A Bayesian Framework for the Precautionary Principle*

Nabil I. Al-Najjar†

First draft: April 2013
This version: April 2015

Abstract

A common misconception is that precautionary motives in public policy cannot be justified within the Bayesian rational-choice framework, and that decision criteria that appeal to ambiguity and pessimism are needed. This paper critically evaluates these claims, arguing that a rational policy process must be Bayesian in order to avoid paradoxical, even absurd, recommendations, such as policies that depend on sunk cost or that suppress costless information. The paper also argues that the distinction between measurable risk and fundamental, or Knightian, uncertainty can be made within the standard framework of Bayesian rationality. Finally, a simple model is proposed to highlight situations where precautionary action may be normatively justified.

Forthcoming in: *Journal of Legal Studies.*

*I thank David Weisbach for his invitation to present this paper at the Conference on Developing Regulatory Policy in the Context of Deep Uncertainty at the University of Chicago Law School, and for his valuable comments as discussant. I also thank participants at the Workshop on Model Inadequacy in Policy Making at LSE, and at seminars in CERGE-EI, Prague, and the University of Vienna. I benefited from many discussions about decision theory with Jonathan Weinstein. Finally, I thank Luciano Pomatto for his comments and research assistance.

† Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston IL 60208.
Research Page: http://www.kellogg.northwestern.edu/faculty/alnajjar/htm/index.html.

Contents

1 Introduction

Debates in public policy often invoke precaution as a guiding principle when the nature and magnitude of risk is unknown. In legal and regulatory contexts, this view is expressed as a *Precautionary Principle*:

> *"When an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause-and-effect relationships are not fully established scientifically."*[1]

The principle has been endorsed in a broad set of conventions, laws, and treaties, including the United Nations Framework Convention on Climate Change, the 1992 Rio Declaration, the Treaty Establishing the European Community, the U.S. National Environmental Policy Act, the U.S. Clean Water Act, among others.[2]

Precautionary decision criteria are frequently criticized as irrational. Sunstein (2003, 2007), for example, argues that the Precautionary Principle is simply incoherent, "a crude and sometimes perverse way of promoting [public policy] goals, which can be obtained through other, better routes." This dissatisfaction has parallels in critiques of appeals to pessimism and ambiguity in macroeconomic, finance, and decision theoretic models.[3]

The special status given to a normative criterion like the Precautionary Principle is puzzling. A conservative attitude towards risky outcomes is neither theoretically novel nor uncommon in practice. The distaste for unpredictability can be expressed in terms of risk aversion within the standard expected utility framework. No new normative principles seem necessary to advise individuals on choosing the right level of life insurance coverage, or businesses on the right level of fire protection. What, then, justifies giving uncertainties, like those arising from inconclusive scientific evidence about climate change, conflicting intelligence reports about a terrorist threat, or incomplete understanding of the risks associated with a new medical treat-

[1] *"Wingspread Statement on the Precautionary Principle,"* Ashford *et al.* (1998).

[2] Other examples include the Energy Charter Treaty, the Cartagena Protocol on Biosafety, the International Joint Commission created by the U.S.-Canada Great Lakes Water Quality Agreement, the Occupational Safety and Health Act, and the Federal Food, Drug, and Cosmetic Act. See Sunstein (2003) and Barrieu and Sinclair-Desgagné (2010) for references and additional examples.

[3] See, for instance, Sims's (2001) critique of the maximin approach to model uncertainty in macroeconomics, and the critique by Al-Najjar and Weinstein (2009) of ambiguity aversion.

ment, a unique status that warrants special pessimistic or precautionary treatment?

The case for precaution often appeals to Knight's (1921) distinction between "measurable risks," where the odds are known, and "unmeasurable uncertainties," where no objective probability can be assigned. In a well-known passage, Keynes (1937) offered his own characterization of what became known as *Knightian uncertainty*:[4]

> *"The sense in which I am using the term [uncertainty] 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."*

Intuitively, fundamental or scientific uncertainties like these seem profoundly different from objectively quantifiable actuarial risks.[5] Criteria like the Precautionary Principle draw their appeal from the feeling that conventional theories of decision making under risk imply inadequate levels of precaution in such contexts, and that alternative frameworks, like the pessimistic criteria proposed by Schmeidler (1989), Gilboa and Schmeidler (1989), and Bewley (1986), may be unavoidable.

This paper argues that the Bayesian rational-choice framework, based on Savage (1954), can serve as a foundation for precaution in public policy. The argument has three parts:

1. Section 2 shows that deviating from Bayesian rationality leads to paradoxical, even absurd, recommendations, such as policies that depend on sunk cost or that suppress costless information. These fallacies are a by-product of mis-interpreting Savage's framework as having a substantive ontological content, well beyond its limited aim of providing a logically consistent calculus for uncertain propositions.

[4] Authors use the terms "unmeasurable," "fundamental," "scientific," "deep," and "Knightian" uncertainty to refer to the idea of risks that cannot be objectively quantified. I uses these terms interchangeably in this paper.

[5] In the same paragraph, Keynes gives examples of phenomena not subject to uncertainty: "By 'uncertain' knowledge [...] 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."

2. Section 3 argues that Knightian uncertainty should not be confounded with pessimism. I propose a distinction between Knightian uncertainty and measurable risk within the standard Bayesian framework. Risk is objective, testable, and not subject to disagreement; uncertainty is just the opposite.

3. Section 4 argues that using probabilities to express subjective uncertainty is consistent with Knightian uncertainty provided that commonly used, and often implicit, assumptions about the separability of payoffs are removed. I present a simple model illustrating this point.

The literature on precaution and the Precautionary Principle is too vast to attempt a survey here. See Sunstein (2003, 2007), and Barrieu and Sinclair-Desgagné (2010) for surveys and critical assessments. Currie and MacLeod (2014) provide an insightful exposition of a Bayesian theory of precaution in legal contexts. Among the omissions of the present paper is a discussion of the role of irreversibility and option values, following Arrow and Fisher (1974) and recent work by Gollier and Treich (2003) and others. I hope to cover the interaction of irreversibility and Knightian uncertainty in future work.

2 Paradoxes of Non-Bayesian Decision Criteria

While the abstract nature of Savage's framework makes it universally applicable, it also make it difficult for nonspecialists to translate into concrete settings or to evaluate the plethora of new exotic axiomatizations that purport to redefine rational choice.

This section aims to provide the nonspecialist with an accessible account of how policy choice problems can be mapped into Savage-style decision problems. I present a version of Savage's rationality principles that follows Ghirardato (2002). This version gives an identical representation to Savage's but has the advantage of explicitly modeling dynamic choice. The exposition aims to convey the main ideas with minimal formalism.[6]

[6] See the Appendix for the formal statements underlying this informal exposition.

2.1 Policy Choice as a Decision Problem

My description will follow Savage's original work. I refer the interested reader to Fishburn (1970), Kreps (1988), or Gilboa (2009) for textbook accounts.[7]

The ingredients of Savage's framework are *states of the world* and *consequences*. Savage's original exposition is hard to improve on: A state is "a description of the world, leaving no relevant aspect undescribed." On the other hand, "[c]onsequences might appropriately be called states of the person" and "might in general involve money, life, state of health, approval of friends, well-being of others, the will of God, or anything at all about which the person could possibly be concerned."

Consider a policy-making problem, such as that of approving a new medical treatment. A state in this case is a description that would include, for instance, the health impact of the treatment (e.g. the time needed to achieve results and how benefits vary by age group and medical history), side effects (e.g. their nature, intensity, and dependence on patient characteristics), and the interaction with other treatments. The description of the state needs to be as extensive as necessary so as not to leave any "relevant aspect undescribed."

Consequences should also be extensively defined. In a climate-change context, a consequence should describe not only monetary measures of well-being, such as GDP growth, but also non-monetary aspects such as public health consequences or the destruction of a natural habitat.

A policy is, formally, an *act* that indicates which consequence obtains at each state of the world. The adoption of a medical treatment, in this language, is an act f that yields consequence $f(s)$ when the state happens to be s. Approving an alternative treatment (or do nothing and maintain the status quo) corresponds to the selection of another act g.

A policy maker is assumed to have a *ranking* over consequences. Recall that consequences describe the entire societal impact of the treatment. Do the lives saved by the new treatment outweigh the potential hazard to the lives of others? Is the value of information about a potential terror plot sufficient to offset the risk to the lives of intelligence personnel? Answers to questions like these capture the policy maker's system of values, including

[7] For a concise account of the theory, Al-Najjar and De Castro (2011) may also be useful.

his appetite for risk taking, and is summarized by a utility function.

Policy makers are assumed to have a ranking, or preference, on policies, and will choose the policy they rank highest among feasible alternatives. If the true state s were known, then choice is easy: simply select the policy (act) that yields the most preferred consequence in that state. The problem of choice under uncertainty is how to choose between policies when the true state is unknown.

So far, this is just a neutral *language* to express *any* decision problems, from choosing statistical procedures to anti-terrorism policies, and anything in between. Savage's theory adds formal principles of rationality to help make consistent decisions by disentangling the roles of tastes, beliefs, and information.

2.2 Bayesian Rationality

Savage's goal was to propose normative principles, or postulates, "to distinguish between coherent behavior and blunder, or demonstrable incoherence, in the face of uncertainty," (Savage (1967)). This framework continues to provide the decision-theoretic foundation for most models in economics, game theory, and finance.

First, the concept of an *event* needs to be introduced. An event E is a set of states and should be thought of as the piece of information: "the true state belongs to E." With a judicious choice of the state space, any piece of information can be represented as an event. New intelligence about the whereabouts of a terror suspect, or scientific findings that a medical treatment causes undesirable side effects, are examples of events.[8]

The policy maker starts with an *(initial) ranking* that incorporates his system of values and the likelihoods assigned to various events given available information. In addition to this initial ranking, for every event E, the policy maker is assumed to have a *conditional ranking* given that event. This conditional ranking reflects the new information that event E actually occurred and captures the intuition that policy evaluation changes in response to new evidence. For example, while restricted electronic surveillance may initially be preferred to a more intrusive one, this preference is likely to be revised after a major terrorist attack.

[8] All events considered below are *non-null*. This roughly means that the decision maker views these events as having a chance of occurring.

To introduce the next two normative principles, we need to consider contingent policies and options. Suppose that g is the default, or status quo, policy and an alternative policy f is proposed. For concreteness, think of g is the current treatment of an illness and f is the proposed alternative. To make the problem interesting, assume that the new treatment is sufficiently more expensive to produce and administer that if one must choose between f and g based on the information available today, then the status quo treatment g will be chosen. Suppose that a soon to be released clinical study will reveal that either:

- f is a miracle treatment that yields superior results that justify the additional costs (Event E occurs); or
- f and g have identical health consequences (E does not occur).[9]

In addition to treatments f and g, the policy maker also has the option to make the treatment contingent on E: "implement treatment f if you learn that E occurred, otherwise implement treatment g." This more flexible policy option defines a new act, which we denote by fEg.

With the above notation, $\{g\}$ represents a situation where the only available policy is the status quo treatment g, while $\{fEg, g\}$ represents the option to choose between g or "implement f if E occurs, otherwise g." Indifference between $\{g\}$ and $\{fEg, g\}$ indicates a policy maker who does not believe that making policy choice contingent on E has value, presumably because this information is irrelevant.

The initial ranking and all conditional rankings are required to satisfy a number of noncontroversial rationality postulates and technical conditions. For instance, all rankings must be transitive (avoid circularity) and rank higher policies with uniformly better consequences (monotonicity).

The first substantive normative principle is:

> **Complete Conditional Rankings:** *The policy maker has a ranking over all policies conditional on any event E.*[10]

Having a conditional ranking for each event E may appear excessive. However, the fact remains that policy makers do not know in advance what

[9] The assumption that the clinical study provides a clear cut indication of the effectiveness of f is used here only to make the example easier to follow, but is not important.

[10] This assumption is implicit in the way Ghirardato (2002) sets up his framework. I single it out here in light of the discussion in Section 2.3.3.

new information will become available in the future that they will be asked to act upon. In the case of medical treatments, the set of all conceivable future laboratory discoveries, statistical findings, or clinical data is vast. The principle states that whatever the future information state may be, a policy maker cannot refuse to make a decision. Not making a decision is itself a decision, namely sticking with the status quo.

The next principle is obvious:

> **Information is Valuable:** *For any event E and acts f, g, if the policy maker prefers f to g when E actually occurs, then he must also prefer fEg to g.*

Violating this principle means that the policy maker views the information that E occurred as valuable (it makes him choose f instead of g), yet he is willing to pay money *not* to make his policy contingent on this information. This is related to the idea discussed below that a non-Bayesian may be willing to suppress costless information.

The final normative principle says that if the policy maker is willing to pay to make his policy contingent on a piece of information, then he must view this information as valuable. In our example, if the clinical study confirms that f is indeed a miracle treatment, then the policy maker must choose it.

> **Consistent Policy Implementation:** *For any event E and acts f, g, if the policy maker prefers fEg to g, then he must prefer f to g when E actually occurs.*[11]

Violating this principle means that the policy maker is willing to pay money to acquire the option to follow policy f if E occurs, but when this event actually occurs, he foregoes f in favor of the status quo policy g.

Using Savage's celebrated theorem, Ghirardato (2002) shows that the above normative principles (and the technical and noncontroversial postulates mentioned earlier) imply that the policy maker has a system of values, represented by a utility function, and beliefs, represented by a probability distribution, and ranks policies based on their expected utility. That is, the policy maker must be Bayesian.

[11] The principles of "value of information" and "consistency of implementation," collectively referred to as *dynamic consistency* in Ghirardato (2002).

2.3 The Non-Bayesian's Predicament

Non-Bayesian policies must necessarily violate one or more of the normative rationality principles outlined above. The non-Bayesian has to decide which principle(s) to compromise on, and to what extent. Considerable efforts went to finding ways out of this conundrum, giving rise to a large literature with a bewildering variety of fixes, compromises, and work-arounds. Here I highlight some of the disturbing, even absurd, policy implications that follow from abandoning the rationality principles.

2.3.1 Fact-Free Policies

The rationality principles introduced above imply that policy making should be based on facts only:

> **Fact-based Policy Making:** *For every event E and pair of policies f, g that yield identical consequences on E, the policy maker is indifferent between f and g given E.*

This is a formal statement of the principle that decision making should be *consequentialist*.[12] Violating this principle leads to fundamental paradoxes, the most disturbing of which is the endorsement of policies that depend on sunk cost. As an example, suppose that society has spent a considerable sum of money to develop a medical treatment f, only to learn that f is no better or worse than the status quo treatment g. The amount spent on developing f is sunk: no current or future decision can recover it fully or partially. Should a rational policy maker be willing to spend additional resources to implement f, or to forgo the cheaper alternative g? Rationality requires that once it is known that f and g have identical health consequences, the policy maker should ignore the sunk development cost of f; bygones-are-bygones.

Yet a non-Bayesian may take sunk cost into account. Al-Najjar and Weinstein (2009) provide simple examples where certain ambiguity-aversion decision maker may strictly prefer to make different choices in two decision problems that are identical in every respect except in the amount of sunk cost incurred at a prior stage.[13]

[12] see Machina (1989). Ghirardato (2002) states this as Axiom 7, but it is, in fact, implied by his other postulates; see the Appendix for details.

[13] Machina (1989) argues that making non-consequentialist choices may make sense for a decision makers who violate Bayesian rationality. From the perspective of this paper,

Although the sunk cost fallacy is common in practice, it seems absurd to recommend it as a basis for policy.[14] Rational policies ought to be based on facts and not on hangovers from past choices, however memorable or vivid they may be.

2.3.2 Information Aversion

A well-known fact in the decision-theoretic literature is that violating Savage's rationality principles may lead to aversion to information.[15] Willingness to pay money to suppress information is especially disturbing in situations of fundamental uncertainty. Assessing the impact of climate change, a new medical treatment, or an emerging terror threat are examples of situations where information ought to be, if anything, even more valuable compared to situations of measurable risk. A non-Bayesian will have to explain why shutting down information sources is good for policy. Should the publication of medical trial results and climate data be suppressed? Or should one recommend that intelligence gathered about new terror threats be kept away from political and military leaders?[16]

2.3.3 The No-Cop-Out Principle

Savage's framework assumes that the decision maker can rank all pairs of policies. This assumption was dropped in Bewley (1986).[17] Bewley's frame-

to recommend deviating from fact-based policy making to correct for violating Bayesian rationality amounts to using one flawed decision principle to offset the flaw in another.

[14] Sunk cost may play a strategic role against an opponent. The discussion here concerns nonstrategic decision problems.

[15] See Wakker (1988) paper, titled "Nonexpected Utility as Aversion of Information," for a detailed discussion. In their paper "Dynamically Consistent Beliefs Must be Bayesian," Epstein and Le Breton (1993) similarly observe that a consequence of violating the rationality principles is that "information will be rejected, even if it is costless." They then "highlight the counterintuitive nature of this implication" in a statistical setting in which a "decision maker would strictly prefer to have no information available to guide the later choice."

[16] In the climate change context, Lange (2003) finds that the maximin criterion can lead to a situation where learning has a negative value. He notes that "although a decision criterion that gives more weight to the worst case can be motivated in several ways, there are problems with applying it to a dynamic framework at least for normative reasons. A criterion [...] where information may have a negative value and thus even costless learning is disregarded, can have merits as a descriptive model but should not serve as a tool for policy advice."

[17] Bewley's paper was eventually published as Bewley (2002).

work is standard except for removing the requirement that decision makers have a complete ranking. Bewley interprets an incomplete ranking as follows: a decision maker who "volunteers" to rank policy f as superior to g is effectively indicating confidence that f is better than g, while abstaining from ranking f and g expresses uncertainty about the underlying probabilities.

Bewley's framework is not irrational. Rather, the problem is that it is silent about what to do in situations of fundamental uncertainty where lacking confidence in one's probability assessment is the norm. Savage's requirement of a complete ranking is a "no-cop-out principle." It captures the idea that real-world decision makers do not have the luxury of abstaining from making hard choices; that not making an active choice is itself a choice, namely the status quo. Bewley's theory is about decisions that policy makers are comfortable making, and thus does not apply to situations, like climate change, medical treatments, and terrorism, where all of the available options may be ambiguous.[18]

A related problem arises in connection with dropping the assumption of complete conditional rankings. This assumption requires that the no-cop-out principle also hold conditionally, denying the policy maker the convenience of saying "I do not know what to do if E happens."[19]

2.3.4 Incomplete State Spaces, Small Worlds, and Pessimism

The Bayesian framework imposes a substantial burden by requiring decision makers to specify a state space and a consequence space that leave out no relevant aspect of the decision problem. Savage recognized that working

[18] The U.S. raid on Osama Bin Laden's compound in May 2011 is a vivid illustration of the no-cop-out principle in a decision with high stakes and incalculable risks. Bob Woodward reports on the debate among the President's advisers: "Several assessments concluded there was a 60 to 80 percent chance that bin Laden was in the compound. Michael Leiter, the head of the National Counterterrorism Center, was much more conservative. During one White House meeting, he put the probability at about 40 percent. When a participant suggested that was a low chance of success, Leiter said, 'Yes, but what we've got is 38 percent better than we have ever had before.' " (*The Washington Post,* May 6, 2011). Despite these ambiguities, a decision had to be made. The president approved the raid at 8:20 a.m., Friday, April 29, 2011.

[19] One approach that gained some popularity recently is the so-called "recursive models" of (non-Bayesian) dynamic choice. Under this approach, one admits only events where updating does not result in paradoxes. Events that lead to inconsistencies are simply not considered. See Al-Najjar and Weinstein (2009) for discussion and references.

with this "grand state space" is practically and cognitively impossible. Any decision problem we will ever work with in practice is necessarily formulated as a "small-world" model that reduces complex real-world problems to manageable sizes. Savage was careful to point out, however, that a proper small-world model must also be one in which his normative postulates hold.[20]

In light of this, I interpret Savage's framework to mean:

1. The world is too complex to study directly.
2. This complexity must be reduced by using coarse "small-world" models.
3. A minimal consistency criterion to require of these models is that choice obeys the normative rationality principles.

Under this interpretation, Savage's principles of rationality are not ontological statements subject to confirmation or refutation by empirical data. The fact that most experimental subjects violate expected utility (*e.g.,* by displaying Ellsberg choices) is no more a refutation of Savage's theory than the fact that most people fail to apply *modus tollens* is a refutation of propositional calculus.

Quiggin (2005) proposes an incompleteness meta-hypothesis as a justification of the Precautionary Principle. The meta-hypothesis states: "Estimates of project outcomes derived from formal models of choice under uncertainty are inherently incomplete. Incomplete estimates will generally be over-optimistic. The errors will be greater, the less well-understood is the problem in question." Precautionary behavior is then justified as an antidote to the inexorable tendency toward over-optimism.

There are two parts to Quiggin's argument: (1) modelers and policy makers are forced to work with coarse models that overlook important parts of the problem; and (2) incomplete models will tend to be over-optimistic. The first part echoes Savage's small-worlds. The second part is more questionable because it refers to over-optimism as if there is an objective reference point to judge what constitutes unbiased probabilities. What was, on April 29, 2011, the unbiased estimate that Bin Laden was in the Abbottabad compound? Or that average global temperature will rise by 4 degrees by 2100?

[20] Kopylov (2007) provides a remarkable extension of Savage's theory to the case where the events, acts, and choices are restricted.

The absence of an objective reference point is one of the defining features of fundamental uncertainty, so arguments justifying caution based on pessimism are, at best, incomplete. And even if an objective reference point existed, one may just as reasonably expect that some decision makers will be over-pessimistic, leading to the reverse conclusion that a more aggressive attitude toward scientific uncertainty is warranted. Finally, this justification of caution does not tell us what to do if policy makers find themselves in violation of the rationality principles above.[21]

In summary, the Bayesian view recognizes that we have no choice but to use "small world" models that are coarse and incomplete. But instead of arbitrarily injecting pessimism into the analysis, Savage's rationality principles provide a tool for checking the consistency and completeness of our models.

2.4 Normative vs. Positive Interpretations of Rationality

Nowhere is human fallibility and limited cognitive abilities more vividly manifested than in the context of decision making under uncertainty. A large body of literature in psychology, too large to cite here, shows the many ways in which temptations, distractions, guilt, regret, disappointments, and elation can impact decisions.

The standing of Bayesian rationality principles as a foundation for public policy seems unassailable. That human actors "get it wrong" so often is an indication of just how hard it is to make consistent choices, and is a powerful vindication of the value of the rationality principles as a guide for action. In the words of the statistician D. Lindley, to let flawed decision practices shape the principles of rational choice would be like "asking people's opinion of 2+2, obtaining an average of 4.31 and announcing this to be the sum. It would be better to teach them arithmetic."[22]

For the policy maker, Savage's framework provides a calculus for un-

[21] Vermeule (2012) argues that our coarse understanding of the world leads not to arbitrary pessimistic decision criteria but to an irreducible diversity of opinions. Since agents may hold different beliefs based on the same common set of evidence, it is not sensible for courts to require parties to deliver further justification for their probability judgement in situations of fundamental uncertainty. The best that courts could do is to monitor these agents to verify whether their decisions (e.g. regarding information gathering) are consistent.

[22] Preface to de Finetti (1974).

certain propositions to help spot "demonstrable incoherence." The framework forces us to come clean about our system of values, assessment of the odds, and interpretation of information. An example is provided by Sunstein (2007) who contrasts American policy toward the risks of terrorism vs. climate change. According to Sunstein, American policy makers take highly precautionary attitudes toward terrorism risk, as exemplified by former Vice President Cheney's "one percent doctrine." On the other hand, some policy makers seem to require an exceptionally high standard of scientific certainty about the causal relationship between human activity and climate change. The Savage framework requires policy makers and societies to make explicit the system of values underlying these decisions: "which is a worse consequence, a 9/11-style attack or the destruction of an American city by a Katrina-like hurricane?" Answers to such questions help rationalize the policy making priorities and reduce the risk of confounding values and beliefs.

Much of the dissatisfaction with the Bayesian framework originates in a well-intentioned, but naive, expectation that a framework of decision making should tell us what decisions to make. "Where do beliefs come from?" is the all-too-common critique of Bayesian framework. What this critique misses is the limited aim of this framework, namely that of providing logical consistency checks for reasoning about uncertainty. The Bayesian framework does not purport to tell us what our judgement of the odds or what our values should be (which is a worse consequence, the destruction of the towers in New York or the flooding of New Orleans?).

An analogy with propositional calculus may be helpful. The rules of logic are powerful in auditing our reasoning to check for "demonstrable incoherence." Their power comes, in part, from their applicability to any deductive reasoning, in any field, time period, or context. On the other hand, the rules of logic do not help in deciding which premises make sense or which theorems are interesting. Similarly, it is not the goal of Savage's framework to help us make substantive judgement calls, but to "search for incoherence among potential decisions [...] The theory itself does not say which way back to coherence is to be chosen, and presumably should not be expected to" (Savage (1967)). Expecting an answer to the question "What beliefs should I hold?" is just as unreasonable a burden to place on the Bayesian framework as expecting propositional calculus to tell us what premises are true.

3 Knightian Uncertainty: A Bayesian Reformulation

"[T]he existence of a problem of knowledge depends on the future being different from the past, while the possibility of the solution of the problem depends on the future being like the past."

Frank Knight (1921)

The term "Knightian uncertainty" is often associated with pessimistic choices and behavioral anomalies, like those presented in Ellsberg (1961).[23] In this section, I argue that this association bears little connection with Knight's conception of risk and uncertainty. Here, I propose a distinction between the two concepts that is consistent with rationality and in the spirit of Knight (1921).

3.1 Knight on Knightian Uncertainty

The misconception that the Bayesian framework precludes Knightian uncertainty is understandable. Knight wrote his book a century ago, decades before the development of modern expected utility theory.[24] Lacking a formal framework to express his views, Knight's ideas can be difficult to follow for the modern reader. He was clear, however, that in one-shot situations, such as Ellsberg's thought experiments, there can be no distinction between risk and uncertainty: "[W]hen an individual instance only is at issue, there is no difference for conduct between a measurable risk and an unmeasurable uncertainty. The individual [...] throws his estimate of the value of an opinion into the probability form of 'a successes in b trials' [...] and 'feels' toward it as toward any other probability situation."

[23] There are too many examples to cite here. Bewley (1986) and Epstein and Wang (1994), for example, make the connection to Knight clear in the titles of their papers. Hansen and Sargent (2001) write that "Knight (1921) distinguished risky events, which could be described by a probability distribution, from a worse type of ignorance that he called uncertainty and that could not be described by a probability distribution." Another example is Routledge and Zin's (2009) paper on liquidity in the financial crisis where they argue that, in Savage's framework, " 'model uncertainty' is indistinguishable from the risk inherent in the asset's stochastic process. The Savage independence postulate implies that one can simply collapse the probability weighting across possible models ('uncertainty') with the probabilities for payoffs ('risk') to represent behavior with a single probability measure for states."

[24] Knight's *"Risk, Uncertainty, and Profit"* began as his Ph.D. thesis. It was eventually published in 1921.

In their study of Knight's work, LeRoy and Singell (1987) conclude that "Knight shared the modern view that agents can be assumed always to act as if they have subjective probabilities." Whatever motivates people to display Ellsberg choices (whether it be paranoia, fear that the experiment is rigged, or whatever) is not Knightian uncertainty—at least not as far as Frank Knight is concerned.

The necessity of a Bayesian point of view is also evident in Keynes (1937). More than a decade before modern subjective expected utility theory, Keynes wrote that even in situations of uncertainty, "the necessity for action and for decision compels us [...] to behave exactly as we should if we had [...] a series of prospective advantages and disadvantages, each multiplied by its appropriate probability, waiting to be summed."

All this suggests that the distinction between "measurable risk and an unmeasurable uncertainty" is not in conflict with the rationality postulates. Knight, in fact, equated the risk-uncertainty distinction with the distinction between *objective* and *subjective probability*.[25] Risk corresponds to situations in which the odds are known/measurable/objective; uncertainty is the opposite. He writes, for example, that "the bursting of bottles does not introduce an uncertainty or hazard into the business of producing champagne; since in the operation of any producer a practically constant and known proportion of the bottles burst." Although the outcome for any single bottle is unpredictable, the odds are objectively measurable through repetition. Here I propose an abstract approach to capture these ideas. Section 4 suggests a concrete application to policy making.[26]

3.2 Measurable Risk

What does it mean to say that a probability distribution P represents a "measurable risk?" Intuitively, saying that an entity is "measurable" means that there is an objective device to measure it. An analogy may be useful: we think of heat and electric current as measurable physical phenomena because we have devices (the thermometer and ammeter) to objectively measure their levels.

[25] "We can also employ the terms 'objective' and 'subjective' probability to designate the risk and uncertainty respectively." (Knight (1921))

[26] See also Al-Najjar and Weinstein (2015) who focus on the role of uncertainty in a model of precautionary saving.

It is essential that the measurement is objective, that it is not open to subjective interpretations and disagreement. The truth or falsehood of the statement: "the objective probability of an event is $p\%$" should be no more open to subjective opinions than the statement: "the temperature of the room where you are reading this paper is 78 degrees."[27]

A natural language to express this idea is that of statistical tests. We imagine a world consisting of "experiments," each representing a case or instance of a problem of interest. For example, a medical case corresponds to a patient's history, symptoms, health consequences of the treatment, and any other relevant aspect of the case.

There is no hope of measuring probability in a single experiment. If you roll a die once, you can measure whether an event like "the die turned 5" occurred. But from a single roll, it is not possible to measure the *probability* that the die turns 5. As Knight clearly understood, when "an individual instance only is at issue, there is no difference [..] between a measurable risk and an unmeasurable uncertainty." The objective measurement of probability requires placing individual experiments in the context of the repetition of similar experiments.

To make this formal, consider an idealized experiment where outcomes are represented by a finite set S. In the medical treatment example, S is the outcome of one patient to whom the treatment is applied. We imagine an environment where the experiment is repeated many times, infinitely often, in fact. Medical treatments are applied to many patients; anti-terrorism policies are implemented in a large number of instances of security threats. Infinite repetition simplifies the exposition, but the ideas can be translated to a context of large but finite repetitions. The state space in this case is the set $S = S^1 \times S^2 \times \cdots$ of all sequences $(s^1, s^2, \ldots)$, where s^1 denotes the outcome of the first experiment, s^2 the outcome of the second, and so on.

Let $\mathcal{P}$ be a set of probability distributions on S. A distribution $P \in \mathcal{P}$ may be viewed as a theory of how outcomes are produced. In the medical treatment example, the "true" effectiveness of a treatment is a distribution that describes, probabilistically and for each patient, how health consequences vary with individual characteristics. In the medical context, $\mathcal{P}$ is a

[27] In the words of Schmeidler (1989), "[t]he concept of objective probability is considered here as a physical concept like acceleration, momentum, or temperature; to construct a lottery with given objective probabilities [..] is a technical problem conceptually not different from building a thermometer."

set of theories of a treatment's effectiveness. Our objective is to distinguish between those elements of $\mathcal{P}$ that represent measurable risk and those that represent unmeasurable uncertainty.

Let the subset $\Theta \subset \mathcal{P}$ represent our candidate for a set of measurable risks. A *statistical test* for Θ is a collection of events $\mathcal{T} = \{T_\theta\}_{\theta \in \Theta}$ with the interpretation that T_θ consists of all sequences of outcomes $(s^1, s^2, \ldots)$ that *confirm* θ. That is, the test defines what observations are consistent with θ being the true distribution generating the data. The next definition proposes a formal notion of measurement for probability:

Definition 1 (Measurement Mechanism) *A* measurement mechanism *for a set of probability distributions Θ is a statistical test $\{T_\theta\}_{\theta \in \Theta}$ satisfying the properties:*

1. *If $\theta \in \Theta$ is the true distribution, then the observed sequence of outcomes $(s^1, s^2, \ldots)$ confirms θ with probability 1; that is,*

$$\theta(T_\theta) = 1.$$

2. *If a sequence of outcomes $(s^1, s^2, \ldots)$ confirms θ, then the same sequence cannot confirm some other $\theta' \neq \theta$; that is,*

$$T_\theta \cap T_{\theta'} \text{ is empty for all } \theta' \neq \theta.$$

We say that Θ is (objectively) measurable *if there is a measurement mechanism for Θ.*

The definition builds on familiar statistical concepts. Requirement 1 says that the test has zero Type I error: if θ is indeed the true distribution, then this fact can be measured by observing what actually happens. Requirement 2 is an identification condition: if an observation confirms that θ is the true distribution, then this observation cannot also confirm some other distribution θ'. Idealized long sequences of observations should be sufficient to resolve any uncertainty as to which θ is the true one.

The definition formalizes the idea that statistical tests can provide an objective measurement mechanism for probability that is independent of subjective judgement or personal values. Thesis definition does not define risk or how it differs from uncertainty; for that, we need more structure.

3.3 Exchangeability

The distinction between risk and uncertainty makes sense only if we believe there are invariant mechanisms that connect past experiences with future outcomes. If the different cases or instances of a problem were completely disconnected from each other, then, as Knight's quote above suggests, it makes little sense to separate the two.

A central idea, associated with Bruno de Finetti, for connecting different experiments is that of *exchangeability*.[28] Suppose that a medical treatment with uncertain consequences is to be applied in a set of cases. A theory about the effectiveness of this medical treatment is a belief P about how the outcomes are generated. Exchangeability is the idea that this theory is not affected by how the cases are labeled. If we initially label the cases $1, 2, \ldots$ then change our mind and switch the labels of cases 3 and 10, say, then the distribution P remains unchanged. The formal definition is that P is exchangeable if it is invariant to permutations of labels.

Exchangeability provides a mathematically precise way to express the idea that the cases considered are similar. The fact that the labels do not matter means that there is nothing special or remarkable distinguishing one case from all the other cases it is exchangeable with. Therefore, the regularities that govern the outcome of one case must be the same as those governing all others.

In addition to its important and deep consequences in statistics, exchangeability also provides a powerful way to separate risk and uncertainty. If we are willing to assume that the experiments are exchangeable, then a good candidate for measurable risks is the class Θ of distributions under which outcomes are independent and identically distributed (i.i.d.). It is clear that any i.i.d. $\theta \in \Theta$ is exchangeable. The following properties of Θ can also be established:

1. Θ is objectively measurable, in the sense of Definition 1.

2. Every exchangeable $\mathcal{P}$ can be uniquely expressed as a subjective uncertainty about a true but unknown θ.

[28] The classic reference is de Finetti (1937), for which English translations are available. Exchangeability is a vast topic in probability and statistics and is covered in any advanced textbook on these subjects. For an introduction with a decision-theoretic interpretation, Chapter 11 in Kreps (1988) is unsurpassed.

3. If Θ' is any other set of exchangeable distributions satisfying 2, then $\Theta' = \Theta$.

The formal statements and proofs of the above assertions are technical. Items 2 and 3 are an informal statement of a fundamental theorem by de Finetti. See Al-Najjar and Shmaya (2015, 2013) for discussion and proofs in a more general context.

In an exchangeable context, the θ's correspond to objective, measurable risks. Imperfect knowledge of θ corresponds to uncertainty. This distinction equates risk with confidence in one's knowledge of the odds: if one is prepared to proclaim a distribution for an experiment's outcome independently of how other similar experiments turn out, then it is reasonable to describe this situation as one where "the odds are known." The requirement that outcomes are independent under risk captures this intuition: given θ, the outcome of one experiment is uninformative about the outcome of another. To say that we are dealing with known probabilities means that we know all there is to be known about how outcomes are generated. Knightian uncertainty, on the other hand, refers to situations where the odds are unknown. Under exchangeability, this is represented by a subjective uncertainty about the true θ.

The proposed distinction between risk and uncertainty is not axiomatic. There is no claim, similar to the one made for the Savage postulate, that this distinction rests on self-evident principles that would be absurd to violate or contradict. Rather, the distinction builds on plausible assumptions about the world, assumptions that reasonable people may question, amend, or disagree on (is the number of experiments large enough? is exchangeability a sensible assumption?). On the other hand, not everything needs to be axiomatic and, given the high standards we set for "axiomaticness," few things are or should be.

4 Rational Precautionary Policies

This sections examines the common misconception that a Bayesian framework cannot be used to justify precautionary action.

4.1 Rational Knightian Uncertainty: A Simple Example

4.1.1 A Repeated-Urn Problem

I start with an example based on Halevy and Feltkamp (2005). An urn contains 100 balls, each colored either white or red. Let $\theta \in \{0, 1, \ldots, 100\}$ denote the number of white balls. A decision maker samples the urn *twice*, with replacement. Let x_1 denote the outcome of the first draw and x_2 the outcome of the second.

A bet on white means that the decision maker receives 10 (dollars) each time a white ball is drawn and 0 (dollars) if red. Betting on red is interpreted similarly. The set of consequences in this problem consists of all pairs (x_1, x_2) so (10,0), for example, means receiving 10 in the first draw and 0 in the second.

There are two types of urn:

The Risky Urn is known to contain 50 white and 50 red balls ($\theta = 50$).

The Uncertain Urn has unknown composition (*i.e.,* θ is unknown).

The urn example appears in Knight (1921) and was subsequently popularized by Ellsberg (1961). The only difference here is that the urn is sampled twice instead of once.

A Bayesian decision maker chooses a color and an urn to bet on. To make this choice, he needs to specify a system of values in the form of a utility function $U(x_1, x_2)$ and to quantify his uncertainty in the form of a probability distribution P. We start with the probabilities. For the risky urn, the joint distribution, denoted $P^{(r)}$, is fully specified and is given in Table 1.

	$x_2 = 10$	$x_2 = 0$
$x_1 = 10$	0.25	0.25
$x_1 = 0$	0.25	0.25

Table 1 – Joint Distribution $P^{(r)}$ of the Risky Urn

For the uncertain urn, a Bayesian decision maker views θ as a random variable with a subjective distribution μ. To make an apples-to-apples comparison, assume that our Bayesian decision maker believes the 101 possible values of θ to be equally likely. The joint distribution, denoted $P^{(u)}$, appears in Table 2.

	$x_2 = 10$	$x_2 = 0$
$x_1 = 10$	0.335	0.165
$x_1 = 0$	0.165	0.335

Table 2 – Joint Distribution $P^{(u)}$ of the Uncertain Urn

Although $P^{(u)}$ and $P^{(r)}$ are obviously different, it is still the case that the marginal probability of drawing a white ball from either urn is the same:

$$P^{(u)}(x_i = 10) = P^{(r)}(x_i = 10) = 0.50, \text{ for } \ i = 1, 2. \tag{1}$$

4.1.2 Aggregative Utility

Halevy and Feltkamp (2005) consider the above setting and assume:

Aggregative Utility: The decision maker values payoffs according to:

$$U_{\text{aggregative}}(x_1, x_2) = u(x_1 + x_2),$$

where u is strictly concave.

Since utility is a function of the sum $x_1 + x_2$, the decision maker will display *uncertainty aversion*, in the sense of a strict preference to bet on the risky urn, and *hedging*, in the sense of strict preference to randomize the choice of color. To understand the intuition, label the three relevant outcomes as:

Outcome 0: two red balls.

Outcome 10: one white ball and one red ball, in any order.

Outcome 20: two white balls.

For the risky urn, the intermediate outcome 10 has a probability of 0.50, while the extreme outcomes, 0 and 20, have a probability of 0.25 each. For the uncertain urn, the three outcomes have a probability about a third each. Since this represents a mean-preserving spread relative to the risky urn, a risk-averse decision maker will choose the risky urn instead. Halevy and Feltkamp (2005) also show that a decision maker will display a strict preference for randomizing his choice.

The intuition behind these findings is that uncertainty introduces a subjective correlation between outcomes. This increases the weight assigned to extreme outcomes, making the uncertain urn less desirable to a risk averse decision maker. Aggregative utility implies the decision maker cares about these extremes.

4.1.3 Separability

The above example delivers uncertainty aversion almost too easily. The point is so intuitive that it is easy to miss an important subtlety. To see this, replace the aggregative utility by:

Separable Utility: The decision maker utility is given by:

$$U_{\text{separable}}(x_1, x_2) = v(x_1) + v(x_2), \tag{$*$}$$

where v is a strictly concave utility function.

Separability is by far the more common assumption in economic, finance, and game theory models.[29]

Under separability, expected utility is additive in the two draws:

$$E_P\, U_{\text{separable}}(x_1, x_2) = E_P\, u(x_1) + E_P\, u(x_2).$$

This expression depends only on the marginal distributions of P; any information about correlation between the two draws is irrelevant. In our example, the probability of white in any given draw is 0.50 under both the risky and the uncertain urns. The decision maker will be indifferent between the two since both generate an expected utility of:

$$\underbrace{[0.50\, u(10) + 0.50\, u(0)]}_{\substack{\text{Expected utility} \\ \text{from the first draw}}} + \underbrace{[0.50\, u(10) + 0.50\, u(0)]}_{\substack{\text{Expected utility} \\ \text{from the second draw}}}.$$

Separability causes uncertainty aversion to disappear!

This example may explain the widespread confusion that a Bayesian framework cannot account for aversion to uncertainty. In the example, the

[29] Dynamic models usually assume separability, except that future payoffs are discounted. Introducing discounting would not affect the point being made here.

risky and uncertain urns correspond to different distributions, Whether the decision maker *cares* about this difference will depend on his utility. A Bayesian decision maker with a separable utility will not care about the difference because separability implies that only the marginal distributions matter.[30]

A formal analysis of aggregative utility may be found in Al-Najjar and Pomatto (2015). Using their model, one can think of uncertainty aversion and hedging as an expression of risk aversion when utility is not separable, rather than as a novel phenomenon that lives outside Savage's framework.

4.2 Bayesian Precautionary Policies

The above example suggests a rational-choice justification for precautionary action. Here I introduce a simple example, leaving further analysis for future work.

4.2.1 Cost and Effectiveness of Precautionary Action

Consider n instances of a policy problem, numbered $1, \dots, n$. In each instance, a binary outcome occurs, either 1 (good outcome) or 0 (bad outcome). Let s_i denote the 0-1 outcome of instance i.[31] In the case of a medical treatment of unknown effectiveness, n is the number of patients to whom the treatment is applied, outcome 1 indicates a successful treatment, while 0 indicates failure.

The true effectiveness of the treatment is parametrized by a number $\theta \in [0, 1]$ that indicates the probability of the good outcome in any given instance. Thus $\theta = 0.70$ means that in any given instance there is 70% chance the outcome will be good and 30% chance it will be bad. Outcomes are independent given θ.

In addition, there is a costly precautionary action $a_i \in [0, \infty)$ that can be taken to improve the odds of a good outcome. Specifically, if precaution a_i is taken and the true effectiveness is θ, then the probability of a good outcome

[30] Separability means that the agent views the problem as a sequence of isolated cases. Knight (1921) clearly recognized that this eliminates the role of uncertainty: "[w]e can only say that 'in so far as' one confronts a situation involving uncertainty and deals with it on its merits as an isolated case, it is a matter of practical indifference whether the uncertainty is measurable or not."

[31] The assumption that outcomes are binary is convenient for the example but not essential to the main point.

in instance i increases from θ to $\vartheta(\theta, a_i)$, where the function ϑ models the effectiveness of precautionary action. Finally, the cost of taking precaution level a_i is a strictly convex function $c(a_i)$.

4.2.2 Aggregative Policy Maker's Utility

Consider a planner who maximizes the following social welfare function:

$$U(a_1, \ldots, a_n; s_1, \ldots, s_n) = u\left(\frac{1}{n}\sum_{i=1}^{n}(s_i - c(a_i))\right), \tag{2}$$

with a strictly concave u. For example, if a_i denotes a precautionary measure to prevent harmful environmental gas emissions at plant i, then welfare is a concave function of the sum, reflecting the policy maker's aversion to aggregate variability. The division by n is convenient to ensure that the utility scale does not vary with n.

The (non-separable) welfare criterion U above ignores important considerations such as distributional concerns. For example, if instances were patients and outcomes measured their well-being, then the policy maker would care only about the average health outcomes, not how they vary across individuals. Incorporating distributional concerns is orthogonal to the risk-uncertainty distinction that is our main focus. All that is needed here is a welfare function that is sensitive to uncertainty about effectiveness of the treatment.

4.3 A Numerical Example

I solve a parameterized version of the model described above (the numerical calculations below are available from the author). First, since the problem is symmetric, I restrict attention to solutions with a common level of precautionary action $a_i = a$ for all i. Second, assume that n is large (in fact, a continuum) so that the law of large numbers holds. This makes it possible to replace random outcomes by their expectations. With these assumptions, the problem becomes:

$$\max_a \int_\Theta u\Big(\vartheta(\theta, a) - c(a)\Big)\, d\mu(\theta). \tag{3}$$

To complete the description of the model, assume:

1. Quadratic cost function $c(a) = \beta a^2, \beta > 0$.

2. The function ϑ takes the form:

$$\vartheta(\theta, a_i) = \theta(1 - a_i\gamma) + a_i\gamma, \tag{4}$$

so a level of precaution a_i proportionally shrinks the interval of possible θ's from [0,1] to $[a_i\gamma, 1]$.

3. CARA utility function u with risk-tolerance parameter R.[32]

We begin with the case of pure risk, defined as an environment where θ is known. Intuitively, the optimal level of precautionary action should increase as θ decreases since precaution is more effective in bad realizations of θ while there is less to gain when θ is already high. More concretely, assume:

- $\beta = 0.7$, $R = 0.2$, $\gamma = 0.5$.

With these values, the level of precautionary action depends on the value of θ, and ranges from 0 for $\theta = 1$ to 0.363 for $\theta = 0$. Figure 1 shows this dependence as a downward sloping line with vertical intercept 0.363.

Fundamental uncertainty refers to environments where the true impact of various policies is unknown. In the case of a new medical treatment, the policy maker is uncertain about the treatment effectiveness, side effects, etc. A Bayesian policy maker represents this uncertainty by a belief μ about θ. Assume:

- The policy maker's belief about θ is the uniform distribution μ on [0,1].

The optimal value of precaution under uncertainty is $a = 0.29$, shown as the horizontal line in Figure 1. It is instructive to consider the following thought experiment: What value of θ would, if known, justify a level of precaution $a = 0.29$? The answer is $\theta = 0.2$, shown in Figure 1 as the intersection between the two lines expressing the optimal levels of precaution under risk and uncertainty.

[32] That is, u is a constant absolute risk aversion utility and can therefore be expressed as $u(x) = 1 - e^{\frac{-x}{R}}$. The risk tolerance parameter R is the inverse of the coefficient of absolute risk aversion. High values of R approximate risk neutrality.

INSERT FIGURE 1 ABOUT HERE.

It is easy to imagine how an outside observer might mistakenly believe that a precaution level of $a = 0.29$ is the result of pessimism rather than an expected utility calculation. The marginal distribution on any instance is $\theta = 0.5$ regardless of whether we are in a risky or uncertain environment, and $\theta = 0.5$ justifies $a = \frac{0.363}{2} = 0.181$, not $a = 0.29$.

Understanding how the high level of precaution $a = 0.29$ can follow from expected utility maximization requires subtle reasoning. First, it requires modeling uncertainty about the parameter θ as a belief about probability distributions (the θ's). Second, the social welfare function cannot be separable. If the policy maker's utility were separable, then, again appealing to symmetry, the optimal precautionary action would solve:

$$\max_a \int_\Theta \Big[\vartheta(\theta, a)v(1 - c(a)) + (1 - \vartheta(\theta, a))\, v(0 - c(a))\Big]\, d\mu(\theta).$$

Separability implies that only the marginal distributions on instances matter and the policy maker would be insensitive to correlation. In our parametric example, the optimal level of precaution under separability is only about 0.075 ! See Figure 1.

In summary, to the outside observer a precaution level of $a = 0.29$ may appear as the result of a contortion of probabilities that puts greater weight on bad outcomes. In fact, increased precaution is a consequence of a non-separable welfare criterion in an otherwise standard expected utility framework.

5 Concluding Observations

A Bayesian's life is hard work: he must specify a system of values and quantify his uncertainty about all the unknowns in his environment.[33] It should therefore not be surprising that common-sense intuitions fail and

[33] Christopher Sims quotes Don Berry as having said: "Bayesian inference is hard, in the sense that thinking is hard." This is a little unfair to non-Bayesians. Full Bayesian analysis in high-dimensional problems can place so high a burden that both modelers and lab subject will understandably resort to non-Bayesian shortcuts. The problem is not with the use of convenient shortcuts per se, but with confusing shortcuts for normatively valid principles.

errors in judgement are the norm when trying to apply the rational-choice paradigm in reasoning about uncertainty.

Human probabilistic intuition fails in settings much simpler than those encountered in public policy.[34] Criteria such as the Precautionary Principle gain currency as well-intentioned shortcuts aimed at correcting human and societal biases. The analysis of this paper leads to conclusions similar to Sunstein's (2003): "those who endorse the precautionary principle are responding to salutary political or moral motivations that the principle might be thought to embody. [...] The problem is that the precautionary principle, as applied, is a crude and sometimes perverse way to promote these various goals."

[34] Halevy's (2007) experimental results show that pessimistic or cautious choices are highly correlated with subjects' inability to apply simple rules of probability calculus.

A Appendix

For the convenience of the reader, I reproduce, in a condensed form, Ghirardato's (2002) framework and postulates. I also report on Hubmer and Ostrizek (2013) who show that Ghirardato's A7 is unnecessary.

The primitives are a state space S and consequence space X. Both are arbitrary sets. Acts are finite-valued functions $f, g : S \to X$. The decision maker is assumed to have a preference relation $\succsim_E$ for every event E representing the decision maker's ranking over acts if he is told event E occurred. The initial or unconditional preference is $\succsim = \succsim_S$. An event E is *null* if the decision maker is indifferent between any pair of acts f, g that differ only on states in E. This is the preference counterpart of sets of zero probability.

A list of Ghirardato's (2002) postulates appears below. The last three postulates are identical to Savage's. They are not controversial and not reproduced here:

A1 *For every event E, $\succsim_E$ is reflexive, transitive and complete.*

A2 *For any non-null event E and pair of acts f, g*

- *(a) $f \succsim_E g$ implies $fEg \succsim g$;*
- *(b) $fEg \succsim g$ implies $f \succsim_E g$;*

A3 *For any two consequences x, x' and non-null event E, $x \succsim x'$ if and only if $x \succsim_E x'$.*

A4 Savage's P4.

A5 Savage's P5.

A6 Savage's P6.

The completeness part of A1 is referred to in the text as complete conditional rankings. A2 (a) corresponds to the assumption that information is valuable, while (b) corresponds to consistent implementation. A3 says that taste over consequences is not affected by information. Ghirardato (2002) also imposes a form of consequentialism:

A7 *For any non-null event E and pair of acts f, g, $f(s) = g(s)$ for every $s \in E$ implies that $f \sim_E g$.*

Hubmer and Ostrizek (2013) show that this postulate is redundant, being implied by 1 and 2.

In the body of the paper I interpreted A2 in terms of options and their value. For that interpretation, a strict-preference version of A2 is more natural. Such version is equivalent to A2 as stated. In one direction, assume that A2 holds. Observe that $g = gEg$, so A2 implies:

$$g \succsim fEg \iff g \succsim_E f.$$

If $f \succ_E g$ but $g \succsim fEg$, then by the earlier observation $g \succsim_E f$; a contradiction. Similarly, suppose that $fEg \succ g$ but $g \succsim_E f$, then $g \succsim fEg$ and we also have a contradiction. In the other direction, suppose that the strict version holds. To prove A2, suppose that $f \succsim_E g$ but $g = gEg \succ fEg$. But then $g \succ_E f$; a contradiction. Finally, if $fEg \succsim g$ but $g \succ_E f$ then $g = gEg \succ fEg$; also a contradiction.

References

Al-Najjar, Nabil I. and Eran Shmaya. 2013. "Learning and Long-run Fundamentals in Stationary Environments." Northwestern University.

Al-Najjar, Nabil I. and Eran Shmaya. 2015. "Uncertainty and Disagreement in Equilibrium Models." *Journal of Political Economy* . Forthcoming.

Al-Najjar, Nabil I. and Jonathan Weinstein. 2009. "The Ambiguity Aversion Literature: A Critical Assessment." *Economics & Philosophy* 25:249–284.

Al-Najjar, Nabil I. and Jonathan Weinstein. 2015. "A Bayesian Model of Knightian Uncertainty." *Theory and Decision* 78:1–22.

Al-Najjar, Nabil I. and Luciano De Castro. 2011. Subjective Probability. In *Encyclopedia of Operations Research and Management Science.* Wiley.

Al-Najjar, Nabil I. and Luciano Pomatto. 2015. "Choice under Aggregate Uncertainty." *Theory and Decision* . Forthcoming.

Arrow, Kenneth J. and Anthony C. Fisher. 1974. "Environmental Preservation, Uncertainty, and Irreversibility." *The Quarterly Journal of Economics* 88(2):312–319.

Ashford *et al.*, Nicholas. 1998. "Wingspread Statement on the Precautionary Principle." *available at:* http://www.gdrc.org/u-gov/precaution-3.html .

Barrieu, Pauline and Bernard Sinclair-Desgagné. 2010. The Paradox of Precaution. Technical report CIRANO.

Bewley, Truman. 1986. "Knightian Decision Theory, Part I." Cowles Foundation Discussion Paper no. 807.

Bewley, Truman F. 2002. "Knightian Decision Theory, Part I." *Decisions in Economics and Finance* 25(2):79–110.

Currie, Janet and Bentley MacLeod. 2014. "Savage Tables and Tort Law: An Alternative to the Precautionary Model." *Chicago Law Review* 81:53–82.

de Finetti, Bruno. 1937. "La prévision: ses lois logiques, ses sources subjectives." *Annales de l'Institut Henri Poincaré* 7:1–68.

de Finetti, Bruno. 1974. *Theory of Probability, Vol. 1-2.* Wiley, New York.

Ellsberg, Daniel. 1961. "Risk, Ambiguity, and the Savage Axioms." *The Quarterly Journal of Economics* 75(4):643–669.

Epstein, Larry G. and Tan Wang. 1994. "Intertemporal Asset Pricing under Knightian Uncertainty." *Econometrica* pp. 283–322.

Epstein, Larry and Michel Le Breton. 1993. "Dynamically Consistent Beliefs Must be Bayesian." *Journal of Economic Theory* 61(1):1–22.

Fishburn, Peter. 1970. *Utility Theory for Decision Making.* New York: John Wiley and Sons.

Ghirardato, Paolo. 2002. "Revisiting Savage in a Conditional World." *Economic Theory* 20(1):83–92.

Gilboa, Itzhak. 2009. *Theory of Decision under Uncertainty.* Cambridge University Press.

Gilboa, Itzhak and David Schmeidler. 1989. "Maxmin Expected Utility with Nonunique Prior." *Journal of Mathematical Economics* 18(2):141–153.

Gollier, Christian and Nicolas Treich. 2003. "Decision-making under Scientific Uncertainty: The Economics of the Precautionary Principle." *Journal of Risk and Uncertainty* 27(1):77–103.

Halevy, Yoram. 2007. "Ellsberg Revisited: An Experimental Study." *Econometrica* 75(2):503–536.

Halevy, Yoram and Vincent Feltkamp. 2005. "A Bayesian Approach to Uncertainty Aversion." *Review of Economic Studies* 72(2):449–466.

Hansen, Lars P. and Thomas J. Sargent. 2001. "Acknowledging Misspecification in Macroeconomic Theory." *Review of Economic Dynamics* 4(3):519–535.

Hubmer, Joachim and Franz Ostrizek. 2013. "A Note on Consequentialism in a Dynamic Savage Framework." IHS, Vienna.

Keynes, John M. 1937. "The General Theory of Employment." *The Quarterly Journal of Economics* 51(2):209–223.

Knight, Frank. 1921. *Risk, Uncertainty and Profit.* New York: Harper.

Kopylov, Igor. 2007. "Subjective Probabilities on "Small" Domains." *Journal of Economic Theory* 133(1):236–265.

Kreps, David M. 1988. *Notes on the Theory of Choice.* Westview Press.

Lange, Andreas. 2003. "Climate Change and the Irreversibility Effect–Combining Expected Utility and Maximin." *Environmental and Resource Economics* 25:417–434.

LeRoy, Stephen F. and Larry D. Singell. 1987. "Knight on Risk and Uncertainty." *The Journal of Political Economy* 95(2):394–406.

Machina, Mark J. 1989. "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty." *Journal of Economic Literature* 27(4):1622–1668.

Quiggin, John. 2005. "The Precautionary Principle in Environmental Policy and the Theory of Choice under Uncertainty." Risk and Sustainable Management Group, University of Queensland.

Routledge, Bryan R. and Stanley Zin. 2009. "Model Uncertainty and Liquidity." *Review of Economic Dynamics* 12:543–566.

Savage, Leonard J. 1954. *The Foundations of Statistics.* New York: John Wiley & Sons Inc.

Savage, Leonard J. 1967. "Difficulties in the Theory of Personal Probability." *Philosophy of Science* 34:305–10.

Schmeidler, David. 1989. "Subjective Probability and Expected Utility Without Additivity." *Econometrica* 57(3):571–587.

Sims, Christopher A. 2001. "Pitfalls of a Minimax Approach to Model Uncertainty." *American Economic Review, Papers and Proceedings* 91(2):51–54.

Sunstein, Cass R. 2003. "Beyond the Precautionary Principle." *University of Pennsylvania Law Review* 151:1003–1058.

Sunstein, Cass R. 2007. "On the Divergent American Reactions to Terrorism and Climate Change." *Columbia Law Review* 107:503.

Vermeule, Adrian. 2012. "Rationally Arbitrary Decisions (in Administrative Law)." Harvard Law School.

Wakker, Peter P. 1988. "Nonexpected Utility as Aversion of Information." *Journal of Behavioral Decision Making* 1:169–175.

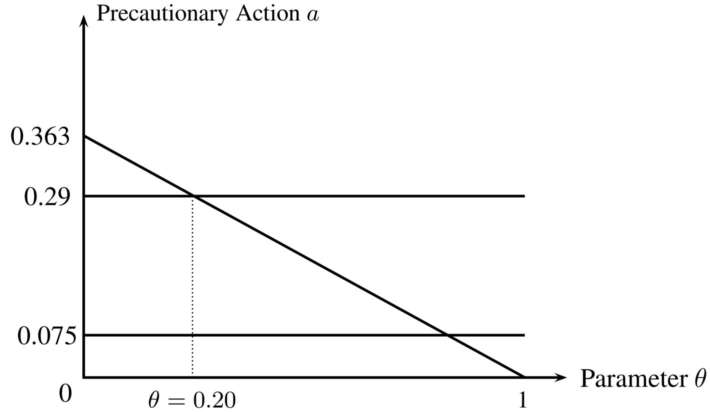

Figure 1 – Optimal Precautionary Action Under Risk and Uncertainty