2010; Giang, 2015; Kay and King, 2020; Rizzo and

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

Dold, 2021). Though Knightian uncertainty is sometimes understood to include ignorance, so understood, my emphasis here is on cases in which outcomes can be defined but probabilities cannot. Some people appear to think that artificial intelligence creates an uncertain risk of catastrophe, including extinction of the human race (Center for AI Safety, 2023). Other people think the same thing about climate change. If so, what is the right response?

Consider in this regard a passage from John Maynard Keynes, also writing in 1921:

> By "uncertain" knowledge, let me explain, I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty; nor is the prospect of a Victory bond being drawn. Or, again, the expectation of life is only slightly uncertain. Even the weather is only moderately uncertain. The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence, or the obsolescence of a new invention, or the position of private wealth owners in the social system in 1970. About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know. (Keynes, 1921)

Like Knight, Keynes urged that some of the time, we cannot assign probabilities to imaginable outcomes.[1] Keynes immediately added, with some bemusement, that

> the necessity for action and for decision compels us as practical men to do our best to overlook this awkward fact and to behave exactly as we should if we had behind us a good Benthamite calculation of a series of prospective advantages and disadvantages, each multiplied by its appropriate probability, waiting to be summed. (Keynes, 1921)

How on earth, he wondered, do we manage to do that? Keynes listed three techniques (and it is worth considering their role in consideration of various hazards in the century since he wrote):

[1]It is important to note that Keynes and Knight had different concerns (Dimand, 2021; Packard *et al.*, 2021; Gerrard, 2022). Their differences are not relevant for my purposes here. Of particular note is this difference between Keynes and Knight:

> Keynes and Knight both grasped the essential difference between probability-as-risk and probability-as-uncertainty, but they travelled along vastly different roads to get there. Knight contextualised risk and uncertainty in the economic theory of profit as the reward for successful entrepreneurial action under uncertainty. The consequence of Knight's emphasis on context is that the philosophical foundations of his approach are less developed. Keynes's road was much longer, more circuitous and initially primarily concerned with the philosophical foundations, culminating in *A Treatise on Probability* before more fully contextualising his logical theory of probability in the behaviour of the economic system as a whole. The different roads followed by Keynes and Knight have had one crucial consequence. Keynes's greater emphasis on the philosophical issues led him ultimately to treat uncertainty as relating to the weight of argument (i.e., the evidential base), not probability *per se*, whereas Knight defined uncertainty in terms of probability (i.e., the degree of belief), not the evidential base that determined the degree of belief. (Gerrard, 2022).

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

(1) We assume that the present is a much more serviceable guide to the future than a candid examination of past experience would show it to have been hitherto. In other words, we largely ignore the prospect of future changes about the actual character of which we know nothing.
(2) We assume that the existing state of opinion as expressed in prices and the character of existing output is based on a correct summing up of future prospects, so that we can accept it as such unless and until something new and relevant comes into the picture.
(3) Knowing that our own individual judgment is worthless, we endeavor to fall back on the judgment of the rest of the world which is perhaps better informed. That is, we endeavor to conform with the behavior of the majority or the average. The psychology of a society of individuals each of whom is endeavoring to copy the others leads to what we may strictly term a conventional judgment (Keynes, 1937).

Keynes did not mean to celebrate those techniques. Actually he thought that they were ridiculous. We might know, for example, that technological innovations have not produced horrific harm in the past, and so we might think that artificial intelligence will not produce such harm in the future (strategy (1)). As good Hayekians, we might look at the price signal, such as flood insurance premiums and the value of coastline real estate, to assess the risks associated with climate change (strategy (2)). We might follow the wisdom of crowds to assess the likelihood of a pandemic (strategy (3)). But under circumstances of uncertainty, should we trust any of these? 'All these pretty, polite techniques, made for a well-panelled Board Room and a nicely regulated market, are liable to collapse', Keynes urged, because and when 'we know very little about the future' (Keynes, 1937). (For a creative institutional response, suggesting the importance of cognitive diversity, see Dold and Rizzo, 2021.)

One reason for uncertainty or ignorance might be that we are dealing with a novel, unique or nonrepeatable event. Another reason might be that we are dealing with a problem involving interacting components of a system, and we cannot know how those components will interact with each other, which means that ex ante predictions are highly unreliable (Taleb *et al.*, 2014).[2] Yet another reason is that the factors that will produce one result, or another result, are so numerous, and so hard to identify in advance, that any assignment of probability has no grounding (Kahneman *et al.*, 2021).

## Any port in a storm

Keynes pointed to three strategies that he thought intuitive but 'liable to collapse'. Are there better ones? It is important to say that over time, some problems that involve ignorance might be transformed into problems of uncertainty, and some problems of uncertainty might turn into problems of risk – a point that may counsel in favor of delay while new information is received. (In important ways, that is the arc of

[2]Of particular note is Taleb *et al.* at 11, emphasizing, "A lack of observations of explicit harm does not show absence of hidden risks. Current models of complex systems only contain the subset of reality that is accessible to the scientist. Nature is much richer than any model of it. To expose an entire system to something whose potential harm is not understood because extant models do not predict a negative outcome is not justifiable; the relevant variables may not have been adequately identified."

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

human history.) With respect to regulation, consider OMB Circular A-4, the Economic Constitution of the United States, as it was in effect until 2023: 'For example, when the uncertainty is due to a lack of data, you might consider deferring the decision, as an explicit regulatory alternative, pending further study to obtain sufficient data'. But as the circular also noted, 'Delaying a decision will also have costs, as will further efforts at data gathering and analysis' (OMB, 2003).

Delay of regulation may produce serious harm (including large numbers of deaths; consider the coronavirus pandemic). In principle, agencies would calculate the costs and benefits of delay. But because of the very problem that counsels in favor of delay (lack of information), that calculation is unlikely to be possible.

If uncertainty is genuinely bounded, agencies might use *breakeven analysis* (Sunstein, 2014). Suppose, for example, that the costs of regulation (involving, say, cybersecurity) are $100 million, that the benefits range from $150 million to $5 billion, and that technical analysts state that at the present time, they cannot assign probabilities to the lower or upper bound, or to points along the range. Even so, it is clear that the regulation should go forward. Or suppose that the monetized costs of some new technology (say, fracking, or certain uses of artificial intelligence) are $500 million, but that the monetized benefits range from $600 million to $500 billion (and we cannot assign probabilities to the various possible outcomes). A regulatory ban would not be a good idea. We could easily imagine variations on these numbers. Breakeven analysis can enable regulators to identify reasonable paths forward even in the midst of uncertainty, so long as it is bounded.

In some cases, however, it will be inadequate. Suppose that we have pure case of Knightian uncertainty: The benefits of some technology are $900 million, and no probabilities can be assigned to a wide spectrum of exceptionally serious harms. Or suppose that we know that the monetized costs of some new technology range from zero to $900 billion (or much more, e.g., extinction of the human race), and that the monetized benefits range from zero to $900 billion (or much more, e.g., $2 trillion) – and that with respect to both benefits and costs, it is impossible to assign probabilities to the various outcomes. If we are really operating under Knightian uncertainty, breakeven analysis cannot solve our problem.

The Principle of Insufficient Reason says that when people lack information about probabilities, they should act as if each probability is equally likely (Dubs, 1942; Luce and Raiffa, 1957; Sinn, 1980; Rawls, 1999). There is some evidence that people follow that principle, at least in surveys (Sunstein, 2006). But why is it rational to do that? By hypothesis, there is no reason to believe that each probability is equally likely. Making that assumption is no better than making some other, very different assumption. The Principle of Insufficient Reason is essentially arbitrary (Kay and King, 2020).

## Maximin

When strategies of avoidance are unappealing or unsuccessful, regulators might be drawn to the maximin rule: *Eliminate the worst-case scenario* (Arrow and Hurwicz, 1972; Elster, 1983). In the context of regulation of new technologies such as artificial intelligence, for example, strong precautions might be justified by reference to the maximin rule, asking officials to identify the worst case among the various options, and to

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

select that option whose worst-case is least bad. Perhaps the maximin rule would lead to what we might call a Catastrophic Harm Precautionary Principle, by, for example, urging elaborate steps to combat potentially horrific risks. It follows that if aggressive measures are justified to reduce the risks associated with artificial intelligence, one reason is that those risks are potentially catastrophic and existing knowledge does not enable us to assign probabilities to the worst-case scenarios. The same analysis might be applied to many problems, including the risks associated with genetically modified food (Taleb *et al.*, 2014), nuclear energy (Elster, 1983), pandemics and terrorism.

To understand these claims, we need to back up a bit. Maximin has sometimes been recommended under circumstances of uncertainty rather than risk (Elster, 1983). In an influential discussion, John Rawls, focusing on justice, offers a justification for a rule that 'directs our attention to the worst that can happen' (Rawls, 1999).[3] As it puts it, 'this unusual rule' is plausible in light of 'three chief features of situations' (Rawls, 1999). The first is that we cannot assign probabilities to outcomes, or at least we are extremely uncertain of them. The second is that the chooser 'has a conception of the good such that he cares very little, if anything, for what he might gain above the minimum stipend that he can, in fact, be sure of by following the maximin rule'. For that reason, it 'is not worthwhile for him to take a chance for the sake of further advantage'. The third is that 'the rejected alternatives have outcomes that one can hardly accept'. In other words, they involve 'grave risks'. Under the stated conditions, the gains are quite limited from running a catastrophic risk, which means that choosers do not much value them, and it is worthwhile giving them up to protect against a downside outcome that choosers deplore. This is an effort to produce a framework, pragmatic in character, for handling Knightian uncertainty.

Rawls emphasizes that the three 'features work most effectively in combination', which means that the 'paradigm situation for following the maximin rule is when all three features are realized to the highest degree' (Rawls, 1999). That means that the rule does not 'generally apply, nor of course is it self-evident' (Rawls, 1971).[4] It is 'a maxim, a rule of thumb, that comes in its own in special circumstances', and 'its application depends upon the qualitative structure of the possible gains and losses in its relation to one's conception of the good, all this against a background in which it is reasonable to discount conjectural estimates of likelihoods' (Rawls, 1971).

Rawls' own argument is that for purposes of justice, the original position, as he understands it, is 'defined so that it is a situation in which the maximin rule applies' (Rawls, 1971) – which helps to justify his principles of justice. The original position, so understood, is one of Knightian uncertainty. We can think of Rawlsian cases as involving something akin to a 'negative freeroll': a situation in which one can incur losses but obtain no (real) gains. Who wants that? In such cases, applying maximin seems quite rational.

[3]Rawls draws on but significantly adapts the work of William Fellner (Fellner, 1965).

[4]I am cheating a little bit here, referring to the original rather than the revised version of Rawls' book. (Sometimes the original is best.) It should be noted that in later work in particular, Rawls emphasized that the Kantian foundations of the veil of ignorance, and those ideas could also be connected with the difference principle. I am bracketing that possibility for my purposes here.

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

## Knightian uncertainty and precautions

These points bear on regulatory policy, where Rawls' defense of maximin has inspired a defense and reconstruction of the Precautionary Principle in an important essay by Stephen Gardiner (Gardiner, 2006). To make the underlying intuition clear, Gardiner begins with the problem of choosing between two options, A and B:

> If you choose A, then there are two possible outcomes: either (A1) you will receive $100, or (A2) you will be shot. If you choose B, there are also two possible outcomes: either (B1) you will receive $50, or (B2) you will receive a slap on the wrist. According to a maximin strategy, one should choose B. This is because: (A2) (getting shot) is the worst outcome on option A and (B2) (getting a slap on the wrist) is the worst option on plan B; and (A2) is worse than (B2). (Gardiner, 2006, p. 46)

It should be immediately apparent that if we can assign probabilities to outcomes, A might turn out to be the better choice. Suppose that if you choose A, there is a 99.99999% chance of (A1), and that if you choose B, there is a 99.99999% chance of (B2). If so, A might seem better. But let us stipulate that assignment of probabilities is not possible. In Gardiner's view, this conclusion helps support what he calls the Rawlsian Core Precautionary Principle in the regulatory setting: When Rawls' three conditions are met, precautions, understood as efforts to avoid the worst-case scenario, should be adopted. (We could consider efforts to create resilience to be precautions, as in the case of climate change.) As he puts it:

> If one really were faced with the genuine possibility of disaster, cared little for the potential gains to be made by avoiding disaster and had no reliable information about how likely the disaster was to occur, then, other things being equal, choosing to run the risk might well seem like a foolhardy and thereby extreme option. (Gardiner, 2006, p. 49)

Gardiner adds, importantly, that to justify use the maximin rule, the threat posed by the worst-case scenario must satisfy some minimal threshold of plausibility. In his view, 'the range of outcomes considered are in some appropriate sense "realistic," so that, for example, only credible threats are considered'. (Gardiner, 2006, page 51). If they can be dismissed as unrealistic, then the maximin rule should not be followed.[5] Gardiner believes that the problem of climate change, and also that of genetically modified organisms, can be usefully analyzed in these terms and that they present good cases for application of the maximin rule:

> The RCPP [Rawlsian Core Precautionary Principle] appears to work well with those global environmental issues often said to constitute paradigm cases for the precautionary principle, such as climate change and genetically-modified crops. For reasonable cases can be made that the Rawlsian conditions are satisfied

[5]There are some conceptual puzzles here. If an outcome can be dismissed as unrealistic, then we are able to assign some probabilities, at least. Gardiner's argument must be that in some cases, we might know that the likelihood that a bad outcome would occur really is trivial.

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

> in these instances. For example, standard thinking about climate change provides strong reasons for thinking that it satisfies the Rawlsian criteria. First, the "absence of reliable probabilities" condition is satisfied because the inherent complexity of the climate system produces uncertainty about the size, distribution and timing of the costs of climate change. Second, the "unacceptable outcomes" condition is met because it is reasonable to believe that the costs of climate change are likely to be high, and may possibly be catastrophic. Third, the "care little for gains" condition is met because the costs of stabilizing emissions, though large in an absolute sense, are said to be manageable within the global economic system, especially in relation to the potential costs of climate change. (Gardiner, 2006, p. 55)

In a similar vein, Jon Elster, speaking of nuclear power, contends that maximin is the appropriate choice when it is possible to identify the worst-case scenario and when the alternatives have the same best consequences (Elster, 1983). Elster urges that 'to the extent that we are in a state of uncertainty, it is rational to act as if the worst that can happen is bound to happen' (again assuming the same best-consequences) (see also Taleb *et al.*, 2014). Revealingly, the examples of genetically-modified crops and nuclear power look badly out of date, which can be taken as a cautionary note about applications of such arguments to specific problems. But what matters is the general ideas, not the particular applications.

## Objections

I now turn to four objections to the argument on behalf of using the maximin rule under circumstances of Knightian uncertainty, in ascending order of force.

### *The argument is trivial*

An evident problem with the argument for the maximin rule under the stated assumptions is that it risks triviality (Kelsey, 1993).[6] If individuals and societies can eliminate an uncertain danger of catastrophe for essentially no cost, then of course they should eliminate that risk! If people are asked to pay $1 to avoid a potentially catastrophic risk to which probabilities cannot be assigned, they might as well pay $1. And if two options have the same best-case scenario, and if the first has a far better worst-case scenario, people should of course choose the first option, at least if we know nothing about probabilities. Consider: Option A has a best-case scenario of wonderful lives for all, and a worst-case scenario of wonderful lives for almost all and nearly wonderful lives for two; Option B has a best-case scenario of wonderful lives for all, and a worst-case scenario of horrible lives for all. Option A seems better.

[6]Kelsey says the following: "It is often argued that lexicographic decision rules such as maximin are irrational, since in economics we would not expect an individual to be prepared to make a small improvement in one of his objectives at the expense of large sacrifices in all of his other objectives. This criticism is less powerful in the current context since we have assumed that the decision maker has a weak order rather than a cardinal utility function on the space of outcomes. Given this assumption the terms 'large' and 'small' used in the above argument are not meaningful" (Kelsey, 1993).

In many contexts, however, decision makers do have a cardinal utility function, not merely a weak order.

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

(I say 'seems' because we are, by hypothesis, dealing with Knightian uncertainty, so on decision-theoretic grounds, it is not quite so simple.)

The real world rarely presents problems of these forms. Where policy and law are disputed, the elimination of uncertain dangers of catastrophe imposes both costs and risks. In the context of climate change, for example, it is implausible to say that regulatory choosers can or should care 'very little, if anything', for what might be lost by following the maximin rule. If nations followed that rule for climate change, they would spend a great deal indeed, right now, to reduce greenhouse gas emissions. The result would almost certainly be far higher prices for energy, probably producing significant increases in suffering, unemployment and poverty. Note too that precautionary measures might have decreasing marginal returns or increasing marginal costs, which is not exactly ideal, and which raises a serious cautionary note against 'doing all one can' to eliminate worst-case scenarios.

If we eliminate the worst-case scenarios associated with artificial intelligence, we will also lose extraordinary gains in terms of money, health, safety and more. If we eliminate the worst-case scenarios for all pandemic risks, people might be required to stay at home, today, tomorrow and the day after. While that might be the right approach, the fact that a very bad worst-case scenario is associated with the pandemic (worse, let us stipulate, than the worst-case associated with the mandate) cannot easily be taken to justify a must-stay-home mandate, without trying to know more about probabilities. It has long been known that something similar can be said about genetic modification of food, because elimination of the worst-case scenario, through aggressive regulation, might well eliminate an inexpensive source of nutrition that would have exceptionally valuable effects on countless people who lives under circumstances of extreme deprivation (Anderson and Nielsen, 2004).

The real question, then, is whether regulators should embrace the maximin rule in real-world cases in which doing so is costly or extremely costly. If they should, it is because condition (3) (the 'care little for gains' condition) is too stringent and should be abandoned. If the costs of following the maximin rule are significant, and if regulators care a great deal about incurring those costs, the question is whether it makes sense to follow the maximin rule when they face uncertain dangers of catastrophe. In the environmental context, some people have so claimed (Elster, 1983; Woodward and Bishop, 1997).[7] To say the least, this claim is not obviously right, and it takes us directly to the next objection to the maximin rule.

### *Infinite risk aversion*

Rawls' arguments in favor of adopting maximin, for purposes of distributive justice, were subject to vigorous objections from economists – objections that many economists accept to this day (Arrow, 1973; Harsanyi, 1975). The central challenge was that the maximin principle would be chosen only if choosers showed *infinite risk*

[7] Elster's specific arguments with respect to nuclear power have not (in my view) stood the test of time, but the analytics hold up exceedingly well. (Some of the datedness of the dated parts is relevant to the discussion here, as in, 'The doubts concerning sun, wind and water are too great for these to be more than interesting side options' (Elster, 1983). With what implicit probability judgment was that sentence written?)

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

*aversion*. In the words of one of Rawls' most influential critics, infinite risk aversion 'is unlikely. Even though the stakes are great, people may well wish to trade a reduction in the assured floor against the provision of larger gains. But if risk aversion is less than infinite, the outcome will not be maximin' (Musgrave, 1974, p. 627).

To be more specific: Suppose that you have a choice between two options. Option A carries with it a 99.9999% likelihood of great wealth and welfare and a 0.0001% likelihood of a terrible outcome. Option B carries with it a 60% chance of a very bad outcome and a 40% chance of a just-short-of-terrible outcome. Would it really make sense to choose Option B? To adapt this objection to the environmental context: It might be plausible to assume a bounded degree of risk aversion with respect to catastrophic harms, in order to support some modest forms of a Catastrophic Harm Precautionary Principle. But even under circumstances of uncertainty – the argument goes – maximin is senseless unless societies are to show infinite risk aversion.

This is a standard challenge, but it is wrong, because maximin does not assume infinite risk aversion. By stipulation, *we are dealing with Knightian uncertainty*. (Chu and Liu, 2001). Perhaps that is rare in the regulatory context. Moreover, it is true that when we are dealing with bounded uncertainty, a version of the infinite risk aversion objection could work: Following an unqualified maximin rule could require regulators to eliminate harms even if all probabilities, in the relevant range or 'band', are small (say, we know that the probability of a truly terrible outcome is between 0.05% and 0.02%, but we know nothing else).

Still, the objection that maximin assumes infinite risk aversion depends on a denial that (pure) Knightian uncertainty exists; it assumes that subjective choices will be made and that they will reveal subjective probabilities. It is true that subjective choices will be made. But such choices do not establish that objective uncertainty does not exist. To see why, it is necessary to engage that question directly.

### *Uncertainty does not exist*

Many economists have denied the existence of uncertainty (for discussion, see Kay and King, 2020; Rizzo and Dold, 2021). Milton Friedman, for example, writes of the risk-uncertainty distinction that 'I have not referred to this distinction because I do not believe it is valid. I follow L.J. Savage in his view of *personal probability*, which denies any valid distinction along these lines. We may treat people as if they assigned numerical probabilities to every conceivable event' (Friedman, 1976).[8] Friedman, Savage and

[8]Consider also Hirshleifer and Riley: "In this book we disregard Knight's distinction, which has proved to be a sterile one. For our purposes risk and uncertainty mean the same thing. It does not matter, we contend, whether an 'objective' classification is or is not possible. For, we will be dealing throughout with a 'subjective' probability concept (as developed especially by Savage, 1954): probability is simply *degree of belief* …. Because we never know true objective probabilities, d]ecision-makers are …never in Knight's world of risk but instead always in his world of uncertainty. That the alternative approach, assigning probabilities on the basis of subjective degree of belief, is a workable and fruitful procedure will be shown constructively throughout this book." (Hirshleifer and Riley, 1992).

For the purposes of the analysis by Hirshleifer and Riley, the assignment of subjective probabilities may well be the best approach. But the distinction between risk and uncertainty is not sterile when regulators are considering what to do but lack information about the probabilities associated with various outcomes.

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

other skeptics[9] are correct to insist that people's choices suggest that they assign probabilities to events. On a widespread view, an understanding of people's choices can be taken as evidence of subjective probabilities. People's decisions about whether to fly or instead to drive, whether to go to a store during a pandemic, whether to walk in certain neighborhoods at night, and whether to take risky jobs can be understood as an implicit assignment of probabilities to events. Indeed, regulators themselves make decisions, including decisions about artificial intelligence and climate change, from which subjective probabilities can be calculated.

But none of this makes for anything like a good objection to Knight, who was concerned with objective probabilities rather than subjective choices (LeRoy and Singell Jr, 1987; Kay and King, 2020; Rizzo and Dold, 2021). Animals, no less than human beings, make choices from which subjective probabilities can be inferred. I have two Labrador Retrievers, Snow and Finley, and both Snow and Finley make decisions about risks (from motor vehicles, strangers and mysterious noises) that reflect judgments about subjective probabilities. But the existence of subjective probabilities does not mean that Snow and Finley do not ever face (objective) uncertainty. Human beings face Knightian uncertainty too (Dibiasi and Iselin, 2021; Rizzo and Dold, 2021). 'From the fact that it is always possible to elicit … subjective probabilities, we should not conclude that one ought rationally to act upon them' (Elster, 1983, p. 199). (For provocative comments on the role of intuitive judgments, see Rizzo and Dold, 2021; I confess that I am skeptical about the usefulness of those judgments, for reasons akin to Keynes' doubts about the strategies conventionally used to manage uncertainty.)

Suppose that the question is the likelihood that at least one hundred million human beings will be alive in 10,000 years For most people, equipped with the knowledge that they have, no probability can sensibly be assigned to that outcome. Perhaps uncertainty is not unbounded; the likelihood can reasonably be described as above 0% and below 100%. (I think.) But beyond that point, there is little to say. Or suppose that I present you with an urn, containing 250 balls, and ask you to pick one; if you pick a blue ball, you receive $1000, but if you pick a green ball, you have to pay me $1000. Suppose that I refuse to disclose the proportion of blue and green balls in the urn – or suppose that the proportion has been determined by a computer, which has been programmed by someone that neither you nor I know. You can make a pick, but what does that tell us about actual probabilities? Suppose that a music company has a new song, by a new artist, that seems terrific and catchy; suppose too that the popularity of songs is a complex matter, depending on social influences that cannot be foreseen in advance. How reliable are probability assessments? Or suppose the question is the probability that artificial intelligence will produce horrific harm within the next 50 years. Is the right number 10%? 20%? 40%? 60%? Frequentists will not be able to answer that question. On what basis might Bayesians do so? That is meant as a rhetorical question.

Regulators may be in a position of Knightian uncertainty, or some form of it, at the early stage of a pandemic or when dealing with a new technology. These examples suggest that it is wrong to deny the possible existence of uncertainty, signaled by the absence of objective probabilities (Elster, 1983).

[9]Frank Ramsey must, alas, be included, I fear (Westgren and Holmes, 2022).

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

## Knightian uncertainty is pretty rare

Notwithstanding what I have said here, it is an understatement to insist that regulatory problems do not typically involve uncertainty, certainly (!) in its pure form, or even within wide bands. Using frequentist strategies, regulators are often able to assign probabilities to outcomes, and Bayesian approaches can also be used (Sunstein, 2020b; OMB, 2024). When they cannot, perhaps regulators can instead assign probabilities to probabilities (or where this proves impossible, probabilities to probabilities of probabilities). There are many techniques to attempt to do that. In many cases, regulators might be able to specify a range of probabilities – saying, for example, that the probability of catastrophic outcomes from a pandemic or climate change is above 2% but below 30%.[10]

Whatever we think of any particular example, we might be able to agree that pure Knightian uncertainty is pretty rare, at least over significant time horizons. Perhaps we can agree that at worst, regulatory problems typically involve problems of (manageably) bounded uncertainty, in which we cannot assign probabilities within specified and not-so-wide bands. It might be possible to think, for example, that the risk of a catastrophic outcome is above 1% but below 5%, without being able to assign probabilities within that band. The pervasiveness and nature of Knightian uncertainty depend of course on what is actually known. If pure uncertainty is pretty rare, then Gardiner's argument, or variations on it, do not apply outside of exotic cases. Fair enough. But even if this is so, exotic cases do arise; in fact we are living with some of them. That is important.

## On sleeping well at night

A great deal of work explores the question whether people should follow the maximin rule under circumstances of Knightian uncertainty (Arrow and Hurwicz, 1972). Some of this work draws on people's practices or intuitions, in a way that illuminates actual beliefs but may tell us little about what rationality requires (Harsanyi, 1975). Other work is highly formal, adopting certain axioms and seeing whether the maximin rule violates them (Luce and Raiffa, 1957). The results of this work are not conclusive (Luce and Raiffa, 1957; Arrow and Hurwicz, 1972). Maximin has not been ruled out as a candidate for rational choice under uncertainty.

In deciding whether to follow the maximin rule under circumstances of Knightian uncertainty, or something close to it (such as bounded uncertainty), a great deal should turn on two questions (Elster, 1983): (a) How bad is the worst-case scenario, compared to other bad outcomes? (b) What, exactly, is lost by choosing the maximin rule? Of course, it is possible that choosers, including regulators, will lack the information that would enable them to answer these questions. But (and this is the central point) in the regulatory context, answers to both (a) and (b) may well be possible *even if it is not possible to assign probabilities to the various outcomes with any confidence*. By emphasizing

[10]I am bracketing here frequentist claims about the pervasiveness of uncertainty (Kay and King, 2020). Even if we are frequentists, regulators are often dealing with repeated cases for which frequentist assignments of probability are perfectly feasible; consider food safety, occupational safety, and air pollution.

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

the relative badness of the worst-case scenario, and the extent of the loss from attending to it, I am attempting to build on the Rawls/Gardiner suggestion that maximin is the preferred decision rule when little is lost from following it.

To see the relevance of the two questions, suppose that you are choosing between two options. The first has a best-case outcome of 10 and a worst-case outcome of −5. The second has a best-case outcome of 15 and a worst-case outcome of −6. It is impossible to assign probabilities to the various outcomes. Maximin would favor the first option, to avoid the worse worst-case (which is −6); but to justify that choice, we have to know something about the *meaning* of the differences between 10 and 15 on the one hand and −5 and −6 on the other. If 15 is much better than 10, and if the difference between −5 and −6 is a matter of relative indifference, then the choice of the first option is hardly mandated. But if the difference between −5 and −6 greatly matters – if it is a matter of life and death – then the maximin rule is much more attractive.

Consider a regulatory analogue. Suppose that as compared with a ban, allowing some new technology would have a best-case outcome of $2 billion in annual net benefits and a worst-case outcome of −$10 million in net benefits. Suppose that we cannot assign probabilities to the various outcomes. Under the maximin rule, we should ban the new technology. But if the net loss of $10 million is not a big deal, we might reject the maximin rule by reference to something like the Rawls/Gardiner theory. Importantly, we might in cases of this kind favor the *maximax* rule: If Option A and Option B have roughly equivalent downsides, and if Option B has an immeasurably better upside, we should choose Option B. Of course we could vary the numbers in such a way as to make the maximin rule much more attractive.

These points have the important implication of suggesting the possibility of a (rough) cost–benefit analysis of whether to follow the maximin rule under conditions of both risk and uncertainty. Sometimes the worst-case is the worst by far, and sometimes we lose relatively little by choosing the maximin rule. It is typically thought necessary to assign probabilities in order to engage in cost–benefit balancing; without an understanding of probabilities, such balancing might not seem able to get off the ground. But a crude version of cost–benefit balancing is possible even without reliable information about probabilities. For the balancing exercise to work, of course, it must be possible to produce cardinal rankings among the outcomes – that is, it must be possible to rank them not merely in terms of their badness but also in at least rough terms of how much worse each is than the less-bad others. That approach will not work if cardinal rankings are not feasible – as might be the case if (for example) it is not easy to compare the catastrophic loss from a pandemic with the loss from huge expenditures on efforts to control a pandemic. Much of the time, however, cardinal rankings are possible in the regulatory context.

Here is a simpler way to put the point. It is often assumed that in order to undertake cost–benefit analysis, it is necessary to assign probabilities, with the understanding that point estimates represent the average or most probable case. But in some cases, a sensible rule-of-thumb can be adopted without assigning probabilities. An understanding of the magnitude of the relevant payoffs can help regulators to navigate difficult situations. If one option has a large downside but no substantial upside, it can be rejected in favor of one that lacks that downside but that has a roughly equivalent upside. And recall cases for maximax: where one option has a terrific upside and a bad downside,

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

it should be favored over another option that has a merely decent upside and a bad downside.

To appreciate the need for some kind of analysis of the effects of following the maximin rule, imagine an individual or society lacking the information that would permit the assignment of probabilities to a series of hazards with catastrophic outcomes; suppose that the number of hazards is 10, 20 or 1000. Suppose too that such an individual or society is able to assign probabilities (ranging from 1% to 90%) to an equivalent number of *other* hazards, with outcomes that range from bad to extremely bad, but never catastrophic. Suppose, finally, that every one of these hazards can be eliminated at a cost – a cost that is high, but that does not, once incurred in individual cases, inflict harms that count as extremely bad or catastrophic. The maximin rule suggests that our individual or society should spend a great deal to eliminate each of the 10, 20 or 100 potentially catastrophic hazards. But once that amount is spent on even one of those hazards, there might be nothing left to combat the extremely bad hazards, even those with a 90% chance of occurring. We could easily imagine that a poorly informed individual or society would be condemned to real poverty and distress, or even worse, merely by virtue of following maximin. In these circumstances, the maximin rule does not have a lot of appeal.

This suggestion derives indirect support from the empirical finding that when asked to decide on the distribution of goods and services, most people reject the two most widely discussed principles in the philosophical literature: average utility, favored by Harsanyi, and Rawls' difference principle (allowing inequalities only if they work to the advantage to the least well-off) (Frohlich and Oppenheimer, 1992). Instead, people choose average utility with a floor constraint – that is, they favor an approach that maximizes overall well-being, but subject to the constraint that no member of society may fall below a decent minimum (Frohlich and Oppenheimer, 1992). Insisting on an absolute welfare minimum to all, they maximize over that floor. Their aversion to especially bad outcomes leads them to a pragmatic threshold in the form of the floor. So too, very plausibly, in the context of precautions against risks. A sensible individual, or society, would not always choose maximin under circumstances of either risk or uncertainty. A great deal depends on what is lost, and what is gained, by eliminating the worst-case scenario (including by increasing resilience); and much of the time, available information makes it possible to answer those questions at least in general terms (on relevant institutional considerations, relevant to government as well as to entrepreneurship, see Dold and Rizzo, 2021).

Nothing here is meant as some kind of proof that maximin is forbidden, or even not required, by rationality (Luce and Raiffa, 1957). My claim is instead that for prudent regulators, attempting to proceed in the midst of (pure) Knightian uncertainty, the maximin rule makes most sense when the worst-case scenario, under one course of action, is much worse than the worst-case scenario under the alternative course of action, when there are no huge disparities in gains from either option, and when the choice of maximin does not result in extremely significant losses. Variations on this case will present harder challenges, but in some situations, they too will allow room for the maximin rule. At the same time, it is important for prudent regulators to focus as well on the best-case scenarios, which may promise miracles (Rowell, 2020); that

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

possibility may provide an important cautionary note about efforts to eliminate risks, including those posed by new technologies.

In the hardest and most intriguing cases, it is not possible to defend any simple rule. Some kind of judgment must be made. Nothing in decision theory can specify that judgment. But in the face of unknown probabilities of genuine catastrophe for which 'wait and learn' is imprudent, it is reasonable to take strong protective measures, whether the problem involves a pandemic, climate change, artificial intelligence, or the kinds of dangers that each of us faces in ordinary life. Those measures will enable us to sleep better at night. And if we end up taking too many precautions, well, so be it.

**Competing interests.** The author declares none.

## References

Anderson, K. and C. P. Nielsen (2004), 'Golden rice and the looming GMO debate: Implications for the poor', Centre for Economic Policy Research, Discussion Paper No. 4195. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=508463.

Arrow, K. (1984), *Individual Choice under Certainty and Uncertainty*, Cambridge: Harvard University Press.

Arrow, K. and L. Hurwicz (1972), 'An Optimality Criterion for Decision-Making under Ignorance', in C. F. Carter and J. L. Ford (eds), *Uncertainty and Expectations in Economics: Essays in Honor of G.L.S. Shackle*, Oxford: Blackwell, 1–12.

Arrow, K. J. (1973), 'Some ordinalist-utilitarian notes on Rawls's Theory of Justice', *Journal of Philosophy*, **70**(9): 245–263.

Aven, T. and R. Steen (2010), 'The concept of ignorance in a risk assessment and risk management context', *Reliability Engineering and System Safety*, **95**(11): 1117–1122.

Bewley, T. F. (1988), 'Knightian Uncertainty', in D. P. Jacobs, E. Kalai, M. I. Kamien and N. L. Scwartz (eds), *Frontiers of Research in Economic Theory*, Cambridge: Cambridge University Press, 71–81.

Caballero, R. J. and A. Krishnamurthy (2008), 'Collective risk management in a flight to quality episode', *The Journal of Finance*, **63**(5): 2195–2230.

Cagliarini, A. and A. Heath (2000), 'Monetary policy-making in the presence of Knightian uncertainty', Reserve Bank of Australia. https://www.rba.gov.au/publications/rdp/2000/2000-10/knightian-uncertainty-and-expected-utility-theory.html [20 February 2024].

Center for AI Safety (2023), 'Statement on AI Risk', Center for AI Safety. https://www.safe.ai/statement-on-ai-risk [20 February 2024].

Chu, C. Y. and W.-F. Liu (2001), 'A dynamic characterization of Rawls's maximin principle: theory and implications', *Constitutional Political Economy*, **12**(3): 268.

Davidson, P. (1991), 'Is probability theory relevant for uncertainty? A post Keynesian perspective', *Journal of Economic Perspectives*, **5**(1): 129–143.

Dibiasi, A. and D. Iselin (2021), 'Measuring Knightian uncertainty', *Empirical Economics*, **61**(4): 2113–2141.

Dimand, R. W. (2021), 'Keynes, Knight, and fundamental uncertainty: a double centenary 1921-2021', *Review of Political Economy*, **33**(4): 570–584.

Dold, M. and M. Rizzo (2021), 'Frank Knight and the cognitive diversity of entrepreneurship', *Journal of Institutional Economics*, **17**(6): 925–942.

Dubs, H. H. (1942), 'The principle of insufficient reason', *Philosophy of Science*, **9**(2): 123–131.

Elster, J. (1983), *Explaining Technical Change: A Case Study in the Philosophy of Science*, Cambridge: Cambridge University Press.

Fellner, W. (1965), *Probability and Profit*, Homewood, Illinois: R. D. Irwin.

Friedman, M. (1976), *Price Theory*, Chicago: Aldine Publishing Company.

Frohlich, N. and J. A. Oppenheimer (1992), *Choosing Justice: An Experimental Approach to Ethical Theory*, Berkeley: University of California Press.

Gardiner, S. M. (2006), 'A core precautionary principle', *Journal of Political Philosophy*, **14**(1): 33–60.

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press

Gerrard, B. (2022), 'The road less travelled: keynes and Knight on probability and uncertainty', *Review of Political Economy*, **36**: 1253–1278, advance online publication.
Giang, P. H. (2015), 'Decision making under uncertainty comprising complete ignorance and probability', *International Journal of Approximate Reasoning*, **62**: 27–45.
Harremoes, P. (2003), 'Ethical aspects of scientific incertitude in environmental analysis and decision making', *Journal of Cleaner Production*, **11**(7): 705–712.
Harsanyi, J. C. (1975), 'Can the maximin principle serve as a basis for morality? A critique of John Rawls' theory', *American Political Science Review*, **69**(2): 594–606.
Hirshleifer, J. and J. G. Riley (1992), *The Analytics of Uncertainty and Information*, Cambridge: Cambridge University Press.
Kahneman, D., O. Sibony, and C. R. Sunstein (2021), *Noise*, New York: Little, Brown, Spark.
Kay, J. and M. King (2020), *Radical Uncertainty*, New York: W. W. Norton.
Kelsey, D. (1993), 'Choice under partial uncertainty', *International Economic Review*, **34**(2): 305.
Keynes, J. M. (1921), *A Treatise on Probability*, London: MacMillan and Company.
Keynes, J. M. (1937), 'The general theory of employment', *Quarterly Journal of Economics*, **51**(2): 209–223.
Knight, F. H. (1921), *Risk, Uncertainty and Profit*, Cambridge: Houghton Mifflin Company.
LeRoy, S. F., L. D. and Singell, Jr (1987), 'Knight on risk and uncertainty', *Journal of Political Economy*, **95**(2): 394–406.
Luce, R. D. and H. Raiffa (1957), *Games and Decisions: Introduction and Critical Survey*, New York: Dover Publications.
Musgrave, R. A. (1974), 'Maximin, uncertainty, and the leisure trade-off', *Quarterly Journal of Economics*, **88**(4): 627–632.
Nishimura, K. G. and H. Ozaki (2017), *Economics of Pessimism and Optimism: Theory of Knightian Uncertainty and Its Applications*, Tokyo: Springer Japan.
OMB (2003), 'Circular A-4', https://www.whitehouse.gov/wp-content/uploads/2023/11/CircularA-4.pdf [20 February 2024].
OMB (2024), 'Reports'. https://www.whitehouse.gov/omb/information-regulatory-affairs/reports/ [20 February 2024].
Packard, M. D., P. L. Bylund and B. B. Clark (2021), 'Keynes and Knight on uncertainty: peas in a pod or chalk and cheese?', *Cambridge Journal of Economics*, **45**(5): 1099–1125.
Rawls, J. (1971), *A Theory of Justice*, Cambridge: Harvard University Press.
Rawls, J. (1999), *A Theory of Justice*, Revised Edition, Cambridge: Harvard University Press.
Rizzo, M. and M. Dold (2021), 'Knightian uncertainty: though a Jamesian window', *Cambridge Journal of Economics*, **45**: 967–988.
Rowell, A. (2020), 'Regulating best-case scenarios', *Environmental Law*, **50**(4): 1105–1172.
Sinn, H.-W. (1980), 'A rehabilitation of the principle of insufficient reason', *Quarterly Journal of Economics*, **94**(3): 493–506.
Smithson, M. (1989), *Ignorance and Uncertainty*, New York: Springer-Verlag.
Sunstein, C. R. (2006), 'Irreversible and catastrophic', *Cornell Law Review*, **91**(4): 841.
Sunstein, C. R. (2014), 'The limits of quantification', *California Law Review*, **102**(6): 1369–1421.
Sunstein, C. R. (2020a), 'Maximin', *Yale Journal on Regulation*, **37**(3): 940–979.
Sunstein, C. R. (2020b), 'On neglecting regulatory benefits', *Administrative Law Review*, **72**(3): 445–459.
Taleb, N. N., R. Read, R. Douady, J. Norman and Y. Bar-Yam (2014), 'The precautionary principle (with application to the genetic modification of organisms)', NYU School of Engineering Working Paper Series. http://www.fooledbyrandomness.com/pp2.pdf.
Westgren, R. E. and T. L. Holmes (2022), 'Entrepreneurial beliefs and agency under Knightian uncertainty', *Philosophy of Management*, **21**(2): 199–217.
Woodward, R. T. and R. C. Bishop (1997), 'How to decide when experts disagree: uncertainty-based choice rules in environmental policy', *Land Economics*, **73**(4): 505.

**Cite this article:** Sunstein CR (2025), 'Knightian uncertainty in the regulatory context', *Behavioural Public Policy* **9**, 614–629. https://doi.org/10.1017/bpp.2024.59

https://doi.org/10.1017/bpp.2024.59 Published online by Cambridge University Press