



HAL
open science

Ambiguity and awareness: a coherent multiple priors model. *

Simon Grant, Ani Guerdjikova, John Quiggin

► **To cite this version:**

Simon Grant, Ani Guerdjikova, John Quiggin. Ambiguity and awareness: a coherent multiple priors model. *. 2020. hal-02550347

HAL Id: hal-02550347

<https://hal.science/hal-02550347v1>

Preprint submitted on 22 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ambiguity and awareness: a coherent multiple priors model.*

Simon Grant[†]

Ani Guerdjikova[‡]

John Quiggin[§]

April 17, 2020

Abstract

Ambiguity in the ordinary language sense means that available information is open to multiple interpretations. We model this by assuming that individuals are unaware of some possibilities relevant to the outcome of their decisions and that multiple probabilities may arise over an individual's subjective state space depending on which of these possibilities are realized. We formalize a notion of *coherent* multiple priors and derive a representation result that with full awareness corresponds to the usual unique (Bayesian) prior but with less than full awareness generates multiple priors. When information is received with no change in awareness, each element of the set of priors is updated in the standard Bayesian fashion (that is, full Bayesian updating). An increase in awareness, however, leads to an expansion of the individual's subjective state and (in general) a contraction in the set of priors under consideration.

Keywords ambiguity, unawareness, multiple priors.

JEL Classification: D81

*We thank Adam Dominiak, Peter Klibanoff, Sujoy Mukerji, Klaus Nehring, and participants at RUD2019 and at the economics seminar at UGA. We are also especially grateful to the guest editor and two anonymous referees for their patience and perseverance in reviewing earlier versions and providing us with many useful comments and suggestions that have considerably improved the clarity of our exposition. This research has been supported by IUF, Labex MME-DII and the French National Research Agency in the framework of the "Investissements d'avenir" program ANR-15-IDEX-02.

[†]Research School of Economics, Australian National University, and School of Economics, University of Queensland, AUSTRALIA Email: simon.grant@anu.edu.au.

[‡]Grenoble Applied Economics Lab, University of Grenoble Alpes, CNRS, INRAE, INP, IUF, FRANCE. Email: ani.guerdjikova@univ-grenoble-alpes.fr.

[§]School of Economics, University of Queensland, AUSTRALIA. Email: j.quiggin@uq.edu.au.

1 Introduction

The idea that choices under uncertainty are subject to problems arising from ambiguity was first put forward by Ellsberg (1961), drawing on the earlier work of Knight (2006). Like Knight, Ellsberg argued that, in many cases, decisionmakers do not, and could not be expected to, act as if they assigned well-defined numerical probabilities to all the possible outcomes of a given choice. His well-known thought experiments illustrating this argument formed the basis of a large subsequent literature both theoretical and empirical.

In most of this literature, the term ‘ambiguity’ has been treated as a synonym for what Knight called ‘uncertainty’ namely the fact that relative likelihoods are not characterized by well-defined numerical probabilities. The standard method of dealing with ambiguity in decision theory is to endow the decisionmaker with multiple priors as in Gilboa and Schmeidler (1989). This approach may be combined with a variety of preference models, notably including the maxmin model of Gilboa and Schmeidler (1989) and the smooth model of Klibanoff et al. (2005).

For a non-specialist this is puzzling; there is no obvious link to the ordinary meaning (or meanings¹) of ambiguity as a characteristic of propositions with more than one interpretation. In its normal usage, ambiguity is a linguistic concept, but in decision theory it is typically treated as a property of preferences.

The now-standard usage is quite different from that in Ellsberg’s (1961) original article. Ellsberg treated ambiguity, not as a property of preferences or relative likelihoods, but as a property of the information on which judgements of relative likelihoods might be based.

“Responses from confessed violators [of the expected utility (EU) axioms] indicate that the difference is not to be found in terms of the two factors commonly used to determine a choice situation, the relative desirability of the possible payoffs and the relative likelihood of the events affecting them, but in a third dimension of the problem of choice: the nature of one’s information concerning the relative likelihood of events. What is at issue might be called the ambiguity of this information, a quality depending on the amount, type, reliability and ‘unanimity’ of information, and giving rise to one’s degree of ‘confidence’ in an estimate of relative likelihoods.”

In this paper, we argue that informational ambiguity, in the ordinary language sense that the available information is open to multiple interpretation, may be modelled using concepts from the literature on unawareness. When individuals are unaware of some possibilities relevant to the outcome of their decisions, there are multiple probability distributions that may be applicable, depending on whether or not these possibilities are realized.

¹Empson (1930) famously distinguished seven types of ambiguity.

To represent this idea, we adopt a syntactic representation, in which the state of the world is characterized by the truth values of a finite set of elementary propositions P . The state space Ω is given by the set of all logically possible combinations of truth values, that is, by the truth table for P . Events in Ω correspond to the truth values of compound propositions in P .

An unboundedly rational decision-maker is aware of all the propositions in P , and in the logical closure of P , and therefore has access to a complete and unambiguous description of the state space Ω . If her unconditional and conditional preferences over acts (mappings from the state space to a set of consequences) conform to expected utility theory, then it is as if the decision-maker can assign a unique subjective probability π to any event E , and update that probability in line with Bayes rule as new information is received.

In contrast, we represent a boundedly rational decision-maker as one who is unaware of at least some propositions in P . For simplicity, consider the case when an agent is aware of a proposition p , but not of a related proposition q . In this situation, the proposition p is ambiguous since it may mean either $p \wedge q$ or $p \wedge \neg q$. From the agent’s viewpoint, her information about p is incomplete, since it is open to multiple interpretations.²

In this paper, we formalize this idea to derive a *coherent multiple priors* (CMP) model. Our goals are twofold. First, we derive a representation theorem for the CMP model and show that, with full awareness, it corresponds to the usual Bayesian model. Second, we consider the problem of updating beliefs. In our setting, updating may arise in response to the receipt of new information or to increased awareness, represented as awareness of new elementary propositions p . When information is received with no change in awareness, each element of the set of priors is updated in the standard Bayesian fashion as in Ghirardato et al. (2008). An increase in awareness is represented by an expansion of the state space to which the decision maker has access, and by a corresponding contraction in the set of priors under consideration. As the decisionmaker approaches full awareness, the set of priors contracts to a singleton $\{\pi^*\}$. Relative to π^* the set of priors at any time t is made of conditional probabilities, depending on the truth values of propositions of which the decisionmaker is unaware.

Roadmap

The paper is organized as follows. We begin with a motivating example, involving a policy-maker’s response to an epidemic disease outbreak. This example illustrates the relationship between bounded awareness and ambiguous beliefs and explains the notion of coherence.

We next set up the description of the decision-making problem in both propositional (syntactic) and state-space (semantic) terms. Awareness, information and acts are defined. Under full

²Ambiguity also arises in interactive contexts as in Grant et al. (2018), where different agents may interpret the same proposition differently.

awareness, individuals have access to a set of propositions P such that the associated truth table 2^P encompasses all relevant distinctions between states of the world. Boundedly aware individuals have access to a more limited set of propositions A which gives rise to a coarser state space S^A , as well as an associated ‘complementary’ space of unawareness states, with each ‘unawareness state’ corresponding to an assignment of truth values for the propositions of which the individual is unaware.

Next we consider preferences and ambiguity. We restate the Ghirardato et al. (2004) axioms. The crucial result of this section is to show that preferences satisfying the Ghirardato et al. (2004) axioms may be derived from the preferences of a fully aware EU-maximizer, by introducing unawareness.

As in Ghirardato et al. (2004), the preference over acts on S^A has a sub-relation which is incomplete and obeys the Independence axiom. This sub-relation may be represented by a unique, closed convex set of priors: an act is preferred over another by this sub-relation if it yields greater expected utility for every prior in the representation. Hence, a DM who is unaware of some of the propositions is endowed with an ambiguous preference relation which is coherent (in a sense that will be made precise) with the expected-utility preferences of a fully aware DM with the same utility function and the appropriate unique prior.

Our new key assumption, Axiom 6, states that if an act is conditionally preferable to another by each conditional preference relation obtained by conditioning the complete expected utility preference of a fully aware agent on each ‘unawareness state’, then it is unambiguously preferred.

We next consider updating in response to increases in information and awareness. We represent updating as a two-stage process. First, in even periods, the DM gets to know an event (in the part of the state space she is aware of). Second, in odd periods the DM becomes aware of propositions she had not previously considered.

For changes in information with constant awareness, we show that the preferences we derive display prior-by-prior Bayesian updating, as in Ghirardato et al. (2008) and Pires (2002). To convey the intuition for changes in awareness, we first address the simplest case where the individual becomes aware of a single additional proposition. We show that the result is to expand the state space, dividing each existing state into two new states, one in which the newly discovered proposition is true and the other in which it is false. Conversely, any pair of priors conditioned on events that differ only on the truth value of the new proposition is replaced by a convex combination of the two. In the finite setting we have here, the state space doubles in size, while the set of priors halves. This result is shown to hold more generally for any changes in awareness consistent with our model structure.

In Section 7, we discuss important links between the concept of bounded awareness used here and problems arising in econometric and statistical theory associated with concept such as ‘latent variables’ and ‘unobserved heterogeneity’ and the techniques for their estimation as developed by

Dempster et al. (1977). The crucial idea is that the relationship between observed variables of interest may be influenced by unobserved variables. In some cases, these variables are known on theoretical grounds to be relevant but cannot be measured. A variety of classical and Bayesian methods may be applied to this case. More generally, any trained statistician understands that relationships estimated on the basis of a given data set may be rendered unreliable by the omission of relevant variables, without necessarily being aware of which variables might be relevant in a particular case. This may be seen as an example of “awareness of unawareness”.

Section 8 relates our work to the existing literature. Finally, in Section 9, we offer some concluding comments.

2 Illustrative example

Consider the regrettably topical problem of developing a public health response to an epidemic disease outbreak. The disease has only recently been discovered in the country in question, but has already had severe impacts elsewhere.

The possible options include a low-cost campaign, focusing on basic hygiene measures such as hand-washing and a high-cost response involving putting the entire population into some form of quarantine. To simplify, we will assume that the high-cost option is guaranteed to control the pandemic if applied sufficiently early before the number of cases reaches some critical proportion of the population. The policy-maker (hereafter, PM) has sufficient data to estimate the probability that the number of cases is below the critical level, which we will denote by r . For the purpose of the numerical exercise below, we will set $r = 4/5$.

The success or failure of the low-cost option depends on a range of factors, only some of which the PM is fully aware of. Consider the proposition

$q =$ “the low-cost option will contain the pandemic” and its negation $\neg q =$ “the low-cost option will result in uncontained spread”

The PM is aware that the success of the low-cost option will depend on the extent of voluntary compliance, but is not explicitly aware of other relevant factors. To make this more precise, she is aware of the propositions: $p_1 =$ “voluntary compliance will be high”. Hence, the initial set of propositions of which the decision maker is aware is $A = \{q, p_1\}$. This defines the relevant state-space: $S^A = 2^{|A|} = \{(q, p_1), (q, \neg p_1), (\neg q, p_1), (\neg q, \neg p_1)\}$

Given knowledge about the relevant population, the PM can form unambiguous statements about the probability of p_1 being true, that is whether the population is likely to comply with her advice. Given the current state of knowledge about the disease, no further information is available.

The PM understands that there are other factors relevant to q , and therefore that information about q is ambiguous in the sense described by Ellsberg. In these circumstances, some advocates

of evidence-based medicine point to randomized controlled trials (RCT) as the ‘gold standard’ for assessing interventions.³

Ideally, an RCT would yield precise estimate of the proportion $\pi(q | p_1)$ and $\pi(q | \neg p_1)$ of programs that succeed in high compliance and low-compliance populations respectively. However, there is no guarantee that RCTs conducted on different populations will yield the same results, even if these populations are similar with respect to compliance. On the contrary, it is possible that results differ too much to be consistent with the hypothesis that low-cost interventions in populations with similar compliance probabilities have the same probability of success.

Suppose, thus, that individual RCTs conducted in different countries have yielded probability estimates $\pi_k(q | p_1) \in [\underline{\pi}(p_1), \bar{\pi}(p_1)]$ and $\pi_k(q | \neg p_1) \in [\underline{\pi}(q | \neg p_1), \bar{\pi}(q | \neg p_1)]$ and that the differences are too large to be explained by chance variation. As Cowan (2020) observes:

“A positive result for treatment against control in a randomized controlled trial shows you that an intervention worked in one place, at one time for one set of patients but not why and whether to expect it to work again in a different context. Evidence based medicine proponents try to solve this problem by synthesizing the results of RCTs from many different contexts, often to derive some average effect size that makes a treatment expected to work overall or typically. The problem is that, without background knowledge of what determined the effect of an intervention, there is little warrant to be confident that this average effect will apply in new circumstances. Without understanding the mechanism of action, or what we call a theory of change, such inferences rely purely on induction.

“The opposite problem is also present. An intervention that works for some specific people or in some specific circumstances might look unpromising when it is tested in a variety of cases where it does not work. It might not work ‘on average’. But that does not mean it is ineffective when the mechanism is fit to solve a particular problem such as a pandemic situation. Insistence on a narrow notion of evidence will mean missing these interventions in favor of ones that work marginally in a broad range of cases where the answer is not as important or relevant.”

In these circumstances, the PM is cognizant of the fact that there are differences between the populations of the various countries that result in differing rates of success for the low-cost intervention, but is unaware what these differences might be.⁴ We may suppose that some of the RCT

³This view is not universally shared. It is also argued that meta-analyses of large numbers of studies provide better evidence.

⁴Here we deliberately use ‘cognizant’ instead of ‘aware’, in light of the observation made by Grant and Quiggin (2013) that the proposition ‘I am aware that there exist propositions of which I am unaware’ is self-contradictory for plausible definitions of awareness.

populations match that with that of the country in question.

How can the PM reason about the probability that the treatment will be successful for a population with known compliance rates?

One possible answer, noted by Cowen, is to impute the mean probability $\pi(q | p)$ (respectively, $\pi(q | \neg p)$) averaged across RCT studies of populations with high (low) compliance. Given that the country is a member of a particular, yet undefined, sub-class, this answer is wrong with probability 1, although the direction and magnitude of the error cannot be determined.

An alternative might be to set up a higher-order probability model. That is, the PM might impute a subjective probability w_k to the proposition ‘the relevant characteristics of my country are most similar to those of the patient population of country k ’. This yields, for compliant populations, the success probability $\pi^*(q | p_1) = \sum_k \pi_k(q | p_1) w_k$. Some version of this subjective approach would be required of an agent satisfying the Savage axioms. Yet the solution is obviously problematic. While the success rates for each country, $\pi_k(q | p_1)$, are objectively defensible, the choices of w_k are just guesses. The resulting weighted average $\pi^*(q | p_1)$ cannot be justified to the public, or to another policy-maker whose guesses w_k are different.

The third response to unawareness is to say that the success probability lies in the interval $[\underline{\pi}(p_1), \bar{\pi}(p_1)]$. With this response, bounded awareness implies ambiguous probability beliefs.

The idea that probabilistic judgements apply to large populations but not to individual cases (in our case, that of individual countries) may be traced back to Knight. It is eloquently expressed by Mukerjee (2015) as applied to medical research:

“We have invented many rules to understand normalcy - but we still lack a deeper, more uniform understanding of physiology and pathology. This is true for even the most common and extensively studied diseases - cancer, heart disease, and diabetes. ... Why do certain immune diseases cluster together in some people, while others have only one variant? Why do patients with some neurological diseases such as Parkinson, have a reduced risk on cancer?

“These ‘outlying’ questions are the Mars problems of medicine...”

In all these questions there is ‘awareness’ of several different outcomes, but ‘unawareness’ of what the cause of these different outcomes might be, and how they apply in particular cases.

Suppose indeed, that the reality in our epidemic example is more complicated: whether or not a low-cost program succeeds depends not only on the population compliance, p_1 but also on whether household size is small, p_2 . What is more, these two factors are interrelated.

In particular, traditional societies with large families are more inclined to follow directives from established authorities. Suppose no other factors are relevant.

Thus the set of propositions $P = (q, p_1, p_2)$ defines an eight element ‘small world’ in the sense of Savage (1954).

The following table gives the joint probability of the validity of the relevant propositions, p_1 and p_2 for the population in question.

	p_2 : small household	$\neg p_2$: large household	
p_1 : compliant.	1/8	3/8	(1)
$\neg p_1$: not compliant	1/3	1/6	

For later reference, $\Pr(p_1|p_2) = 3/11$, $\Pr(p_1|\neg p_2) = 9/13$.

Given the interplay between p_1 , p_2 and q explained above, we specify the following conditional probabilities of q given the truth realizations of p_1 and p_2 :⁵

	p_2 : small household	$\neg p_2$: large household	
p_1 : compliant.	$\Pr\{q p_1, p_2\} = 1$	$\Pr\{q p_1, \neg p_2\} = 1/2$	(2)
$\neg p_1$: not compliant	$\Pr\{q \neg p_1, p_2\} = 1/2$	$\Pr\{q \neg p_1, \neg p_2\} = 0$	

This implies that the joint probability distribution of the truth values of q , p_1 and p_2 is:

q : low-cost success	p_2 : small household	$\neg p_2$: large household	
p_1 : compliant.	1/8	3/16	(3)
$\neg p_1$: not compliant	1/6	0	
$\neg q$: low-cost failure	p_2 : small household	$\neg p_2$: large household	
p_1 : compliant.	0	3/16	
$\neg p_1$: not compliant	1/6	1/6	

and the joint (marginal) probability distribution of q , p_1 is:

	q : low cost success	$\neg q$: failure	
p_1 : compliant.	5/16	3/16	(4)
$\neg p_1$: not compliant	1/6	1/3	

A fully aware PM can thus state the unconditional probability of q as $\Pr\{q\} = 23/48$.

In this context, learning the truth value of p_2 generates a partition

$$\mathcal{F}_1 = \left\{ \begin{array}{l} \{(q, p_1, p_2), (q, \neg p_1, p_2), (\neg q, p_1, p_2), (\neg q, \neg p_1, p_2)\}, \\ \{(q, p_1, \neg p_2), (\neg q, p_1, \neg p_2), (q, \neg p_1, \neg p_2), (\neg q, \neg p_1, \neg p_2)\} \end{array} \right\}$$

⁵In the general model specified in Section 6, we require that all relevant events be non-null so as to avoid discussing updating of beliefs upon the occurrence of a null event. For the purposes of the example, we assume that some of the states do occur with probability 0. Note that this does not generate a contradiction in as far as the example does not require updating conditional on such 0-probability events.

and the updated conditional probabilities are calculated as usual, e.g.,

$$\Pr \{q, p_1 \mid p_2\} = \frac{\Pr \{(q, p_1, p_2)\}}{\Pr \{(q, p_1, p_2)\} + \Pr \{(q, \neg p_1, p_2)\} + \Pr \{(-q, p_1, p_2)\} + \Pr \{(-q, \neg p_1, p_2)\}}.$$

Hence, for each realization of the truth value of p_2 , we obtain the probability distribution on the state space S^A :

$$\begin{aligned} \Pr \{q, p_1 \mid p_2\} &= 3/11 & \Pr \{q, p_1 \mid \neg p_2\} &= 9/26 \\ \Pr \{\neg q, p_1 \mid p_2\} &= 0 & \Pr \{\neg q, p_1 \mid \neg p_2\} &= 9/26 \\ \Pr \{q, \neg p_1 \mid p_2\} &= 4/11 & \Pr \{q, \neg p_1 \mid \neg p_2\} &= 0 \\ \Pr \{\neg q, \neg p_1 \mid p_2\} &= 4/11 & \Pr \{\neg q, \neg p_1 \mid \neg p_2\} &= 4/13 \end{aligned} \tag{5}$$

Now consider a partially aware PM. We suppose that the PM entertains two priors over S^A , corresponding to the possible truth values of p_2 . These priors represent the PM's perception of the ambiguity of the information she has available concerning S^A .

We will be particularly concerned with the case of coherent priors. Coherent priors correspond to the conditional probabilities the agent would assign if she were actually aware of the truth values. The underlying idea is that, while the PM cannot articulate p_2 , her beliefs on S^A correctly reflect the beliefs she would hold if she were fully aware. That is, each row in (5) defines an interval of probabilities a PM may entertain about the occurrence of a state in S^A .

In our example, the probability distribution in the right column in (5) corresponds to a data set from an RCT study conducted in a country with small households whereas the left column would correspond to RCT data generated in a country with large households.

Thus, with $k \in \{1, 2\}$, where p_2 is true for country 1 and $\neg p_2$ holds for country 2, we would have

$$\begin{aligned} \pi_1(q \mid p_1) &= 1, \pi_1(q \mid \neg p_1) = \frac{1}{2} \\ \pi_2(q \mid p_1) &= \frac{1}{2}, \pi_2(q \mid \neg p_1) = 0 \end{aligned}$$

Realistically, for most countries, the proportions of households would be between the extreme values, giving rise to a convex set of probabilities.

The range of possible probabilities over S^A are given by an interval in $[0, 1]^4$ with extreme points given by the two conditional distributions $(3/11, 0, 4/11, 4/11)$ and $(9/26, 9/26, 0, 4/13)$ identified in (5).

Upon learning the truth value of p_1 , that is, the compliance characteristics of the population, and using full Bayesian updating on the set of priors, the beliefs over the value of q reduce to:

	q : low-cost program successful
p_1 : compliant.	$\Pr \{q \mid p_1\} \in [\pi_1(q \mid p_1), \pi_2(q \mid p_1)] = [1/2, 1]$
$\neg p_1$: not compliant	$\Pr \{q \mid \neg p_1\} \in [\pi_2(q \mid \neg p_1), \pi_1(q \mid \neg p_1)] = [0, 1/2]$

(6)

where the extreme points of the interval in each case reflect the two distributions conditional on the truth of p_2 presented in Table 2.

The resulting convex set of priors is illustrated in Figure 1. The upper red line shows the set of priors for the pair $(\Pr(p_1), \Pr(q | p_1))$. Point A represents the prior belief derived from the implicit assumption of p_2 . Point B represents the prior belief derived from the implicit assumption of $\neg p_2$.

Similarly, the lower blue line represents the set of priors for the pair $(\Pr(p_1), \Pr(q | \neg p_1))$. Point A' represents the prior belief conditional on the implicit assumption of p_2 . Point B' represents the prior belief conditional on the implicit assumption of $\neg p_2$.

The information represented by Points A and A' is sufficient to represent the prior belief over the state space $\{(q, p_1), (q, \neg p_1), (\neg q, p_1), (\neg q, \neg p_1)\}$ derived from the implicit assumption of p_2 . Similarly, the information represented by Points B and B' is sufficient to represent the prior belief derived from the implicit assumption of $\neg p_2$.

Now suppose our PM becomes aware of p_2 and assigns the ‘correct’ prior probability $11/24$ to p_2 being true. (If the available data is representative, this prior could also be inferred from the data, once the truth value of p_2 has been measured for each observation). Now the two priors on S^A , $(3/11, 0, 4/11, 4/11)$ and $(9/26, 9/26, 0, 4/13)$, are collapsed into one, namely that given by (4):

$$\begin{aligned} & \Pr(p_2) \pi_1(q | p_1) + \Pr(\neg p_2) \pi_2(q | p_1) \\ &= \frac{11}{24} \left(\frac{3}{11}, 0, \frac{4}{11}, \frac{4}{11} \right) + \frac{13}{24} \left(\frac{9}{26}, \frac{9}{26}, 0, \frac{4}{13} \right) = \left(\frac{5}{16}, \frac{3}{16}, \frac{1}{6}, \frac{1}{3} \right) \end{aligned}$$

and the interval convex hull shrinks to a single point.

Note that the so-obtained probability distribution will coincide with π^* if and only if the relative frequency of p_2 observed in the RCTs, coincides with the actual probability of p_2 for the population in question.

In Figure 1, upon becoming aware of p_2 , the pairs (A, B) and (A', B') are replaced by the single priors C and C' respectively, where $C = \lambda A + (1 - \lambda) B$ with $\lambda = \Pr(p_2 | p_1)$, and similarly $C' = \lambda' A' + (1 - \lambda') B'$ with $\lambda' = \Pr(p_2 | \neg p_1)$. The vertical axis values of C and C' are the prior conditional probabilities $\Pr(q | p_1)$ and $\Pr(q | \neg p_1)$, taking account of available information about p_2 .⁶

As we will show below, under appropriate conditions on beliefs and preferences, this coherence result holds in general. Becoming aware of some new propositions, while remaining unaware of others, and updating each element of the set of coherent priors according to Bayes rule yields a set of posteriors which is the same as the set of coherent priors obtained by beginning with expanded awareness and information, and then deriving priors as above. Moreover, the derivation above shows

⁶Similarly, the horizontal axis values of C and C' are convex combinations of the probability of p_1 conditional on the realization of p_2 with the same coefficients λ and λ' . The interpretation of these values is more complicated due to the fact that the figure is attempting to represent an 8-dimensional state space in a two-dimensional graph.

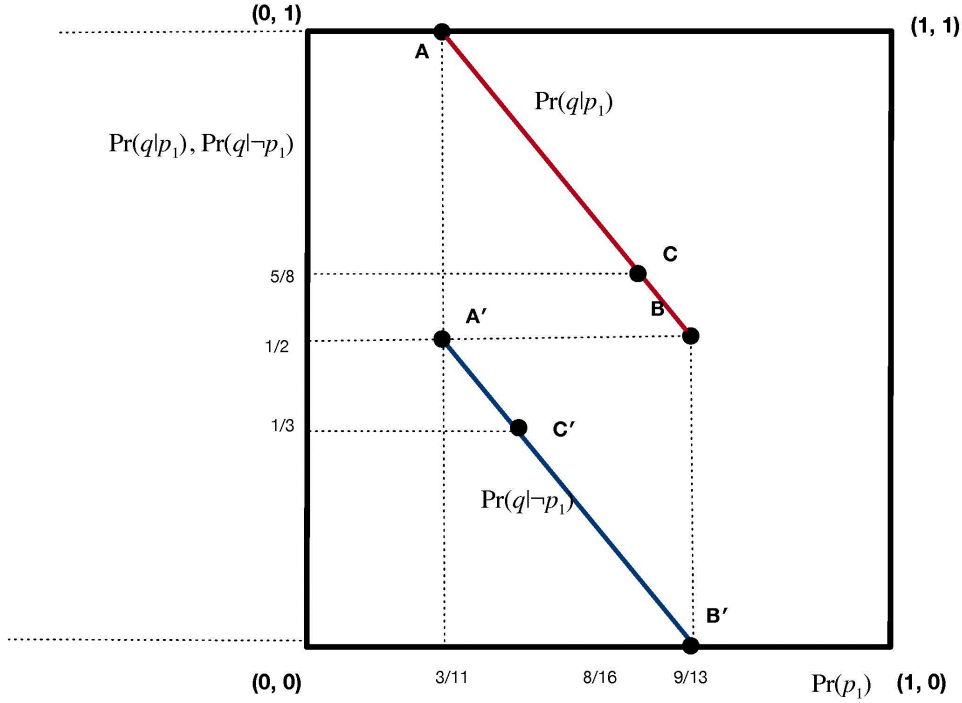


Figure 1: Convex sets of priors for illustrative example

that coherence is, in principle, testable. Once the PM is fully aware, a probability distribution over Ω can be elicited, and this is sufficient to derive the set of priors associated with any state of partial awareness and any information set.

Now consider the choice faced by the PM, between the low-cost and high-cost responses. We may consider three possible outcomes: an uncontrolled pandemic, a successful high cost response and a successful low-cost response. We will associate these outcomes with payoffs 0, $1/2$, and 1.

The high-cost option α_h generates an unambiguous lottery yielding $1/2$ with probability $r = 4/5$ and 0 with probability $1 - r = 1/5$. The expected utility of this lottery is $2/5$.

Under full awareness, the low cost option α_ℓ also generates an unambiguous lottery yielding 1 with probability $23/48$ and 0 with probability $25/48$.

By contrast, a PM unaware of p_2 entertains the two priors $(3/11, 0, 4/11, 4/11)$ and $(9/26, 9/26, 0, 4/13)$.

If the PM is uncertain about the degree of voluntary compliance of the population (the validity of p_1), the MMEU of α_ℓ is determined using the minimal probability of success for the low-cost option, $\Pr\{q\} = \min\{7/11, 9/26\} = 9/26$.

With these parameter values

$$9/26 < 2/5 < 23/48.$$

That is, under full awareness the low-cost option α_ℓ will be preferred to the high cost option, but

under partial awareness with MMEU preferences, the high-cost option α_h will be preferred. The PM, realizing that there are relevant factors of which she is yet unaware, might thus prefer to choose a costlier and more aggressive course of action, which would however help control the pandemic with a known probability r , rather than facing the worst possible scenario, in which the unknown factor drives the probability of success to its lowest possible value $9/26$. This can be interpreted as an instance of the precautionary principle in face of unawareness.

3 Setup

3.1 The state of the world: propositional and state-space descriptions

Information and awareness evolve over time. However, we will initially consider an individual with fixed information and awareness, suppressing time subscripts.

The world is described by the truth values of a finite set of elementary propositions $P = \{p_1, \dots, p_N\}$. Compound propositions formed by conjunction and negation of elementary propositions are denoted by boldface \mathbf{p} .

Individuals have bounded awareness, represented by a set $A \subseteq P$ of elementary propositions which they can express. Awareness and information are mutually dependent. On the one hand, as will be described in more detail below, the individual's awareness depends on the information they have. On the other hand, that information must be expressed in terms of propositions $p \in A \subseteq P$ expressible by the individual.

This propositional description of the world may be represented equivalently in state space terms more familiar to decision theorists. The state space associated with the truth table for P may be represented by $\Omega = 2^N$ with ω a representative element/state.⁷ Let Σ denote the field of subsets of Ω with elements E .

Similarly, for an individual with awareness A the state space of which she is aware can be expressed as

$$S^A = 2^A,$$

with a generic element s^A and field of subsets Σ^A with elements E^A , referred to as events. Note that each event E^A corresponds to the truth of a compound proposition \mathbf{p} made up of elementary propositions in P^A .

Clearly, each state $\omega \in \Omega$ uniquely identifies a state $s^A \in S^A$ by its projection to S^A denoted by ω_{S^A} . Conversely, each $s^A \in S^A$ corresponds to an event in Ω , that is, the set of states ω whose

⁷Since a state specifies the truth value of each proposition in P , it can be expressed as a binary number, i.e., $\omega \in \{0, 1\}^N$. While such a representation may be useful in some contexts, it is mostly unnecessary for the purposes of this paper and is thus evoked only once in Section 6.2.

projection to S^A is exactly s^A : $E_{s^A} = \{\omega \mid \omega_{S^A} = s^A\}$. Such events are measurable with respect to Σ^A . More generally, an event $E \subseteq \Omega$ is measurable with respect to Σ^A iff

$$\omega \in E \text{ implies } \omega' \in E \text{ for all } \omega' \in \Omega \text{ with } \omega_{S^A} = \omega'_{S^A}. \quad (7)$$

We define a probability as a mapping $\pi : \Sigma^A \rightarrow [0, 1]$ with the usual properties, and define the conditional probability operator $\pi(\cdot \mid E)$ according to Bayes rule

$$\pi(E' \mid E) = \frac{\pi(E \cap E')}{\pi(E)}$$

whenever $\pi(E) > 0$.

Note that a probability measure π on Σ induces a unique probability measure $\tilde{\pi}$ on Σ^A by setting

$$\tilde{\pi}(E) = \pi\{\omega \mid \omega_{S^A} \in E\}.$$

Let $\bar{A} = P \setminus A$ be the set of propositions of which the individual is unaware. The ‘complementary state space’ of which she is *unaware* can be expressed as $\bar{S}^A = 2^{\bar{A}}$, with generic element \bar{s}^A . Notice that $\Omega = S^A \times \bar{S}^A$.

Furthermore, each (awareness) state $s^A \in S^A$ corresponds to the event $\{s^A\} \times \bar{S}^A$ in Ω and each $\bar{s}^A \in \bar{S}^A$ corresponds to the event $S^A \times \{\bar{s}^A\}$ in Ω . That is, awareness of the form described above leads to a ‘coarsening’ of the state space as, for example, in Quiggin (2016), represented by the projection of Ω onto S^A . In the case of full awareness, \bar{S}^A must be a singleton and similarly for S^A in the case of maximal unawareness.

Unawareness in this sense may be distinguished from the case of ‘reduction’ or ‘restriction’ of the state space, in which some possible elements of Ω are disregarded or, equivalently, in which some propositions that are possibly true are implicitly assumed to be false. This leads to the possibility of ‘surprise’, see, for example, Grant and Quiggin (2015).

3.2 Acts

Acts will be represented formally in the usual Anscombe-Aumann framework. To understand how this framework is applicable in situations of differential awareness, it is useful to recall that each state and event in Ω corresponds to the truth value of a (typically compound) proposition.

In order to focus on beliefs rather than preferences over outcomes, we concentrate on a set comprising just two final consequences $Z = \{0, 1\}$.⁸ Hence, each simple act may be expressed in

⁸In this paper, we abstract from the problem of how the agent deals with ambiguity and explore the relation between the perception of ambiguity and partial awareness concentrating on agent’s beliefs. Since only two outcomes are necessary for belief elicitation, the binary outcome assumption is without loss of generality and all results in the paper hold for more general finite sets of outcomes. Using a richer set of outcomes would allow us (as in the standard Anscombe-Aumann setting) to uncover the agent’s risk attitude, and, under some additional conditions, as for instance in Ghirardato et al. (2004) elicit his attitude towards ambiguity.

propositional terms as a bet of the form “If \mathbf{p} , receive 1, otherwise receive 0”. The extension to general outcome sets is straightforward but adds complexity without additional insight.

Let $\Delta(Z)$ denote the set of all lotteries on Z , with a generic element denoted by $x \in [0, 1]$, yielding the (‘good’) outcome 1 with probability x and the (‘bad’) outcome 0 with complementary probability $1 - x$.

An act α maps Ω to $\Delta(Z)$. The set of acts available under full awareness is denoted \mathcal{A} , and is equal to $\Omega^{\Delta(Z)}$. Let \mathcal{C} denote the set of all constant acts. Let \mathcal{B} denote the set of ‘bets (on events)’, that is, $\alpha \in \mathcal{B}$ if and only if $\alpha(\omega) \in \{0, 1\}$ for all $\omega \in \Omega$.

The outcomes of acts considered by an individual with limited awareness must be conditional on propositions of which the individual is aware. Hence, for given awareness $A \subseteq P$, any act α must be measurable with respect to Σ^A , and we denote by $\mathcal{A}(A)$ the subset of such acts. Formally:

$$\mathcal{A}(A) = \{\alpha : \Omega \rightarrow [0, 1] \mid \alpha(\omega) = \alpha(\omega') \text{ for all } \omega, \omega' \text{ with } \omega_{S^A} = \omega'_{S^A}\}.$$

Clearly, an agent with limited awareness cannot perceive such mappings. Hence, the set of acts available to an agent with awareness A is given by:

$$\mathcal{A}^A = \{\alpha : S^A \rightarrow [0, 1]\}.$$

Note that there is a one-to-one correspondence between the acts in $\mathcal{A}(A)$ and those in \mathcal{A}^A , but that the mappings defining these acts have different domains (Ω vs. S^A). Analogously, \mathcal{C}^A denotes the set of constant acts \mathcal{A}^A and \mathcal{B}^A denotes the set of bets in \mathcal{A}^A .

Each act in \mathcal{A}^A induces a mapping of elements of S^A into lotteries on Z . In particular, for any s^A an act α specifies the probability of obtaining 1, which with slight abuse of notation we shall denote by $\alpha(s^A)$.

As is standard, convex mixtures of acts are defined as state-by-state probability mixtures: for α and $\alpha' \in \mathcal{A}$, and $\lambda \in (0, 1)$, we define $\lambda\alpha + (1 - \lambda)\alpha'$ by

$$(\lambda\alpha + (1 - \lambda)\alpha')(\omega) = \lambda\alpha(\omega) + (1 - \lambda)\alpha'(\omega)$$

for all $\omega \in \Omega$. Convex mixtures in \mathcal{A}^A are defined analogously with respect to the state space S^A .

Example Suppose that an agent’s initial awareness consists of a single proposition, $A = \{p_1\}$, and that the agent later becomes aware of the full set of propositions $P = \{p_1, \dots, p_N\}$. In the initial condition of partial awareness, the state space is $S^{\{p_1\}} = S^A = \{p_1, \neg p_1\}$, and an act may be represented as an ordered pair of consequences: $(\alpha(p_1), \alpha(\neg p_1))$. There are only four simple acts available to the agent: the two constant acts $(0, 0)$ and $(1, 1)$ the bet on p_1 $(1, 0)$ and the bet against p_1 , $(0, 1)$. The set of acts $\mathcal{A}^A = [0, 1]^2$ consists of convex mixtures over these simple acts. Trivially, each of these acts corresponds to an act in Ω measurable with respect to Σ^A . For example, the bet on p_1 can be represented as $\alpha(\omega) = 1$ if p_1 is true in ω and $\alpha(\omega) = 0$ if $\neg p_1$ is true in ω . The logic of this simple example is fully general.

4 Preferences and ambiguity

For each level of awareness $A \subseteq P$ we define preferences on \mathcal{A}^A by \succsim^A . We will denote $\succsim = \succsim^P$ preferences under full awareness. That is, we consider an agent with a family of preference orderings over state spaces S^A , one for each A .

4.1 The GMM approach

We first impose the Ghirardato et al. (2004) (hereafter, GMM) axioms: for each \mathcal{A}^A , we assume the preference relation \succsim^A satisfies:

Axiom 1 *Completeness and transitivity*

Axiom 2 *Archimedean axiom: for all $\alpha, \alpha', \alpha'' \in \mathcal{A}^A$, if $\alpha \succ^A \alpha' \succ^A \alpha''$, then there are λ and $\mu \in (0, 1)$ such that*

$$\lambda\alpha + (1 - \lambda)\alpha'' \succ^A \alpha' \succ^A \mu\alpha + (1 - \mu)\alpha''.$$

Axiom 3 (a) *Certainty independence: for $\alpha, \alpha' \in \mathcal{A}^A$ and $\bar{\alpha} \in \mathcal{C}^A$, $\alpha \succsim^A \alpha'$ iff $\lambda\alpha + (1 - \lambda)\bar{\alpha} \succsim^A \lambda\alpha' + (1 - \lambda)\bar{\alpha}$ for all $\lambda \in [0, 1]$.*

(b) *Independence under full awareness: if $A = P$, then for any $\alpha, \alpha', \alpha'' \in \mathcal{A}$, $\alpha \succsim^P \alpha'$ iff $\lambda\alpha + (1 - \lambda)\alpha'' \succsim^P \lambda\alpha' + (1 - \lambda)\alpha''$ for all $\lambda \in [0, 1]$.*

Axiom 4 *Monotonicity: for $\alpha, \alpha' \in \mathcal{A}^A$ if $\alpha(s^A) \geq \alpha'(s^A)$ for all $s^A \in S^A$, then $\alpha \succsim^A \alpha'$.*

Axiom 5 *Non-degeneracy: there are α and $\alpha' \in \mathcal{A}^A$ such that $\alpha \succ^A \alpha'$.*

Axiom 2 is purely technical, Axiom 4 is intuitive and uncontroversial, while Axiom 5 ensures the setting and the analysis that follows is not vacuous. Axiom 1 is of more interest. The standard completeness axiom, adopted by GMM, is very demanding, since it requires agents to have complete preferences over all conceivable acts with respect to a state space which may be arbitrarily large. Since we are restricting the axiom to apply to acts expressed in terms of propositions of which the agent is explicitly aware, our Axiom 1 is consistent with reasonable bounds on cognitive capacity. The same reasoning extends to transitivity.

For the case of partial awareness $A \subset P$, Axiom 3 is less restrictive than the standard independence axiom (as in GMM) except in the special case of full awareness, for which independence does hold. This implies that in our setting partial awareness is the only source of deviations from expected utility.

Definition 1 *For $\alpha, \alpha' \in \mathcal{A}^A$, α is \succsim^A -unambiguously preferred to α' , if*

$$\lambda\alpha + (1 - \lambda)\alpha'' \succsim^A \lambda\alpha' + (1 - \lambda)\alpha'' \text{ for all } \lambda \in [0, 1] \text{ and all } \alpha'' \in \mathcal{A}^A. \quad (8)$$

Whenever (8) is satisfied for two acts α and α' , we write $\alpha \succ_*^A \alpha'$. As in GMM, the relationship \succ_*^A is a preorder satisfying all the Anscombe-Aumann axioms (except completeness), and so admits a representation as given in the Lemma below, GMM, Proposition 5, p. 144. For the case of full awareness, this relationship is complete.

Lemma 1 *Axioms A1-A5 imply for each A the existence of a unique convex and (weak*) closed set of priors Π^A such that for any two acts α and $\alpha' \in \mathcal{A}^A$*

$$\alpha \succ_*^A \alpha' \text{ iff } \sum_{s^A \in S^A} \pi(s^A) \alpha(s^A) \geq \sum_{s^A \in S^A} \pi(s^A) \alpha'(s^A) \text{ for all } \pi \in \Pi^A. \quad (9)$$

In particular, Π^P is a singleton $\{\pi^*\}$.

Remark 1 *Just as in Anscombe-Aumann, for the case of full awareness, Axiom 3(b) allows us to define conditional preferences and thus, identify conditional beliefs $\pi^*(\cdot | E)$ which satisfy Bayes rule.*

Since each constant act can be identified with the probability $x \in [0, 1]$, with which it results in the outcome 1 in every state, we write $\alpha_x \in \mathcal{C}$. As in GMM (pp. 153-4), for $\alpha \in \mathcal{A}^A$, define the interval of possible certainty equivalents:

$$CE^*(\alpha) = \left\{ x \in [0, 1] \mid \text{for any } y \in [0, 1], \alpha_y \succ_*^A \alpha \text{ implies } y \geq x \text{ and } \alpha \succ_*^A \alpha_y \text{ implies } x \geq y \right\}.$$

GMM show that $x \in CE^*(\alpha)$ iff

$$\min_{\pi \in \Pi^A} \sum_{s^A \in S^A} \pi(s^A) \alpha(s^A) \leq x \leq \max_{\pi \in \Pi^A} \sum_{s^A \in S^A} \pi(s^A) \alpha(s^A).$$

Interpreting constant acts x as utilities, and noting that we have restricted utilities to the unit interval, $CE^*(\alpha)$ corresponds to the set of utilities for α consistent with some $\pi \in \Pi^A$. In the case of full awareness, $CE^*(\alpha)$ is a singleton.

4.2 Unawareness and ambiguity

The preferences described in the previous section are normally interpreted in terms of ambiguity. Given our setup, there is a natural interpretation in terms of awareness.

Consider EU preferences over \mathcal{A} , \succ , described by the full awareness probability distribution π^* derived using Axiom 3(b) in the previous section.⁹ In this setting, conditional preferences can be defined as usual.

⁹Without loss of generality, we assign utilities 0 and 1 to the final consequences 0 and 1, respectively.

Definition 2 For $\hat{\alpha}, \hat{\alpha}' \in \mathcal{A}$, $\hat{\alpha}$ is preferred to $\hat{\alpha}'$ conditional on event $E \in \Sigma$, denoted $\hat{\alpha} \succsim_E \hat{\alpha}'$ iff $\tilde{\alpha} \succ \tilde{\alpha}'$ holds for some (and thus, by Axiom 3(b), for all) $\tilde{\alpha}, \tilde{\alpha}' \in \mathcal{A}$ with $\tilde{\alpha}(s) = \hat{\alpha}(s)$, $\tilde{\alpha}'(s) = \hat{\alpha}'(s)$ for all $s \in E$ and $\tilde{\alpha}(s) = \tilde{\alpha}'(s)$ for all $s \notin E$.

Remark 2 The so-defined conditional preferences under full awareness can be shown to satisfy the same Axioms 1–5 as above.

For any $\bar{s}^A \in \bar{S}^A$, that is, for any set of truth values for the propositions outside of a given set $A \subseteq P$ (which describes a possible level of awareness), π^* induces a conditional probability distribution $\pi^*(\cdot | \bar{s}^A)$ over S^A . Correspondingly, the induced conditional preferences $\succsim_{\bar{s}^A}$ over $\mathcal{A}(A)$, are given by $\hat{\alpha} \succsim_{\bar{s}^A} \hat{\alpha}'$ iff

$$\sum_{s^A \in S^A} \pi^*(s^A | \bar{s}^A) \hat{\alpha}(s^A) \geq \sum_{s^A \in S^A} \pi^*(s^A | \bar{s}^A) \hat{\alpha}'(s^A).$$

Now consider the family of preferences $(\succsim^A)_A$ which satisfy Axioms 1–5. The following axiom relates the preference relation at a level of awareness A to the conditional preferences under full awareness.

Axiom 6 Unanimity: for $\alpha, \alpha' \in A^A$, $\alpha \succsim_*^A \alpha'$ if and only if for every $\bar{s}^A \in \bar{S}^A$ and every $\hat{\alpha}_{\bar{s}^A}, \hat{\alpha}'_{\bar{s}^A} \in \mathcal{A}$ with $\hat{\alpha}_{\bar{s}^A}(s \times \bar{s}^A) = \alpha(s)$ and $\hat{\alpha}'_{\bar{s}^A}(s \times \bar{s}^A) = \alpha'(s)$ for each $s \in S^A$ and $\hat{\alpha}_{\bar{s}^A}(s \times \bar{s}^A) = \hat{\alpha}'_{\bar{s}^A}(s \times \bar{s}^A)$ for all $s \in S^A$ and $\bar{s}^A \in \bar{S}^A \setminus \bar{s}^A$, $\hat{\alpha}_{\bar{s}^A} \succ \hat{\alpha}'_{\bar{s}^A}$.

Note that this condition can be equivalently rewritten in terms of the conditional preferences in Definition 2. Recall that to each act $\alpha \in A^A$ paying off on the state-space S^A , there corresponds an act $\hat{\alpha} \in \mathcal{A}(A)$ paying off on the state-space Ω and measurable with respect to Σ^A such that $\hat{\alpha}(\tilde{s}) = \alpha(s)$ for each $\tilde{s} \in E_s$ and each $s \in S^A$. Then Axiom 6 says that for any two acts $\alpha, \alpha' \in A^A$, α is unambiguously preferred to α' , $\alpha \succsim_*^A \alpha'$, if and only if for $\hat{\alpha}, \hat{\alpha}' \in \mathcal{A}(A)$ with $\hat{\alpha}(\tilde{s}) = \alpha(s)$ and $\hat{\alpha}'(\tilde{s}) = \alpha'(s)$ for each $\tilde{s} \in E_s$ and each $s \in S^A$, $\hat{\alpha}$ is preferred to $\hat{\alpha}'$ conditionally on any state $\bar{s}^A \in \bar{S}^A$, $\hat{\alpha} \succsim_{\bar{s}^A} \hat{\alpha}'$. Recalling that $\hat{\alpha}$ and $\hat{\alpha}'$ are measurable with respect to S^A , the only effect of \bar{s}^A is to determine the conditional probability distribution $\pi^*(\cdot | \bar{s}^A)$. So, we are evaluating α and α' with respect to a set of probability distributions. Axiom 6 says that if α is preferable with respect to each such distribution, then it must be preferred *unambiguously*.

This property may be viewed as a version of the sure-thing principle. Unambiguous preferences of the type $\alpha \succsim_*^A \alpha'$ here correspond to comparisons for which existing evidence is uncontroversial, i.e., for which the truth realizations of the yet unknown propositions and thus of the states in \bar{A} all yield the same preference ranking for a given choice. Axiom 6 then requires a certain consistency when the preferences of a decision maker in a state in which he is not yet aware of \bar{A} are revised upon becoming fully aware of all propositions in \bar{A} . If it is the case that the realizations of the unknown states were irrelevant for the comparison between α and α' ex-ante, then this should also

be true ex-post for every such possible realization, i.e., conditional on any possible truth value of the propositions in \bar{A} . Vice versa, if there is a truth value specification for the propositions in \bar{A} for which the comparison between α and α' is reversed, representative and accurate empirical evidence would reflect this, even if the decision maker is yet unaware of the cause for such a reversal. Thus, the decision maker would not have unambiguously preference for α versus α' at awareness level A .

Since Axiom 6 uses the conditional preferences derived from the full awareness preference relation for each level of awareness A , its effect is to tie all the preference relations \succsim^A to the fully aware preferences \succsim and thereby to each other.

4.3 The awareness-based multiple priors model

We now establish the claim that unawareness, as we have defined it, naturally lends itself to a multiple priors model to represent preferences.

Definition 3 *Multiple priors preferences are awareness-based if for each $A \subseteq P$, the set of probabilities Π^A derived in Lemma 1 satisfies:*

$$\Pi^A = \overline{CH} \{ \pi^* (\cdot | \bar{s}^A) | \bar{s}^A \in \bar{S}^A \},$$

where \overline{CH} stands for the convex hull of a set, and π^* is a probability as described in Section 1.

Clearly, for awareness-based preferences, under full awareness, Π is a singleton (recall that in this case \bar{S}^A is a singleton).

Given these definitions we have the following implication of the axioms, the proof of which is relegated to the Appendix:

Proposition 2 *Under axioms A1–A6, the family of preferences $(\succsim^A)_A$ are awareness-based.*

Notice that we do not expect the converse to hold. As we shall show, in the absence of changes in awareness, the probabilities Π derived as conditional distributions based on unawareness follow Bayesian updating in response to the arrival of new information. Axioms 1–6 are insufficient to ensure this – we need additional properties as discussed by Ghirardato et al. (2008). These additional properties give rise to a *coherent* multiple priors model.

5 Time, information, awareness and histories

We now consider changes in *information* as well as awareness *over time*, and the induced changes in beliefs and preferences. Time $t = 0, 1, 2, \dots, T$ is discrete and finite.

Information is formally modeled by partitions: \mathcal{F}_t denotes a partition of Ω at time t . As with acts, understanding is assisted by considering the interpretation in terms of propositions. Each

element of a partition corresponds to the truth value of a compound proposition, and the set of such propositions must be exhaustive and mutually exclusive. Information consists of learning that precisely one of these propositions is true.

The element of \mathcal{F}_t that obtains at time t is denoted $f_t \in \mathcal{F}_t$. The collection $\mathcal{F} = \{\mathcal{F}_t\}_{t=0}^T$ constitutes a *filtration*. That is, for each $t = 0, \dots, T-1$, \mathcal{F}_{t+1} is a refinement of \mathcal{F}_t , or equivalently, if $f_{t+1} \in \mathcal{F}_{t+1}$ then $f_{t+1} \subseteq f_t$ for some $f_t \in \mathcal{F}_t$. We write $\rho_t(f)$ for the element of the partition at time t which contains $f \subseteq \Omega$ (provided such an element exists). That is,

$$\rho_t(f) = \{f_t \in \mathcal{F}_t \mid f \subseteq f_t\}$$

In particular, for $f_{t+1} \in \mathcal{F}_{t+1}$, $\rho_t(f_{t+1})$ is the immediate predecessor of f_{t+1} .

We write $\sigma_{t'}(f_t) = \{f_{t'} \in \mathcal{F}_{t'} \mid \rho_{t'}(f_{t'}) = f_t\}$ for the set of successors of f_t at time $t' > t$. In particular, $\sigma_{t+1}(f_t)$ is the set of immediate successors of f_t .

Without loss of generality, we will assume that non-trivial new information arrives only at odd periods, that is, $|\sigma_{t+1}(f_t)| > 1$ only if $t = 2k$ for some $k \in \mathbb{N}_0$. For even periods, $t = 2k + 1$ for some $k \in \mathbb{N}_0$, we have $\mathcal{F}_{t+1} = \mathcal{F}_t$ and no new information is revealed.

We assume that no uncertainty is resolved at date $t = 0$, that is, $\mathcal{F}_0 = \{\Omega\}$ and all uncertainty is resolved by date T , so that for each $\omega \in \Omega$, $\{\omega\} \in \mathcal{F}_T$.¹⁰

The explanation above shows that information must be measurable with respect to the awareness of the agent. An agent cannot learn the truth of a proposition of which she is unaware. It is possible to become aware of a proposition because its truth becomes evident. Next, we introduce the joint dynamics of information and awareness.

To do so, we associate with each pair (t, f) such that $f \in \mathcal{F}_t$, an awareness level $A(t, f)$. We impose the following restrictions on the awareness structure defined by $A(t, f)$:

Definition 4 An awareness structure $\{A(t, \cdot) : \mathcal{F}_t \rightarrow 2^P \setminus \{\emptyset\}\}_{t=0}^T$ is admissible relative to an information partition \mathcal{F} if:

- (i) Awareness is non-empty, increasing and can only change in even periods: $\emptyset \neq A(t, f) \subseteq A(t+1, \sigma_{t+1}(f))$ for all t and $A(t+1, \sigma_{t+1}(f)) = A(t, f)$ if $t = 2k$ for some $k \in \mathbb{N}_0$.
- (ii) Information is measurable with respect to awareness: for an even period $t = 2k$ for some $k \in \mathbb{N}_0$ and a given $f \in \mathcal{F}_t$, each $f_{t+1} \in \sigma_{t+1}(f)$ is measurable with respect to $\Sigma^{A(t, f)}$ as defined in (7).

We will assume that at time T , the decision maker is fully aware: $A(T, f) = P$ for all $f \in \mathcal{F}_T$.

¹⁰Alternatively, we may assume that logical contradictions such as ‘the weather is sunny’ AND ‘the weather is rainy’ are ruled out at $t = 0$, and that information received after $t = 0$ relates only to conceivable states, which may therefore be assumed to have non-zero probability.

Remark 3 The fact that for even $t = 2k$, $\sigma_{t+1}(f_t)$ is measurable with respect to the awareness level at t , $A(t, f_t) = A$ implies that f_t itself is measurable with respect to Σ^A . We write f_t^A for the projection of f_t to A , whenever measurability is satisfied.

Definition 5 An information structure $\{\mathcal{F}_t\}_{t=0}^T$ and an admissible awareness structure $\{A(t, \cdot) : \mathcal{F}_t \rightarrow 2^P \setminus \{\emptyset\}\}_{t=0}^T$ give rise to a set of histories

$$\mathcal{H} = \{h = (t, f, A) \mid t \in \{0, 1, \dots, T\}, f \in \mathcal{F}_t, A = A(t, f)\}.$$

For a given $h = (t, f, A)$, the set of one-step-ahead histories is:

$$\mathcal{H}_{+1}^h = \{h_{+1} = (t+1, f', A') \mid f' \in \sigma_{t+1}(f), A' = A(t+1, f')\}$$

Definition 6 The awareness-adapted history h^A corresponding to $h = (t, f, A)$ is given by:

$$h^A = (t, f^A, A)$$

where f^A is the projection of f on Σ^A . For $t = 2k$, $k \in \mathbb{N}$, and $h = (t, f, A)$, the set of one-step-ahead awareness-adapted histories is:

$$\mathcal{H}_{+1}^{h,A} = \{h_{+1}^A \mid h_{+1} \in \mathcal{H}_{+1}^h\}.$$

For $t = 2k$, and a history $h = (t, f, A)$, we recall that both f and every $f' \in \sigma_{t+1}(f)$ are measurable with respect to Σ^A . Thus, for every such h and the corresponding awareness-adapted history h^A , the set of one-step-ahead awareness-adapted histories, i.e., histories expressible in terms of propositions in A is well-defined.

Remark 4 For a given information structure \mathcal{F} and an admissible awareness structure $\{A(t, \cdot) : \mathcal{F}_t \rightarrow 2^P \setminus \{\emptyset\}\}_{t=0}^T$, consider two consecutive periods $\hat{t} = 2k$, and $\hat{t} + 1 = 2k + 1$. Setting $\hat{\mathcal{F}}_0 = \{f\}$, $\hat{\mathcal{F}}_1 = \sigma_{\hat{t}+1}(f)$, and the corresponding awareness structure to $\hat{A}(0, f) = A(t, f) = A(t+1, f') = A$ for all $f' \in \sigma_{\hat{t}+1}(f)$ results in an information structure together with an awareness structure admissible with respect to $\hat{\mathcal{F}}$. For $h_0 = (0, f, A)$, the corresponding set of one-step-ahead awareness-adapted histories is given exactly by:

$$\mathcal{H}_1^{h_0,A} = \{h_1^A = (1, f', A) \mid f' \in \sigma_{\hat{t}+1}(f)\}$$

Note that up to renumbering of the periods, these histories correspond to the one-step-ahead awareness adapted histories starting at $h = (t, f, A)$ under the original partition structure \mathcal{F} and awareness structure $A(t, f)$:

$$\mathcal{H}_{+1}^{h,A} = \{h_{+1}^A \mid h_{+1} = (t+1, f', A), f' \in \sigma_{\hat{t}+1}(f)\}.$$

5.1 Conditional Preferences over Acts

For a given history $h = (t, f, A)$, the acts of which the decision-maker is aware at h are those measurable with respect to her awareness level A , that is,

$$\mathcal{A}^A = \{\alpha : S^A \rightarrow [0, 1]\}.$$

As usual, the information f available to the decision maker at history h allows us to define conditional preferences.

In particular, for $t = 2k$ and $h = (t, f, A)$, let $\Sigma^{\mathcal{H}_{+1}^{h,A}}$ be the σ -algebra (the set of all subsets) of all one-step-ahead awareness adapted histories. Preferences on \mathcal{A}^A conditional on an event $H \in \Sigma^{\mathcal{H}_{+1}^{h,A}}$ are written \succsim_H^A . Such conditional preferences reflect the fact that for a given awareness level, the arrival of information that the true state is in H will in general change the beliefs of the decision maker and thus, the evaluation of each of the available acts.

In contrast, for $t = 2k + 1$ and $h = (t, f, A)$, the direct successor of h is $h' = (t + 1, f, A')$, where $A \subseteq A'$ with strict inclusion being the non-trivial case. The transition between h and h' is a change in awareness level, the information f remaining unchanged. Thus, an agent with conditional preferences \succsim_h^A on $\mathcal{A}^A = \{\alpha : S^A \rightarrow [0, 1]\}$ will adjust her preferences to the new awareness level, that is, to $\succsim_h^{A'}$ on $\mathcal{A}^{A'} = \{\alpha : S^{A'} \rightarrow [0, 1]\}$.

These two adjustment processes will obey different principles. The former will incorporate new information in a deductive, Bayesian way, leaving existing ambiguity unchanged. The latter will incorporate newly learned states into the model by expanding the state-space and simultaneously reducing the level of perceived ambiguity, that is, the set of priors.

We now describe our desideratum, the coherent multiple prior model, which explains how these principles can be combined in a coherent way.

5.2 The coherent multiple prior model

The following definition of the coherent multiple prior model combines the intuition presented so far. For a probability distribution $\pi \in \Delta(\Omega)$, denote by π^h the Bayesian update of π conditional on history h .

Definition 7 *For a given information structure $\{\mathcal{F}_t\}_{t=0}^T$ and an admissible awareness structure $\{A(t, \cdot) : \mathcal{F}_t \rightarrow 2^P \setminus \{\emptyset\}\}_{t=0}^T$, let \mathcal{H} be the set of corresponding histories. The beliefs of an agent satisfy the coherent multiple prior model (CMP) if there exists a probability distribution $\pi^* \in \Delta(\Omega)$ and for each $h = (t, f, A) \in \mathcal{H}$ a set of priors $\Pi^h \subseteq \Delta(S^A)$ s.t.*

(i) for $t = 0$, and $h_0 = (0, \Omega, A_0 = A(0, \Omega))$,

$$\Pi^{h_0} = \overline{CH} \{ \pi^* (\cdot | \bar{s}^{A_0}) | \bar{s}_{A_0} \in \bar{S}^{A_0} \}$$

(ii) for any $t = 2k$, $k \in \mathbb{N}$, and any $h_{t+1} = (t+1, f', A) \in \mathcal{H}_{t+1}^h$, with $f' \in \sigma_{t+1}(f)$

$$\Pi^{h_{t+1}} = \left\{ \pi^{h_{t+1}^A} \mid \pi \in \Pi^h \right\}$$

(iii) for any $t = 2k+1$, $k \in \mathbb{N}_0$ and the direct successor of h , $h' = (t+1, f, A') \in \mathcal{H}$,

$$\Pi^{h'} = \overline{CH} \left\{ \pi^{*h} \left(\cdot \mid \bar{s}^{A'} \right) \mid \bar{s}_{A'} \in \bar{S}^{A'} \right\}.$$

Part (i) states that for the trivial history, h_0 , the model mimics the static awareness-based multiple prior model presented above. That is, there exists a probability distribution π^* of the fully aware agent, which, when applied to the situation with partial awareness A_0 results in multiple priors $\pi^* \left(\cdot \mid \bar{s}^{A_0} \right)$, one for each of the states of which the agent with awareness level A_0 is unaware.

Part (ii) is an analogue of the generalized Bayesian updating rule for multiple priors introduced by Ghirardato et al. (2008) and is applied to those periods, in which new information arrives.

Finally, Part (iii) is the inductive extension of the static awareness-based multiple prior model: whenever awareness increases, the set of priors is redefined on a larger set of states and the number of priors is reduced to the number of those states $\bar{S}^{A'}$, of which the agent is still unaware. In particular, this part of the definition implies that as long as information remains unchanged, the sequence in which awareness is updated does not matter for the final set of priors.

6 Axiomatizing the coherent multiple prior model

In this section, we provide axiomatic foundations for the coherent multiple prior model. We already provided a condition, Axiom 6, that relates beliefs under full awareness to beliefs under partial awareness and thus, establishes property (i) of the coherent multiple prior model. This axiom will now be extended to hold at all conceivable histories, which will allow us to establish property (iii). In contrast, to obtain generalized Bayesian updating upon arrival of new information, we will make use of the axiomatization of this updating rule in Ghirardato et al. (2008).

6.1 Updating with constant awareness

In Definition 2, we defined conditional preferences for the case of full awareness. Here, we first extend this definition to arbitrary levels of awareness and provide a characterization of multiple prior updating with partial, but constant awareness. In a second step, we incorporate changes in awareness and characterize beliefs satisfying the coherent multiple prior representation.

For the remainder of the subsection, we will focus on the process of belief updating between two consecutive periods $t = 2k$, and $t+1 = 2k+1$. As explained in Remark 4, for a given $h = (t, f, A)$, setting $\hat{\mathcal{F}}_0 = \{f\}$, $\hat{\mathcal{F}}_1 = \sigma_{t+1}(f)$, and the corresponding awareness structure to $\hat{A}(0, f) = A(t, f) =$

$A(t+1, f') = A$ for all $f' \in \sigma_{t+1}(f)$ results in an information structure together with an admissible awareness structure. For $h_0 = (0, f, A)$, the corresponding set of awareness-adapted histories is given exactly by:

$$\mathcal{H}_{+1}^{h,A} = \{h_1^A \mid h_1 = (1, f', A), f' \in \sigma_{t+1}(f)\}$$

Let $\Sigma^{\mathcal{H}_{+1}^{h,A}}$ denote the σ -algebra on such histories and define for each $H \in \Sigma^{\mathcal{H}_{+1}^{h,A}}$ the conditional preferences \succ_H^A on \mathcal{A}^A . Recall that the unconditional preferences are given by \succ_h^A .

For a fully aware agent, conditional preferences are defined on the one-step-ahead histories

$$\mathcal{H}_{+1}^h = \{h_1 = (1, f', A = P) \mid f' \in \sigma_{t+1}(f)\}$$

In fact, for a fully aware agent, awareness can no longer change and thus, only the information partition matters. Thus a history $h = (1, f, A = P)$ can be uniquely identified by f . The relevant events are given by the algebra generated by $\hat{\mathcal{F}}_1, \Sigma^{\hat{\mathcal{F}}_1}$. The conditional preferences of a fully aware agent are defined using Definition 2 for each $E \in \Sigma^{\hat{\mathcal{F}}_1}$.

We are, in a first step, interested in the change in beliefs of an agent who observes and reacts to changes in information while keeping awareness constant at A . In order to ensure that such beliefs are well-defined we require in the following all information events to be non-null, regardless of the relevant awareness level. For an act $\alpha \in \mathcal{A}^A$, let $\alpha^{-1}(1)$ (respectively, $\alpha^{-1}(0)$) denote the event in S^A , on which α pays 1 (respectively 0).

Axiom 7 *Non-null (information-)events:* Let $h = (t, f, A)$ and the corresponding $\hat{\mathcal{F}}_1$ and $\mathcal{H}_{+1}^{h,A}$ be defined as above:

- (a) For any $H \in \Sigma^{\mathcal{H}_{+1}^{h,A}}$, for the bets $\alpha, \alpha_0 \in \mathcal{A}^A \cap \mathcal{B}$ defined as $\alpha^{-1}(1) = H$ and $\alpha_0^{-1}(0) = H$, we have $\alpha \succ_H^A \alpha_0$.
- (b) For any $E \in \Sigma^{\hat{\mathcal{F}}_1}$, for the bets α and α_0 defined as $\alpha^{-1}(1) = E$ and $\alpha_0^{-1}(0) = E$ we have $\alpha \succ_E \alpha_0$.

The first part of the axiom ensures that an agent with awareness A at a given history h considers every information event H corresponding to an element of the information partition $\hat{\mathcal{F}}_1$ adapted to her level of awareness, A as non-null. The second part ensures the same property for the agent who observes the same history h , but is aware of all possible propositions. In particular, since $\mathcal{F}_T = \{\{\omega\}_{\omega \in \Omega}\}$, applying Axiom 7(b) to histories at $T-1$ ensures that each state ω is considered non-null by the fully aware agent.

First consider the case of a fully-aware individual. As usual, Axiom 3(b) implies Bayesian updating at every possible history $h = (t, f, P)$: a fully aware individual, $A = P$, with prior probabilities π^* replaces them by the conditional at history h , $\pi^{*h} = \pi^*(\cdot \mid f)$ upon learning f . Under Axiom 7(b), conditional preferences are not-trivial and thus, every history and thus, every f is not null with respect to the preferences \succ . It follows that the Bayesian updating process is always well-defined.

Extending this reasoning to all possible subsets of histories, E measurable with respect to the information partition \mathcal{F} , Axioms 1–3(b), 4, 5 and 7(b) imply:

Lemma 3 *The class of fully-aware preferences $(\succsim_E)_{E \in \Sigma^{\mathcal{F}^t}, t \in \{0 \dots T\}}$ has an expected utility representation with a prior given by π^* and a system of conditional beliefs $(\pi^{*E})_E$ which satisfy Bayesian updating, i.e., $\pi^{*E} = \pi^*(\cdot | E)$ for every $E \in \Sigma^{\mathcal{F}^t}, t \in \{0 \dots T\}$.*

In particular, the so-defined system of beliefs uniquely identifies, for each history $h = (t, f, A)$, beliefs conditional on events of the type¹¹ $(h, \bar{s}^A), \pi^{*h}(\cdot | \bar{s}^A)$.

We next turn to an agent who at history $h = (f, t, A)$ has awareness level A . We assume that for each $H \in \Sigma^{\mathcal{H}_{+1}^{h,A}}$, conditional preferences \succsim_H^A over awareness adapted acts \mathcal{A}^A satisfy the same set of axioms, 1–3(a), 4, 5, as the unconditional preferences \succsim^A . Adding 7(a), implies that for each H, \succsim_H^A is non-degenerate. Unambiguous preferences \succsim_{*H}^A are defined in analogy to \succsim_*^A . Thus, since A remains constant, for any H , we can associate with the preference \succsim_{*H}^A a set of probability distributions Π^H on S^A such that for $\alpha, \alpha' \in \mathcal{A}^A$,

$$\alpha \succsim_{*H}^A \alpha' \text{ iff } \sum_{s^A \in S^A} \pi(s^A) \alpha(s^A) \geq \sum_{s^A \in S^A} \pi(s^A) \alpha'(s^A) \text{ for all } \pi \in \Pi^H.$$

The next two axioms suggested by Ghirardato et al. (2008) impose the desired structure on Π^H across all possible measurable events $H \in \Sigma^{\mathcal{H}_{+1}^{h,A}}$.

Axiom 8 *Consequentialism: For a given $h = (t, f, A)$ and $H \in \Sigma^{\mathcal{H}_{+1}^{h,A}}$, if α and $\alpha' \in \mathcal{A}^A$ such that $\alpha(\tilde{h}) = \alpha'(\tilde{h})$ for all $\tilde{h} \in H, \alpha \sim_H^A \alpha'$.*

Axiom 9 *Dynamic consistency of \succsim_{*H}^A : For a given $h = (t, f, A)$ and $H \in \Sigma^{\mathcal{H}_{+1}^{h,A}}$, and any acts α and $\alpha' \in \mathcal{A}^A$ such that $\alpha(\tilde{h}) = \alpha'(\tilde{h})$ for all $\tilde{h} \notin H, \alpha \succsim_{*h}^A \alpha'$ iff $\alpha \succsim_{*H}^A \alpha'$.*

Intuitively, Axiom 8, Consequentialism, requires that acts which have identical payoffs on any element of a set H are considered indifferent given H . Notably, this axiom is weaker than the corresponding version in Ghirardato et al. (2008), since it is only imposed on acts of which the agent is aware, i.e., on acts measurable with respect to Σ^A . Axiom 9 requires unambiguous preferences to be dynamically consistent for a given level of awareness A . No such restriction is imposed across histories with different levels of awareness.

The following corollary follows from the main result in Ghirardato et al. (2008):

¹¹Eliciting such beliefs can be done by offering the agent at history h , an act α paying 1 in case \bar{s}^A occurs and nothing otherwise ($\alpha(s) = 1$ if $s \in S^A \times \bar{s}^A, \alpha(s) = 0$, else) and asking him to identify the constant lottery $\alpha' \in \mathcal{C}$ ($\alpha'(s) = x$ for all $s \in \Omega$) he considers indifferent to α conditional on $h, \alpha \sim_h \alpha'$.

Corollary 4 *If for a given $t = 2k$, and a corresponding history $h = (t, f, A)$ the family of preferences \succsim_h^A and $(\succsim_H^A)_{H \in \Sigma^{\mathcal{H}_{t+1}^{h,A}}}$ satisfies Axioms 1–3(a), 4, 5, 7(a), 8 and 9, then there exist sets of probability distributions, Π^h and $(\Pi^H)_{H \in \Sigma^{\mathcal{H}_{t+1}^{h,A}}}$ such that*

(i) *unambiguous preferences at h , $\succsim_{*h}^A = \succsim_{*\mathcal{H}_{t+1}^{h,A}}^A$ are represented by:*

$$\alpha \succsim_{*h}^A \alpha' \text{ iff } \sum_{s^A \in S^A} \pi(s^A) \alpha(s^A) \geq \sum_{s^A \in S^A} \pi(s^A) \alpha'(s^A) \text{ for all } \pi \in \Pi^h;$$

(ii) *unambiguous preferences conditional on $H \in \Sigma^{\mathcal{H}_{t+1}^{h,A}}$ are represented by:*

$$\alpha \succsim_{*H}^A \alpha' \text{ iff } \sum_{s^A \in S^A} \pi(s^A) \alpha(s^A) \geq \sum_{s^A \in S^A} \pi(s^A) \alpha'(s^A) \text{ for all } \pi \in \Pi^H;$$

(iii) *for every H the set of posteriors Π^H is given by the generalized Bayesian updating of Π^h :*

$$\Pi^H = \{ \pi^H \mid \pi \in \Pi^h \},$$

where π^H is the Bayesian update of π conditional on H .

Proposition 5 *For any given even period $t = 2k$, and a corresponding history $h = (t; f; A)$, the coherent multiple prior model satisfies consequentialism. Furthermore, if $\pi^h \in \text{int}(\Delta(S^A))$ for all $\pi^h \in \Pi^h$, the coherent prior model at h satisfies dynamic consistency.*

6.2 Changes in awareness

We now consider pure changes in awareness. For $t = 2k + 1$, consider a history $h = (t, f, A)$. To simplify the exposition, we first consider the case in which the individual becomes aware of a single new proposition p . Suppose the individual's beliefs at h can be represented by a set of coherent probability measures $\Pi^h \subseteq \Delta(S^A)$ generated by the prior probability $\pi^* \in \Delta(\Omega)$ and the information f available at h . Now suppose the individual becomes aware of a proposition $p \in \bar{A}$ at $h' = (t + 1, f, A')$ so that $A' = A \cup \{p\}$.¹² Consider any $s^A \in S^A$ corresponding to a set of truth values for all propositions $p^A \in P^A$, and giving rise to a compound proposition \mathbf{p}^A .

For each such \mathbf{p}^A , the individual at h' considers two possible compound propositions $\mathbf{p}^A \wedge p$, $\mathbf{p}^A \wedge \neg p$ corresponding to the truth or falsity of p . Noting that $s^A \in \{0, 1\}^{|A|}$ is a binary number, we may define the states $(s^A, 1)$ (for p true) and $(s^A, 0)$ (for p false) in $S^{A'}$. Similarly, any $\bar{s}^{A'} \in \bar{S}^{A'}$ corresponds to two complementary states $(\bar{s}^{A'}, 1)$ (for p true) and $(\bar{s}^{A'}, 0)$ (for p false) in \bar{S}^A .

We wish to compare the priors of the agent at h' with those she would have held at h if she were already aware of p . As we showed in our example in Section 2, under coherence these will coincide.

¹²Given the alternating dates setup the individual at h_{+1} does not learn whether p is true, although this may be resolved by subsequent revelation of information.

Proposition 6 Consider an information structure $\{\mathcal{F}_t\}_{t=0}^T$ and an admissible awareness structure $\{A(t, \cdot) : \mathcal{F}_t \rightarrow 2^P \setminus \{\emptyset\}\}_{t=0}^T$ such that at history $h = (t, f, A)$ for $t = 2k+1$ the change in awareness involves a single elementary proposition p : that is, $h_{+1} = (t+1, f, A' = A \cup \{p\})$ for some $p \notin A$. The coherent multiple prior model implies

$$\begin{aligned} \Pi^h &= \overline{CH} \left\{ \pi^{*h} \left(\cdot \mid \left(\bar{s}^{A'}, 0 \right) \right), \pi^{*h} \left(\cdot \mid \left(\bar{s}^{A'}, 1 \right) \right) \mid \bar{s}^{A'} \in \bar{S}^{A'} \right\} \\ \Pi^{h_{+1}} &= \overline{CH} \left\{ \pi^{*h} \left(\cdot \mid \bar{s}^{A'} \right) \mid \bar{s}^{A'} \in \bar{S}^{A'} \right\} \end{aligned} \quad (10)$$

where π^{*h} and $\pi^{*h_{+1}}$ represent the conditional beliefs of the fully aware agent and for each of the priors $\pi^{*h} \left(\cdot \mid \bar{s}^{A'} \right)$,

$$\pi^{*h} \left(\cdot \mid \bar{s}^{A'} \right) = \frac{\pi^{*h} \left(\bar{s}^{A'}, 1 \right)}{\pi^{*h_{+1}} \left(\bar{s}^{A'} \right)} \pi^{*h_{+1}} \left(\cdot \mid \left(\bar{s}^{A'}, 1 \right) \right) + \frac{\pi^{*h} \left(\bar{s}^{A'}, 0 \right)}{\pi^{*h_{+1}} \left(\bar{s}^{A'} \right)} \pi^{*h_{+1}} \left(\cdot \mid \left(\bar{s}^{A'}, 0 \right) \right) \quad (11)$$

We next adapt Axiom A6 to the intertemporal setting. Recall that definition 2 specifies the conditional preferences under full awareness for each subset in Σ and hence, allows us to talk about preferences conditional on the realization of any event in Σ in conjunction with any history h .

Axiom 10 *Conditional Unanimity:* Let $h = (t, f, A)$. For $\alpha, \alpha' \in A^A$, $\alpha \succ_{*h}^A \alpha'$ if and only if for every $\bar{s}^A \in \bar{S}^A$ and every $\hat{\alpha}_{\bar{s}^A}, \hat{\alpha}'_{\bar{s}^A} \in \mathcal{A}$ with $\hat{\alpha}_{\bar{s}^A}(s \times \bar{s}^A) = \alpha(s)$ and $\hat{\alpha}'_{\bar{s}^A}(s \times \bar{s}^A) = \alpha'(s)$ for each $s \in S^A$ and $\hat{\alpha}_{\bar{s}^A}(s \times \tilde{s}^A) = \hat{\alpha}'_{\bar{s}^A}(s \times \tilde{s}^A)$ for all $s \in S^A$ and $\tilde{s}^A \in \bar{S}^A \setminus \bar{s}^A$, $\hat{\alpha}_{\bar{s}^A} \succ_h \hat{\alpha}'_{\bar{s}^A}$.

Just as Axiom 6, Axiom 10 can be equivalently expressed in terms of conditional preferences to state that an act α is unambiguously preferred to act α' at history h and awareness level A , $\alpha \succ_{*h}^A \alpha'$, if and only if for every two acts $\hat{\alpha}, \hat{\alpha}' \in \mathcal{A}(A)$ with $\hat{\alpha}(\tilde{s}) = \alpha(s)$ and $\hat{\alpha}'(\tilde{s}) = \alpha'(s)$ for each $\tilde{s} \in E_s$ and each $s \in S^A$, $\hat{\alpha}$ is preferred to $\hat{\alpha}'$ conditional on history h and state \bar{s}^A , $\hat{\alpha} \succ_{\bar{s}^A, h} \hat{\alpha}'$, for all $\bar{s}^A \in \bar{S}^A$. Axiom 10 thus requires that at any history $h = (t, f, A)$ an act α is unambiguously preferred to α' if and only if this ranking is preserved for a fully aware agent, conditional on every possible truth value of the propositions not contained in A .

Proposition 7 Consider an information structure $\{\mathcal{F}_t\}_{t=0}^T$ and an admissible awareness structure $\{A(t, \cdot) : \mathcal{F}_t \rightarrow 2^P \setminus \{\emptyset\}\}_{t=0}^T$ with a corresponding set of histories \mathcal{H} . If

- (i) for each h , (conditional) preferences under full awareness $(\succ_h)_{h \in \mathcal{H}}$ satisfy Axioms 1–3(b), 4, 5 and 7(b);
- (ii) for each $t = 2k$, $h = (t, f, A)$, (conditional) preferences under partial awareness $(\succ_H^A)_{H \in \Sigma^{\mathcal{H}_{+1}^{h, A}}}$ satisfy Axioms 1–3(a), 4, 5, 7(a), 8 and 9;
- (iii) for each $h = (t, f, A)$, \succ_h^A satisfies Axiom 10,

then there exists a prior π^* on (Ω, Σ) and for each $h = (t, f, A) \in \mathcal{H}$ a set of priors Π^h on (S^A, Σ^A) s.t. for any two acts α and $\alpha' \in \mathcal{A}^A$ and the corresponding $\hat{\alpha}, \hat{\alpha}' \in \mathcal{A}(A)$ with $\hat{\alpha}(\tilde{s}) = \alpha(s)$ and $\hat{\alpha}'(\tilde{s}) = \alpha'(s)$ for each $\tilde{s} \in E_s$ and each $s \in S^A$,

$$\alpha \succsim_{*h}^A \alpha' \text{ iff} \tag{12}$$

$$\hat{\alpha} \succsim_{\bar{s}^A, h} \hat{\alpha}' \text{ for all } \bar{s}^A \in \bar{S}^A \text{ iff}$$

$$\sum_{s^A \in S^A} \pi(s^A) \alpha(s^A) \geq \sum_{s^A \in S^A} \pi(s^A) \alpha'(s^A) \text{ for all } \pi \in \Pi^h$$

and for any two acts α and $\alpha' \in \mathcal{A}$, any h and any \bar{s}^A ,

$$\alpha \succsim_h \alpha' \text{ iff } \sum_{s \in \Omega} \pi^{*h}(s) \alpha(s) \geq \sum_{s \in \Omega} \pi^{*h}(s) \alpha'(s) \text{ and} \tag{13}$$

$$\alpha \succsim_{\bar{s}^A, h} \alpha' \text{ iff } \sum_{s \in \Omega} \pi^{*h}(s | \bar{s}^A) \alpha(s) \geq \sum_{s \in \Omega} \pi^{*h}(s | \bar{s}^A) \alpha'(s)$$

Furthermore, π^* , together with the family of multiple priors $(\Pi^h)_{h \in \mathcal{H}}$ is a coherent multiple prior model.

Proposition 7 shows that under Axiom 10, Conditional Unanimity, the set of priors the decision-maker entertains at h , before becoming aware of a certain set of propositions is given exactly by the convex hull of the posteriors conditional on the truth values of these propositions once she has become aware of them at h_{+1} . Using the fact that under Axioms 1–5 and 7–9, conditional beliefs are formed via generalized Bayesian updating, we obtain the coherent multiple prior representation of beliefs.

Finally, since for each history $h = (t, f, A)$, the extreme points of the set Π^h are derived from conditional probabilities obtained from the (full awareness) prior $\pi^* \in \Delta(\Omega)$, conditioning on h and in turn on each \bar{s}^A in \bar{S}^A , it follows that $\pi^{*h} \in \Pi^h$, since from the iterative law of expectations, we have:

$$\pi^{*h} = \sum_{\bar{s}^A \in \bar{S}^A} \pi^{*h}(\bar{s}^A) \pi^{*h}(\cdot | \bar{s}^A).$$

More generally, for any history $h = (t, f, A)$, with t odd, and with an immediate successor $h_{+1} = (t+1, f, A')$ embodying a pure increase in awareness (that is, $A \subset A'$), we have from the construction that $\Pi^{h_{+1}} \subset \Pi^h$. That is, the mapping from Π^h to $\Pi^{h_{+1}}$ may be viewed as a contraction with π^{*h} as a fixed point.

7 An application: Latent variables and mixture models

The interpretation of multiple priors models offered here may be considered in relation to mixture and latent variables used in a variety of statistical settings. Consider a model in which the dataset

consists of M observations, with each observation consisting of a variable of interest y , taking the values 0 and 1, and a vector of explanatory variables X .

Begin by considering a standard binary response model

$$y = \begin{cases} 1 & \text{if } f(X, \beta, \varepsilon) > 0 \\ 0 & \text{otherwise} \end{cases}$$

where β is a vector of parameters and ε is a random variable determined by the state of nature and characterized by parameters θ .

The parameters β and θ may be estimated by maximizing the log-likelihood function

$$\ell(\beta, \theta) = \sum_m \log p(y_m | X_m, \beta, \theta)$$

In particular, if β is known, this estimation procedure amounts to Bayesian updating with respect to θ .

To relate this model to the decision theory problems discussed suppose that the variables in X may be partitioned into a vector of exogenous variables X_1 and a vector of control variables X_2 . As before, utility may be given as $u(y) = y$. The agent seeks to choose X_2 to maximize expected utility $p(y = 1)$, given the observable value of X_1 and the estimated parameters (β, θ) .

A latent variable model associates with each m a vector of K unobserved latent variables $(u_{km})_{k=1}^K$, commonly taken to be discrete. We will focus on the case where each u_{km} is a dummy variable taking the values 0 and 1. Hence, each latent variable may be interpreted as the truth value of a proposition. Consider for simplicity the case $K = 1$. In this case, the distribution of ε is determined by parameters θ_0 if $u_{km} = 0$, and θ_1 if $u_{km} = 1$.

One way to approach this problem would be to formulate two separate models, one assuming $u_k = 0$, and the other assuming $u_{km} = 1$. Now, we have two possible likelihoods, $p_0(y_m | X_m, \beta) = p(y_m | X_m, \beta, \theta_0)$ and $p_1(y_m | X_m, \beta) = p(y_m | X_m, \beta, \theta_1)$. This corresponds to the simplest case of multiple priors discussed above. Hence, expected utility lies in the interval $[p_0, p_1]$ and the choice of X_2 might be given by a criterion such as maxmin. In terms of the discussion above, this approach corresponds to the case where the agent is unaware of the latent variable u .

Now consider the case where the agent is aware of u but cannot observe its value, instead assigning a probability $p_m = p(u_m = 1 | X_{1m})$. Hence, the log-likelihood is given by a weighted sum

$$\ell(\beta, \theta_0, \theta_1) = \sum_m \log p(y_m | X_m, \beta, \theta_0, \theta_1)$$

where

$$\begin{aligned} p(y_m | X_m, \beta, \theta_0, \theta_1) &= p(y_m | X_m, \beta, \theta_1) (1 - p(u_m = 1 | X_{1m})) \\ &\quad + p(y_m | X_m, \beta, \theta_0) p(u_m = 1 | X_{1m}) \end{aligned}$$

This corresponds to a latent variable model, more precisely a mixture model. A standard iterative algorithm for estimating the parameters of a model of this kind is expectation-maximization or EM, due to Dempster et al. (1977). The algorithm proceeds with alternating steps: an ‘expectation’ step in which estimates of $p_m = p(u_m = 1 \mid X_{1m})$ are updated and a maximization step in which the log-likelihood $\ell(\beta, \theta_0, \theta_1)$ is maximized with respect to the parameter vector $(\beta, \theta_0, \theta_1)$.

There is an obvious analogy with the dynamic process set out above, where the alternating steps involve changes in awareness of previously unconsidered propositions (that is, latent variables) and updating in response to new information. We conjecture that it might be possible to represent a boundedly rational individual as following a real-time dynamic version of the Dempster et al. (1977) algorithm in which the expectation and maximization steps incorporate changes in awareness and information respectively.

Finally, note that in the case when the agent is aware of u and can observe its value, we have a standard dummy variable.

8 Related literature

Our paper is related to the growing literature on unawareness, see Schipper (2014) for an introduction. Two main approaches have arisen in this literature: the first relies on explicitly modelling the knowledge and the awareness of the decision maker at each possible state of the world. The second, consists in axiomatizing choice behavior: either for a given awareness structure which is exogenously specified, or for a subjective awareness structure, which can then be deduced from preferences.

Fagin and Halpern (1988) were the first to introduce awareness structures into a model of knowledge. Their approach was then followed by that of Heifetz et al. (2006). The main characteristic of this approach is the exogenous specification of awareness at each state of the world. Combined with the information structure, which at each state specifies the event the agent is informed about, this gives rise to two knowledge operators: implicit knowledge – that captured by the information partition and explicit knowledge, which necessitates the agent to simultaneously implicitly know the event and be aware of it.¹³

A simple awareness structure, which captures this distinction is provided by Li (2009). In the present paper, we adopt her framework by modeling unawareness of propositions (“questions” in her framework). At each state, the model exogenously defines the subset of propositions the agent is aware of, as well as the event known to have occurred. Li (2009) shows that such structures can be used to generate meaningful “knowledge” and “unawareness” operators. Our structure is simpler

¹³Less expressive structures, in which the exact specification of awareness at each state are replaced by less informative propositions such as: “the agent is (is not) fully aware at this state” have been introduced by Agotnes and Alechina (2008) and further developed by Walker (2014).

in that we require information obtained at each state (node) to be measurable with respect to the agent’s awareness at this node, which need not be the case in Li (2009). This simplification proves convenient for the analysis and the resulting model captures well the phenomena we have in mind. A generalization to information structures non-measurable with respect to the awareness level would involve techniques similar to those used by Li (2009), whereby the information actually revealed to the agent is the finest coarsening measurable with respect to her current awareness level. This complicates the notation without providing meaningful insights into our main results.

The relation between the model of Li (2009) and the awareness structures of Fagin and Halpern (1988) and unawareness structures by Heifetz et al. (2006) are discussed in Schipper (2014, p. 4). In as far as these models rely on implicit knowledge, Schipper (2014, p. 3) suggests that such awareness structures can be viewed as capturing “features of logical non-omniscience”.

The class of axiomatic models of unawareness can be roughly divided into two: in the first category fall the models, which take the awareness structure as exogenously given and model preferences related to such structures. This comprises the work by Ahn and Ergin (2010), Grant and Quiggin (2013, 2015), Karni and Viero (2013, 2017), Viero (2018), Lehrer and Teper (2014), Alon (2015), Dominiak and Tserenjigmid (2018), Dietrich (2018). These papers study conditions, which relate preferences across different levels of awareness.

In a seminal paper, Karni and Viero (2013) study growing awareness as a consequence of discovering new acts or new consequences.¹⁴ This results in a refinement or in an expansion of the original state space. Karni and Viero (2013) impose a restriction on beliefs – reverse Bayesianism, which implies that relative likelihoods of events considered at lower levels of awareness are preserved when awareness increases.

Dietrich (2018) also considers partial awareness in terms of both coarsening and reduction of both the state-space and the outcome space. He models a subjective expected utility maximizer, who satisfies Savage axioms at all levels of awareness. His axiomatization imposes stringent constraints on utility and beliefs across awareness structures: both utility and probability at lower awareness levels are intimately related to those utilities and probabilities the agent would entertain were she fully aware.¹⁵ While we study ambiguity as arising from partial awareness, our approach is similar

¹⁴In this model, we are primarily interested in predictions (i.e., in agent’s beliefs). Hence, we keep the binary set of outcomes $\{0, 1\}$ constant. Acts can be thought as bets on the occurrence of an event. Thus, the framework of Li (2009), in which the agent becomes aware of finer contingencies through new propositions (which in turn expands the set of possible predictions / available acts) is more appropriate for our purposes than that of Karni and Viero (2013), where new states have to be constructed either from new acts or new outcomes.

¹⁵Dietrich (2018) views such consistency requirements as normative. “One may legitimately question the plausibility of such a hybrid agent: why should someone who can come up with objective evaluations fail to come up with objective outcomes, states and acts in the first place? The point of defining classical EU rationalizations is not to defend ‘objective evaluations of subjective objects’ as genuinely realistic, but to spell out the classical benchmark from which our less classical rationalizations depart.”

both to those of Karni and Viero (2013) and Dietrich (2018) in that beliefs under full awareness uniquely pin down beliefs at lower levels of awareness.¹⁶ The normative aspect of such constraints as captured by the axioms we propose is of interest in problems of inference, such as the latent variable model discussed in Section 7.

While the approach taken by Karni and Viero (2013) and Dietrich (2018) is purely Bayesian, Grant and Quiggin (2015), Karni and Viero (2017) and Viero (2018) introduce the notion of unknown contingencies and associate with them a decision weight and a utility. In a similar spirit, Alon (2015) models awareness of unforeseen contingencies by introducing an “unforeseen event” with an associated weight and an associated “worst-case-contingency”. While in these models, such decision weights and utilities are purely subjective, our coherent multiple prior model imposes restrictions on the way in which a partially aware agent incorporates the possibility of becoming more aware into her decisions. In particular, partial awareness leads to ambiguity, and thus, multiple priors reflecting all possible ways in which the realization of the yet undiscovered states might affect the probabilities of the currently known ones.

The papers by Lehrer and Teper (2014) and Dominiak and Tserenjigmid (2018) advance the idea that the agent’s confidence / perception of ambiguity may depend on her awareness. Both papers consider two preference relations, one corresponding to a larger state-space (higher level of awareness), the other to a coarsening of it. Both papers argue that the agent behaving as an EUM when partially aware might face more ambiguity / be less confident with a larger state-space. This is captured by incompleteness of preferences on the larger state-space in Lehrer and Teper (2014) and by perceived ambiguity on the expanded state-space in Dominiak and Tserenjigmid (2018). In a similar spirit, Grant et al. (2019) characterize a family of beliefs conditional on realizations of a stochastic process such that: (i) as new states are observed, there is maximal ambiguity with respect to the probability of these new states; (ii) as information about already known states accumulates, the agent updates her priors in a Bayesian way, eventually learning the correct probabilities of each state. In the limit, the expectations taken with respect to the posterior of each of the initial priors are arbitrarily close.

Similarly to these works, we establish a relation between preferences and beliefs with “varying degrees of awareness” and impose consistency requirements across such beliefs. Our point of departure, however is that ambiguity arises due to partial awareness and disappears as the agent becomes fully aware. We thus propose a deductive (forward-looking) rather than an inductive approach in that beliefs at lower levels of awareness already incorporate all possible unforeseen contingencies as multiple priors. Furthermore, we concentrate on the characterization of ambiguous beliefs (multiple priors) for partial awareness and remain agnostic as to how the agent deals with such ambiguity. This

¹⁶Karni and Viero (2015) allow for non-additivity by modelling unawareness with probabilistically sophisticated beliefs. While this captures Allais-style violations of additivity, it does not incorporate ambiguity.

allows for a simpler framework with a binary outcome space and fewer restrictions on preferences.

Finally, Ahn and Ergin (2010)'s model of framing effects can be interpreted as a model of preferences for different levels of awareness. Similarly, to the models in the preceding paragraph, their premise is that the agent has non-additive beliefs on the complete state-space, but forms additive beliefs for each subset of this space (each possible frame). These additive beliefs are derived from the non-additive capacity over the complete state-space.

The second class of axiomatic models uses preferences over extended classes of objects (e.g., menus or consumption streams) and proposes behavioral conditions to identify the awareness of the agent. In Epstein et al. (2007) the coarse state space is subjective and derived from preferences over menus, which violate indifference to randomization. When choosing over menus, the agent takes into account for each such coarse state, the worst payoff realization, thus exhibiting complete ignorance about the probability of the individual utility realizations.

Preferences on richer domains can provide a distinction between ambiguity and unawareness in an axiomatic framework. In particular, Piermont (2017) relates awareness of unawareness to the agent's unwillingness to commit to any contingent plan even when delaying the decision is costly. Kochov (2017) interprets awareness of unawareness as the agent recognizing her inability to correctly judge the autocorrelation of payoffs over time. In his framework, awareness of unawareness is revealed whenever a payoff stream with state-contingent outcomes to which the agent assigns identical utility is assigned a utility different from that assigned to the outcomes.

In contrast, our paper seeks to deduce ambiguity from unawareness rather than distinguishing the two concepts. In that sense, it is related to the papers by Mukerji (1997), Ghirardato (2001), Epstein et al. (2007) and Billot and Vergopoulos (2018). For example, Mukerji (1997) shows that probability weighting may be derived from a decision-maker's anticipation that her perception of future contingencies is incomplete. Ghirardato (2001) models ambiguity as a consequence of the coarse perception of state-contingent payoffs. Billot and Vergopoulos (2018) show how ambiguity can arise when the observable state space differs from the relevant state space.

The difference between these papers and our model is two-fold: first, we study changes in awareness and in information and impose consistency requirements across preferences and beliefs at different levels of awareness and information for the same agent. To do so, we rely extensively the model of generalized Bayesian updating for multiple prior preferences as developed by Ghirardato et al. (2008).

Second, we constrain the agent to reason only about acts which are measurable with respect to her coarse state space. Thus, ambiguity is not due to the fact that the agent conceives of multiple outcomes related to a given coarse contingency, but to the fact that the probabilities of the contingencies of which the decision maker is aware depend on factors of which she is not. Rather than assuming that such ambiguity leads to full ignorance and to behavior which is guided by the

worst possible outcome, in our context, ambiguity is objectively related to the stochastic process which determines the realization of the factors of which the decision maker is unaware.

Finally, there is a well-known formal result, see Ghirardato and Breton (2000) and Gilboa and Schmeidler (1994), showing that the Choquet integral (the standard tool for modelling ambiguity-sensitive preferences) can be represented in an additive way on an extended state-space (the set of $\{0, 1\}$ -valued capacities). Our construction is different: we do model additive beliefs on a larger state space and non-additive ones on a smaller state-space, but the two spaces are not related as in Ghirardato and Breton (2000) and Gilboa and Schmeidler (1994) and the resulting evaluations of acts in general differ between the two spaces, reflecting the decrease in ambiguity which accompanies growing awareness.

9 Conclusion

Beginning in the late 1970s, alternatives to and generalizations of Expected Utility theory have proliferated in response to behavioral violations of EU predictions and theoretical criticism of the axiomatic foundations of EU. Examples have included probability weighting models for choice under risk (Allais (1953), Kahneman and Tversky (1979), Quiggin (1982), Yaari (1987)), ambiguity models for choice under uncertainty (Gilboa and Schmeidler (1989), Schmeidler (1989), Ghirardato et al. (2004), Klibanoff et al. (2005)) and the rapidly growing literature on unawareness (Schipper (2014)).

We reviewed some of the attempts at unification of the two theories in Section 1. Similarly, in this paper, we have shown that the invariant biseparable model of Ghirardato et al. (2004) model of choice under ambiguity (which incorporates α -maxmin EU as a special cases), may be derived from the preferences of an EU maximizer with coarse awareness. Updating in response to both new information and refined awareness is well-behaved.

This development raises the possibility of a more general unified theory of EU behavior with bounded awareness that might encompass a wide range of observed behavior as well as being consistent with the fundamental postulate that all humans have bounded cognitive capacity.

Such models can prove useful as a decision-theoretical underpinning of statistical methods related to latent variables.

Appendix

Proof of Proposition 2:

Suppose that the family of preferences $(\succsim^A)_A$ satisfies Axioms 1–6. Fix an awareness level A . By Lemma 1, for the so-chosen A , there exists a unique closed and convex set of priors Π^A such that the unambiguous preferences induced by \succsim^A, \succsim_*^A can be represented by (9). Furthermore, by

Axiom 6 and by Definition 2 of conditional preferences, we obtain $\alpha \succsim_*^A \alpha'$ iff for any $\bar{s}^A \in \bar{S}^A$, the corresponding acts $\hat{\alpha}_{\bar{s}^A}$ and $\hat{\alpha}'_{\bar{s}^A}$ defined in statement of Axiom 6, satisfy $\hat{\alpha}_{\bar{s}^A} \succ \hat{\alpha}'_{\bar{s}^A}$, or, equivalently:

$$\sum_{s \in S} \pi^*(s) \hat{\alpha}_{\bar{s}^A}(s) \geq \sum_{s \in S} \pi^*(s) \hat{\alpha}'_{\bar{s}^A}(s). \quad (14)$$

Furthermore, since $\hat{\alpha}_{\bar{s}^A}(s) = \hat{\alpha}'_{\bar{s}^A}(s)$ for every $s \in S^A \times \bar{S}^A \setminus \bar{s}^A$, we have that (14) is equivalent to:

$$\sum_{s^A \in S^A} \pi^*(s^A | \bar{s}^A) \hat{\alpha}_{\bar{s}^A}(s^A) \geq \sum_{s^A \in S^A} \pi^*(s^A | \bar{s}^A) \hat{\alpha}'_{\bar{s}^A}(s^A)$$

and since $\hat{\alpha}_{\bar{s}^A}(s) = \alpha(s)$ and $\hat{\alpha}'_{\bar{s}^A}(s) = \alpha'(s)$ for every $s \in S^A \times \bar{s}^A$, we have:

$$\sum_{s^A \in S^A} \pi^*(s^A | \bar{s}^A) \alpha(s^A) \geq \sum_{s^A \in S^A} \pi^*(s^A | \bar{s}^A) \alpha'(s^A). \quad (15)$$

We thus conclude that the requirement that $\hat{\alpha}_{\bar{s}^A} \succ \hat{\alpha}'_{\bar{s}^A}$ for every $\bar{s}^A \in \bar{S}^A$ is satisfied if and only if (15) holds for every $\bar{s}^A \in \bar{S}^A$, or, equivalently, if and only if

$$\sum_{s^A \in S^A} \pi(s^A) \alpha(s^A) \geq \sum_{s^A \in S^A} \pi(s^A) \alpha'(s^A) \text{ for all } \pi \in \hat{\Pi} = \overline{CH} \{ \pi^*(\cdot | \bar{s}^A) | \bar{s}^A \in \bar{S}^A \}. \quad (16)$$

Axiom 6 requires (16) to be equivalent to $\alpha \succsim_*^A \alpha'$. But from Axioms 1–5 and Lemma 1, we have that $\alpha \succsim_*^A \alpha'$ iff

$$\sum_{s^A \in S^A} \pi(s^A) \alpha(s^A) \geq \sum_{s^A \in S^A} \pi(s^A) \alpha'(s^A) \text{ for all } \pi \in \Pi^A,$$

where Π^A is unique. Thus, we obtain $\hat{\Pi} = \Pi^A = \overline{CH} \{ \pi^*(\cdot | \bar{s}^A) | \bar{s}^A \in \bar{S}^A \}$ as required by the definition of awareness-based beliefs. Since the awareness level A was chosen arbitrarily, this holds for every A and the claim of the proposition obtains.

Proof of Proposition 5:

Assume a representation by an CMP model. Fix $h = (t, f, A)$. To see that Consequentialism holds, note that generalized Bayesian updating implies that conditional on $H \in \Sigma^{\mathcal{H}_{t+1}^h, A}$, all $\tilde{h} \notin H$ are assigned 0-probability under all $\pi \in \Pi^H$. Let

$$S_h = \{s^A \in S^A | s^A \in f\}$$

$$S_H = \{s^A \in S^A | s^A \in \sigma_{t+1}(f) \text{ with } (t+1, \sigma_{t+1}(f), A) \in H\}$$

be the set of states in S^A consistent with histories in H . By the definition of α and α' in A8, we have that for all $\pi \in \Pi^H$,

$$\sum_{s^A \in S_H} \pi(s^A) \alpha(s^A) = \sum_{s^A \in S_H} \pi(s^A) \alpha'(s^A)$$

so that $\alpha \sim_{*H}^A \alpha'$, which implies $\alpha \sim_H^A \alpha'$.

As for dynamic consistency of \succsim_H^* , note that for any acts α and α' as defined in A9 and every $\pi \in \Pi^h$

$$\begin{aligned} & \sum_{s^A \in S_h} \pi(s^A | S_h) \alpha(s^A) - \sum_{s^A \in S_h} \pi(s^A | S_h) \alpha'(s^A) \\ &= \pi(S_H) \left[\sum_{s^A \in S_h} \pi(s^A | S_H) \alpha(s^A) - \sum_{s^A \in S_h} \pi(s^A | S_H) \alpha'(s^A) \right]. \end{aligned}$$

$\alpha \succ_{\sim_h^A}^A \alpha'$ holds iff the first difference is positive for all $\pi \in \Pi^h$ and, as long as all such π are in the interior of $\Delta(S^A)$, this is clearly equivalent to the last difference being positive for all $\pi \in \Pi^H$, or to $\alpha \succ_{\sim_H^A}^A \alpha'$.

Proof of Proposition 6:

Notice that

$$\pi^{*h+1}(\bar{s}^{A'}) = \pi^{*h+1}(\bar{s}^{A'}, 1) + \pi^{*h+1}(\bar{s}^{A'}, 0)$$

where π^* is the probability on Ω and its arguments are considered as events. Furthermore, by condition (iii) of the definition of CMP,

$$\Pi^{h+1} = \overline{CH} \left\{ \pi^{*h+1}(\cdot | \bar{s}^{A'}) \mid \bar{s}^{A'} \in \bar{S}^{A'} \right\} = \overline{CH} \left\{ \pi^{*h}(\cdot | \bar{s}^{A'}) \mid \bar{s}^{A'} \in \bar{S}^{A'} \right\}$$

where the second equality follows from the fact that no new information arrives between h and $h+1$, and hence $\pi^{*h} = \pi^{*h+1}$.

For $h_{-1} = (t-1, \rho_{t-1}(f), A)$, we have by property (iii) of CMP,

$$\Pi^{h-1} = \overline{CH} \left\{ \pi^{*h-1}(\cdot | \bar{s}^A) \mid \bar{s}^A \in \bar{S}^A \right\}$$

and hence, by property (ii),

$$\begin{aligned} \Pi^h &= \left\{ \pi(\cdot | h) \mid \pi \in \Pi^{h-1} \right\} = \left\{ \pi^{*h}(\cdot | \bar{s}^A) \mid \bar{s}^A \in \bar{S}^A \right\} \\ &= \overline{CH} \left\{ \pi^{*h}(\cdot | \bar{s}^A) \right\} = \overline{CH} \left\{ \pi^{*h}(\cdot | (\bar{s}^{A'}, 0)), \pi^{*h}(\cdot | (\bar{s}^{A'}, 1)) \mid \bar{s}^{A'} \in \bar{S}^{A'} \right\}. \end{aligned}$$

Hence, Bayesian updating of the beliefs of the fully aware agent π^* , together with $\pi^{*h} = \pi^{*h+1}$ implies:

$$\pi^{*h}(\cdot | \bar{s}^{A'}) = \frac{\pi^{*h}(\bar{s}^{A'}, 1)}{\pi^{*h+1}(\bar{s}^{A'})} \pi^{*h+1}(\cdot | (\bar{s}^{A'}, 1)) + \frac{\pi^{*h}(\bar{s}^{A'}, 0)}{\pi^{*h+1}(\bar{s}^{A'})} \pi^{*h+1}(\cdot | (\bar{s}^{A'}, 0)).$$

Proof of Proposition 7:

The representation of \succsim_h follows directly from the conditions in (ii) and Lemma 3. These conditions uniquely identify π^* and ensure Bayesian updating of fully aware beliefs.

The representation of the conditional $\succ_{\sim_h^A}^A$ preference relations follows from Axioms 1–3(a), 4, 5, 7(a) stated in condition (i) and Lemma 1. These axioms identify the corresponding sets of priors

Π^h . Axiom 10 relates the partially aware preferences conditional on h to the conditional preferences under full awareness and implies condition (12).

Just as Axiom 6 for the static case, Axiom 10 relates Π^{h_0} (the partially aware preference with no information) to the conditionals of $\pi(\cdot | \bar{s}^{A_0})$ and by Lemma 2 implies property (i) of the CMP model.

Property (ii) of the CMP model follows from Axioms 8 and 9, which by Corollary 4, imply generalized Bayesian updating of every prior in Π^h .

Finally, the proof of property (iii) of the CMP model is a consequence of Axiom 10 and is shown exactly as in the proof of Proposition 2. Indeed, fix a history h and in the proof of Proposition 2, replace preferences \succ_*^A by \succ_{*h}^A and \succ by \succ_h , whereas the set of priors Π^A is replaced by Π^h and the full awareness probability distribution π^* is replaced by π^{*h} . Repeating the arguments in the proof of Proposition 2 gives property (12). Property (13) follows from the fact that the fully aware preference at history h satisfies all axioms of expected utility maximization for the probability distribution π^{*h} , Axioms 1–3(b), 4, 5 and by Axiom 7(b) all relevant events are non-null.

References

- Agotnes, T. and N. Alechina (2008). Proceedings of the 11th conference on theoretical aspects of rationality and knowledge (tark). In *Full and Relative Awareness: A decidable logic for reasoning about knowledge and unawareness*, pp. 6–14. Presses Universitaires De Louvain.
- Ahn, D. and H. Ergin (2010). Framing contingencies. *Econometrica* 78, 655–695.
- Allais, M. (1953). Le comportement de l’homme rationnel devant le risque: critique des postulats et axiomes de l’école Américaine. *Econometrica* 21(4), 503–546.
- Alon, S. (2015). Worst-case expected utility. *Journal of Mathematical Economics* 60, 43–48.
- Billot, A. and V. Vergopoulos (2018). Expected utility without parsimony. *Mathematical Social Sciences* 93, 14–21.
- Cowan, N. (2020). Pandemic responses are beyond Evidence-based Medicine. <https://notesonliberty.com/2020/03/23/pandemic-responses-are-beyond-evidence-based-medicine/>.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society, Series B* 39(1), 1–38.
- Dietrich, F. (2018). Savage’s theorem under changing awareness. *Journal of Economic Theory* 176, 1–54.

- Dominiak, A. and G. Tserenjigmid (2018). Ambiguity under growing awareness.
- Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics* 75(4), 643–669.
- Empson, W. (1930). *Seven Types of Ambiguity*. Read Books Limited.
- Epstein, L., M. Marinacci, and K. Seo (2007). Recursive smooth ambiguity preferences. *Theoretical Economics* 2, 355–394.
- Fagin, R. and J. Y. Halpern (1988). Belief, awareness, and limited reasoning. *Artificial Intelligence* 34, 39–76.
- Ghirardato, P. (2001). Coping with ignorance: Unforeseen contingencies and non-additive uncertainty. *Economic Theory* 17, 247–276.
- Ghirardato, P. and M. L. Breton (2000). Choquet rationality. *Journal of Economic Theory* 90, 277–285.
- Ghirardato, P., F. Maccheroni, and M. Marinacci (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory* 118(2), 133–173.
- Ghirardato, P., F. Maccheroni, and M. Marinacci (2008). Revealed ambiguity and its consequences: updating. In *Advances in decision making under risk and uncertainty*, pp. 3–18. Springer.
- Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18(2), 141–153.
- Gilboa, I. and D. Schmeidler (1994). Maxmin expected utility with non-unique prior. *Annals of Operations Research* 52, 43–65.
- Grant, S., J. Kline, and J. Quiggin (2018). Contracting under uncertainty: a principal-agent model with ambiguity averse parties. *Games and Economic Behavior* 109, 582–597.
- Grant, S., I. Meneghel, and R. Tourky (2019). Learning under unawareness.
- Grant, S. and J. Quiggin (2013). Inductive reasoning about unawareness. *Economic Theory* 54, 717–755.
- Grant, S. and J. Quiggin (2015). A preference model for choice subject to surprise. *Theory and Decision* 79(2), 167–180.
- Heifetz, A., M. Meier, and B. C. Schipper (2006). Interactive unawareness. *Journal of Economic Theory* 130, 78–94.

- Kahneman, D. and A. Tversky (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47(2), 263–291.
- Karni, E. and M.-L. Viero (2013). “Reverse Bayesianism”: a choice-based theory of growing awareness. *American Economic Review* 103(7), 2790–2810.
- Karni, E. and M.-L. Viero (2015). Probabilistic sophistication and reverse bayesianism. *Journal of Risk and Uncertainty* 50, 189–208.
- Karni, E. and M.-L. Viero (2017). Awareness of unawareness: A theory of decision making in the face of ignorance. *Journal of Economic Theory* 168, 301–328.
- Klibanoff, P., M. Marinacci, and S. Mukerji (2005). A smooth model of decision making under ambiguity. *Econometrica* 73(6), 1849–1892.
- Knight, F. H. (1921, 2006). *Risk, Uncertainty, and Profit*. Dover Publications.
- Kochov, A. (2017). A behavioral definition of unforeseen contingencies. *Journal of Economic Theory* 175, 265–290.
- Lehrer, E. and R. Teper (2014). Extension rules or what would the sage do? *American Economic Journal: Microeconomics* 6, 5–22.
- Li, J. (2009). Information structures with unawareness. *Journal of Economic Theory* 144, 977–993.
- Mukerjee, S. (2015). *The Laws of Medicine: field notes from an uncertain science*. Simon Schuster.
- Mukerji, S. (1997). Understanding the non-additive probability decision model. *Economic Theory* 9(1), 23–46.
- Piermont, E. (2017). Introspective unawareness and observable choice. *Games and Economic Behavior* 106, 134–152.
- Pires, C. P. (2002). A rule for updating ambiguous beliefs. *Theory and Decision* 53(2), 137–152.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization* 3(4), 323–343.
- Quiggin, J. (2016). The value of information and the value of awareness. *Theory and Decision* 80(2), 167–185.
- Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley & Sons.
- Schipper, B. C. (2014). Unawareness - a gentle introduction to both the literature and the special issue. *Mathematical Social Sciences* 70, 1–9.

- Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica* 57(3), 571–587.
- Viero, M.-L. (2018). An intertemporal model of growing awareness.
- Walker, O. (2014). Unawareness with "possible" possible worlds. *Mathematical Social Sciences* 70, 23–33.
- Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica* 55(1), 95–115.