Games and Economic Behavior 106 (2017) 134–152

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Introspective unawareness and observable choice ☆

Evan Piermont

Department of Economics, Royal Holloway, University of London, United Kingdom

A R T I C L E I N F O

Article history:
Received 14 August 2016

JEL classification:
D01
D810
D830

Keywords:
Unawareness
Uncertainty
Epistemic preferences
Dynamic choice

A B S T R A C T

This paper explores the behavior of a decision maker (DM) who is unaware of some of the options available to her. The DM has a preference over consumption alternatives that is informed by her epistemic state: what she knows and what she is aware of. The main result is a characterization, via observable choice, of *introspective unawareness*—a DM who is both unaware of some information and aware she is unaware. Under dynamic introspective unawareness, the DM is unwilling to commit to future choices, even when given the flexibility to write a contingent plan that executes a choice conditional on the realization of uncertain events. This is a behavior that cannot be explained by uncertainty alone (i.e., without appealing to unawareness). In a simple strategic environment, the acknowledgment of unawareness can lead to strategic concealment of choice objects (i.e., actions), in turn, leading to a desire for incomplete contracts.

1. Introduction

This paper explores the behavior (i.e, observable actions) of a decision maker (DM) who is unaware of some of the options available to her. Due to the consideration of observability, the primary interest is in the DM's preferences (hypothetically embodied by choice data) and how patterns in preference change in response to the structure of awareness. I argue, unawareness produces distinct patterns, and so, attempting to model unawareness with uncertainty, regardless of how complex, will fail. As an example of when such issues arise and how they might alter predictions, I consider a simple contracting environment. I show that when unawareness is taken into account, players can have an incentive to conceal mutually beneficial actions.

To highlight the distinction between uncertainty and unawareness, consider Hal, who will buy a new smartphone in six months. He will have three options at the time of purchase: x, y, and z. Hal might not know which phone he would most like to purchase six months from now. This uncertainty could arise because he does not know the technical specifications of the phones, their price, etc., and his true preference depends on the realization of these variables. Contrast this to the case where Hal has never heard of phone z. Here, he is unaware of z, and so naturally, of any preferences there regarding. Importantly, if Hal is unaware of a piece of information (the existence of phone z), he is unable to make any choice based directly on this information.

☆ I thank Felipe A. Araujo, Mark Azic, Eddie Dekel, Simon Grant, Joseph Halpern, Edi Karni, Peter Klibanoff, Bart Lipman, Daniel Martin, Ian Morrall, Sofia Moroni, Tymofiy Mylovanov, Teddy Seidenfeld, Marciano Siniscalchi, Burkhard Schipper, Asher Wolinsky, the conference participants at RUD 2016 and LOFT 2016, my advisors Luca Rigotti and Roee Teper, Francoise Forges, and two anonymous referees for their helpful and insightful comments and suggestions. This project began while I was a graduate student at the University of Pittsburgh.

E-mail address: Evan.Piermont@rhul.ac.uk.

http://dx.doi.org/10.1016/j.geb.2017.09.008
0899-8256/

More subtle, but just as fundamental, is our acknowledgment of our own unawareness. Indeed, most people would readily admit they cannot conceive of all future technologies or trends, or exhaustively list the set of choices to be confronted in the upcoming week. This recognition of unawareness is important because it suggests that the things a DM is unaware of may play an indirect role in her decision making, even if they cannot be directly acted upon. Central to the analysis, then, is the DM who is (1) unaware and (2) aware she is unaware. A DM in such an epistemic state is referred to as *introspectively unaware*. By contrast, a DM who does not satisfy the second criterion would be referred to as *naively unaware*. In the presence of introspective unawareness, Hal might envision a world in which he prefers something to x and y; of course, he cannot know this *something* is z, as that would require him to be aware of z.

Under either uncertainty or introspective unawareness, Hal has a natural inclination to delay making his choice: if he cannot start using the phone for six months, he might as well wait until then to choose. However, the motivation for delay is different under the different types of ignorance. Under uncertainty, he would like to wait so as to make a decision based on the realization of the relevant variables (for example, the price of the phones). Under introspective unawareness, he would like to wait in case he becomes aware of something better than whatever he would have chosen today. If Hal had been naively unaware, he would have had no reason to delay; he would not consider the possibility of becoming aware of new information.

Hal is going to get his mother, Avril, to purchase the phone on his behalf, and has to instruct her *today* about which phone to purchase in six months. If Hal is either uncertain or introspectively unaware of his preference, it will not be optimal for him to specify any single phone. In the case of uncertainty, however, he could leave detailed instructions that would carry out his optimal choice: in the event the prices are ($\$x, \$y, \$z$), purchase phone x, ... etc. A commitment to consume (in the future) a particular alternative given the state of affairs is referred to as a *contingent plan*. If Hal's optimal decision depends only on the realization of some variables, it is enough to specify a contingent plan that depends on said variables. Contrast this to the case in which Hal is introspectively unaware. He cannot articulate any plan he is sure will carry out his optimal decision. This is because he would need to describe objects he is currently unaware of and to include such information in a contingent plan would require he is aware of it.

The main results of this paper show that the observable criterion for introspective unawareness is a strict preference for delaying choices at a positive cost rather than committing to a contingent plan, even when *any* plan is available. In particular:

(★1) *When the DM is fully aware, she is always willing to commit to some contingent plan.*
(★2) *Without full awareness, the DM might prefer costly delay to every contingent plan.*
(★3) *If the DM is unwilling to commit to any contingent plan, then she must be introspectively unaware.*

So a preference for costly delay that cannot be appeased by the appeal to contingent planning is the behavioral indication—in an exact sense—of introspective unawareness. The intuition is exactly as in the above example: the DM's language is not rich enough to specify the optimal contingent plan (unawareness), but is rich enough that she knows this fact (awareness of unawareness). Because the DM is not aware of *what* she is unaware of, the strength of the her aversion to commitment to a contingent plan is purely subjective. This can lead to behavior, particularly in strategic environments, that is substantially different than what is predicted by standard models.

After buying him his phone, Avril wants to hire Hal to write a computer program over his summer break. If Hal does not accept the offer, he could spend the summer working on developing his own app. Hal, being introspectively unaware, knows he does not fully understand what he could accomplish on his own. So, when contemplating a contract, he will weigh the benefit against his subjective assessment of this outside option.

The best language for Avril's job depends on the soon to be realized state s_1 or s_2. Hal is currently only aware of the programing languages: C++ (which is better if state s_1 is realized) and JAVA (better if state s_2 is realized). Hal, after thinking about what kind of app he could design with his current skill set, will accept the contract $c = [\text{C++}, \text{JAVA}]$ in which he writes the program for Avril, using C++ if state 1 is realized and JAVA in state 2 is realized.

In state s_2, HASKELL is truly the best language and Avril knows this. She believes, however, that were Hal to become aware of HASKELL, he would begin to investigate purely functional programming. This would draw his attention to the paucity of his awareness, and ultimately lead him to increase his subjective assessment of the outside option. In other words, by expanding his awareness a little, Hal becomes *more* averse to commitment, because he now believes there are many great possibilities that he might shortly become aware of. Therefore, it is possible that although $\hat{c} = [\text{C++}, \text{HASKELL}]$ is a strict improvement over c for both parties, it will be rejected; although the use of HASKELL is mutually beneficial, its existence will be concealed.

1.1. Decision theory, logic, and unawareness

Directly incorporating unawareness into a decision theoretic model introduces subtleties that need to be dealt with judiciously. First, one must take care to ensure the process of eliciting preferences from a DM does not affect her preferences. While asking a DM to rank risky prospects ostensibly does not affect her risk preference, asking her to contemplate objects of which she was formerly unaware would most certainly affect her awareness. Second, the *type* of unawareness considered (e.g., naive or introspective, object-based or state-based, etc.) must be rich enough to produce observable patterns, even when keeping in mind the previous concern. Finally, Modica and Rustichini (1994), Dekel et al. (1998) show that within the context of state space models, simply assuming that the DM is unaware of certain states (while retaining desirable properties of knowledge) is insufficient: the DM will either be fully aware or fully unaware.

To overcome these obstacles, I predicate my analysis on an epistemic modal logic with quantification (ranging over objects), based on a set of logical statements that include a formal description of the DM's preference, and adapted from Board and Sau Chung (2011). Each *state of the world* is defined by a set of statements which are true: statements that directly regarding the relation between objects (e.g., "x costs more than y" or "all objects are yellow or blue") and include how the DM ranks objects (e.g., "x is preferred to y"). In addition, there are statements describing the DM's epistemic state: what the DM *implicitly knows* (e.g., "the DM implicitly knows 'x is preferred to y' ") and which objects the DM *is aware of* (e.g., "the DM is aware of x"). The intersection of implicit knowledge and awareness is *explicit knowledge*. Implicit knowledge can be thought of as idealized knowledge—what the DM would know if she was fully aware and logically omniscient. In contrast, explicit knowledge can be thought of as working knowledge, subject to the DM's cognitive limitation of awareness.

The reason to consider a first order logic, rather than either a propositional logic (as in Fagin and Halpern, 1988) or generalized state space models (as in Heifetz et al., 2006; Heifetz et al., 2008) is to allow for introspection. The DM displays an aversion to commitment because she *anticipates* the expansion of her awareness, and thus, the logic must be able to capture *reasoning about unawareness*. Without quantification we run into problems: the statement "I know I am unaware of a quantum computer," while expressible, is inconsistent with the intuitive notion of awareness and is generally precluded axiomatically.[1] This problem can be circumvented by appealing to quantification: "I believe there exists something I am unaware of" is both expressible and intuitive.

1.2. Observability and unawareness

If a modeler were to ask a person on the street, or a subject in the laboratory, to choose between phone x and phone y, it is unreasonable to believe, at the time of her answer, that she is unaware of either x or y. The very act of asking forces the DM's hand.[2] This issue is exacerbated in the identification of introspection: any question regarding even the *existence* of unforeseen objects has the potential to change the DM's epistemic state.

Such issues are dealt with by relaxing what is meant by the revealed preference approach. Instead of providing the DM with a set and asking her to indicate her preferences thereover, the modeler asks the DM to provide both the relevant alternatives and her assessment of them. Elicitation can be incentivized as follows. Ask the DM for a contingent plan, c, and the dollar amount, δ, required for the DM to commit to the plan. Then a BDM mechanism (Marschak et al., 1964) is run as follows: a number, $d \in [0, \infty)$, is drawn from a known distribution. If $d < \delta$ the DM is not committed and delays her choice until some pre-specified time; if $d \geq \delta$ the DM's receives \$$d$ and her future consumption is implemented using the reported contingent plan. Given a contingent plan c, it is clearly incentive compatible to report the true cost of delay. Moreover, if c is preferred to c' (so that $\delta < \delta'$), it is similarly incentive compatible to report (c, δ) rather than (c', δ'). As such, if there exists a contingent plan to which the DM is willing to commit, she will report it with a corresponding $\delta = 0$.

A crucial aspect to the identification of unawareness contained in this paper is that it never requires the DM to contemplate objects she herself could not have conceived; it suffices for the modeler to consider the DM's preference over the set of objects she herself reported. The main contribution of this paper is, therefore, the assertion of a framework that characterizes introspective unawareness from choices regarding only information of which the DM is aware. Moreover, this approach works irrespective of the modeler's awareness or conception of the objective set of alternatives.

1.3. Organization

The structure of the paper is as follows. Section 2 introduces the logical underpinnings of the decision theory and expounds upon the choice patterns based on static preference. The main results are contained in Section 3, which introduces contingent plans and the notions of acceptability. Section 4 explores a simple strategic contracting game. A survey of the relevant literature can be found in Section 5. Appendix A discusses the connection to subjective state space models and a preference for flexibility. Additional results and proofs omitted from the text are contained in the appendix.

---

[1] Knowledge here should be interpreted as explicit knowledge. Such a statement violates KU introspection Dekel et al. (1998), Heifetz et al. (2008) and closure under subformulae Fagin and Halpern (1988). It is an important and open question as to if and how introspection can be incorporated into state space models.

[2] As pointed out by a referee, the issue may be partially skirted if the consumption objects are *subjective acts*, interpreted as a label for a function that maps subjective states of the DM to consequences. This is the route taken by Minardi and Savochkin (2017), Kochov (2015). However, this modeling choice creates as many issues as it resolves: it necessarily assumes that either (i) the modeler does not himself understand the objective alternatives to which the label refer, in which case any identification is filtered through the lens of modeler's subjective world view, or (ii) the modeler fully understands the mapping between labels and objective alternatives, in which case nothing can be identified in DMs who are more aware than the modeler.

2. Logical foundations: preferential logic

This section outlines the formal construction of the logic used in this paper. First Section 2.1 provides the syntax for well defined formulae. That is, a purely mechanical account of which strings of characters will be *well defined*. Then Section 2.2 endows well defined formulae with meaning by providing a semantic interpretation: *possible worlds semantics* adapted to consider preferential statements and awareness structures. Finally, section 2.3, considers an axiomatization (a method of deriving new true statements from old ones) corresponding to the semantic models.

2.1. Syntax

Preferential choices will be described directly by an epistemic logic. To this end, for each $n \geq 1$, define a(n at most countable) set of n place predicates denoted by $\alpha, \beta, \gamma, \ldots$. Assume the existence of a countably infinite set of variables denoted by $\mathcal{X} = a, b, c \ldots$. Then, any n place predicate followed by n variables is a well formed *atomic* formula. That is, if α is a 2 place predicate, then αab is a well formed atomic formula, with the interpretation that a and b stand in the α relation to one another. For example, if α is "greater than", then αab states that a is greater than b.

There are two distinguished predicates. The first, a binary predicate $\succcurlyeq$, represents weak preference (where $(a \succcurlyeq b)$ is used rather than $(\succcurlyeq ab)$). The second, a unary predicate, E, denotes existence. Semantically, variables will range over a set of consumption alternatives. Importantly, the set consumption alternatives is itself a component of the world, and therefore, will be allowed to be different in different situations. Hence, $E(a)$ postulates that in the true state of affairs, *a exists*. In addition, this language allows for universal quantification, $\forall$: $\forall a(b \succcurlyeq a)$ states that b is preferred to all a. Take note that variables are placeholders, and, until endowed with an interpretation, do not refer to any specific object.

There are also three (time indexed) modalities. L_t, A_t and K_t for $t \in \{0, 1\}$. Given a well formed formula φ (defined below), the interpretation of the modal operators is as in Fagin and Halpern (1988). L_t is implicit knowledge at time t; an agent implicitly knows φ, denoted $L_t\varphi$, if φ is true in every state of affairs she considers possible (at time t). A_t is awareness at time t; $A_t\varphi$ is interpreted as the DM is aware of φ at time t. Lastly, K_t is explicit knowledge—the conjunction of L_t and A_t. The DM explicitly knows φ if she implicitly knows it and is aware of it: $L_t\varphi \wedge A_t\varphi$.

Define the set of well formed formulae recursively: for any well formed formulae, φ and ψ, $\neg\varphi$, $\varphi \wedge \psi$, $\forall a\varphi$, $L_t\varphi$, $A_t\varphi$ and $K_t\varphi$ are also well formed. The resulting language is $\mathcal{L}$.

Taking the standard shorthand, $\varphi \vee \psi$ is short for $\neg(\neg\varphi \wedge \neg\psi)$, $\varphi \implies \psi$ is short for $\neg\varphi \vee \psi$, and $\exists a\varphi$ is short for $\neg\forall a\neg\varphi$. In addition, let P_t denote $\neg K_t\neg$, with the intended interpretation of $P_t\varphi$ as the DM considers φ possible; she does not explicitly know it is not the case. Per usual, an occurrence of a variable a is *free* in a formula φ if a is not under the scope of a quantifier, and is *bound* otherwise. A formula with no free occurrences is called a *sentence*.

2.2. Semantics

For a given language, $\mathcal{L}$, each DM is characterized by the tuple

$$M = \langle S, X, \{X_s\}_{s\in S}, \mathcal{V}, \{R_t\}_{t=0,1}, \{\mathcal{A}_t\}_{t=0,1}, \{u_s\}_{s\in S}\rangle.$$

M is referred to as a model (or, a model of decision making). $S = \{s, s', \ldots\}$ is a non-empty set of states of the world. X denotes a non-empty denumerable domain of the individual variables, the set of all possible values a variable might take. Elements of X are referred to using $x, y, z \ldots$. $X_s \subseteq X$ is a non-empty subset of consumption alternatives which exist in state s.[3] Truth values of atomic formulae will be assigned by $\mathcal{V}$, a function that assigns to each n place predicate and state of the world s, a class of n-tuples from X. If $(x_1 \ldots x_n) \in \mathcal{V}(\alpha, s)$, with $x_1 \ldots x_n \in X$, then $\alpha x_1 \ldots x_n$ is true in that model in state s. $\{R_t\}_{t=0,1}$ is an of accessibility relations on S for each time period; the interpretation of $R_t(s) = \{s'|sRs'\}$ is the states the DM considers possible when the true state is s. $\mathcal{A}_t(s) \subseteq X$ is a (possibly-empty) set of objects which the DM is aware of at time t in state s. Finally, $u_s : X \to \mathbb{R}$ is a utility representation of DM's true preference over X in state s.

Let $\mathcal{M}$ be the class of all models based on $\mathcal{L}$. Let $\mathcal{M}^{BND}$ denote the subclass of models in which $u_s|_{X_s} : X_s \to \mathbb{R}$ is a bounded function which attains its supremum for each $s \in S$.

The truth of a particular formula depends on the assignment of the free variables in that formula. Let an *assignment* be a function from the set of individual variables into objects: $\mu : \mathcal{X} \to X$. If μ and μ' are assignments that differ only in the object assigned to a then they are referred to as a-variants, and related by $\mu \sim_a \mu'$. Then a DM, M, is represented semantically via the operator $\models$, recursively, as

[3] I work with a varying domain model. A word should be said on this, as there is considerable philosophical debate regarding constant/varying domains. On one hand, it simplifies matters considerably to assume the same objects *hypothetically* exist in each possible world. One the other, the very intention that possible worlds be distinct means they might be defined by different objects. Here, I adopt the later view. This modeling choice allows for the possibility that a DM who is fully aware (aware of all *existing* objects) is nonetheless uncertain she is fully aware. For a lengthier discussion on constant v. varying domains within the context of unawareness, see Board and Sau Chung (2011), Halpern and Rêgo (2013).

$$
\begin{array}{ll}
(M,s) \models_\mu E(a) & \text{iff } \mu(a) \in X_s, \\
(M,s) \models_\mu (a \succcurlyeq b) & \text{iff } u_s(\mu(a)) \geq u_s(\mu(b)), \\
(M,s) \models_\mu \alpha a_1 \dots a_n & \text{iff } (\mu(a_1)\dots\mu(a_n)) \in \mathcal{V}(\alpha, s), \\
(M,s) \models_\mu \neg\varphi & \text{iff not } (M,s) \models_\mu \varphi, \\
(M,s) \models_\mu (\varphi \wedge \psi) & \text{iff } (M,s) \models_\mu \varphi \text{ and } (M,s) \models_\mu \psi, \\
(M,s) \models_\mu \forall a\varphi & \text{iff for all } \mu' \sim_a \mu' \text{ with } \mu'(a) \in X_s, (M,s) \models_{\mu'} \varphi \\
(M,s) \models_\mu L_t\varphi & \text{iff for all } s' \in R_t(s), (M,s') \models_\mu \varphi, \\
(M,s) \models_\mu A_t\varphi & \text{iff } \mu(x) \in \mathcal{A}_t(s) \text{ for all } a \text{ free in } \varphi, \\
(M,s) \models_\mu K_t\varphi & \text{iff } (M,s) \models_\mu L_t\varphi \text{ and } (M,s) \models_\mu A_t\varphi.
\end{array}
$$

A formula φ is *satisfiable* if there exists an M, and a state thereof, s, and an assignment μ, such that $(M,s) \models_\mu \varphi$. If $(M,s) \models_\mu \varphi$ for every assignment μ, write $(M,s) \models \varphi$. Given a DM, M, φ is *valid* in M, denoted as $M \models \varphi$, if $(M,s) \models \varphi$ for all s. Likewise, for some class of DMs, $\mathcal{N}$, φ is *valid* in $\mathcal{N}$, denoted as $\mathcal{N} \models \varphi$, if $N \models \varphi$ for all $N \in \mathcal{N}$. Finally, φ is *valid* (i.e., without qualification) if $M \models \varphi$, for all models M.

2.3. Axioms

Board and Chung show that the ***AWARE*** axiom system (presented in Appendix B) is a sound and complete axiomatization of the semantic structure introduced in section 2.2 (without the inclusion of preferential element statements).[4]

The semantics of preferential statements require two additional axioms positing the completeness and transitivity of preferences.

[CMP] $\forall a \forall b \big(\neg(a \succcurlyeq b) \implies (b \succcurlyeq a)\big)$.

[TRV] $\forall a \forall b \forall c \big((a \succcurlyeq b) \wedge (b \succcurlyeq c) \implies (a \succcurlyeq c)\big)$.

[BND] $\exists a \forall b (a \succcurlyeq b)$.

Because the set of consumption alternatives is assumed to be denumerable, completeness and transitivity suffice to ensure the existence of a representation. BND axiomatically imposes the existence of maximal elements. This will ensure, later, that a DM's unwillingness to commit to an action is because optimal decision do not exist. BND is valid in all finite models.

Proposition 2.1. *$\boldsymbol{AWARE} \cup$ CMP $\cup$ TRV (resp. $\cup$BND) is a sound and complete axiomatization of $\mathcal{L}$ with respect to $\mathcal{M}$ (resp. $\mathcal{M}^{BND}$).*

Proof. In Appendix D □

2.4. The structure of knowledge and awareness

Further axioms can impose structure on the DMs knowledge and awareness, and hence, semantically, on the accessibility relations and awareness sets. Consider the following:

[T] $L_t\varphi \implies \varphi$.	[LE] $E(a) \implies L_1 E(a)$.
[4] $L_t\varphi \implies L_t L_t\varphi$.	[L↑] $L_0\varphi \implies L_1\varphi$.
[5] $\neg L_t\varphi \implies L_t \neg L_t\varphi$.	[A↑] $A_0\varphi \implies A_1\varphi$.

It is well known, in the presence of a base axiom system, T, 4, and 5 correspond to the class of models where R_t is reflexive, transitive, and Euclidean,[5] respectively (see Fagin et al., 1995 for the propositional case and Hughes and Cresswell, 1996 for a first order treatment). Of note is the system $\boldsymbol{S5} = (\boldsymbol{AWARE} \cup \mathrm{T} \cup 4 \cup 5)$, corresponding to the class of models where R_t is an equivalence relation, and therefore, partitions the state space.

Axiom LE states that by time 1 the DM implicitly knows which alternatives are feasible. Because the model tacitly assumes that all consumption takes place at time 1, this requirement is tantamount to assuming that, at the time of consumption, the DM implicitly knows all *feasible* alternatives. While included for technical reasons,[6] this dictate has little conceptual bite, since existence could be modeled as the subset of objects about which the DM could feasibly choose at

[4] Given an axiom system $\boldsymbol{AX}$, and a language $\mathcal{L}$, we say that the formula $\varphi \in \mathcal{L}$ is a *theorem* of $\boldsymbol{AX}$ is it is an axiom of $\boldsymbol{AX}$ or derivable from previous theorems using rules of inference contained in $\boldsymbol{AX}$. Further, $\boldsymbol{AX}$ is said to be *sound*, for the language $\mathcal{L}$ with respect to a class of structures $\mathcal{N}$ if every theorem of $\boldsymbol{AX}$ is valid in $\mathcal{N}$. Conversely, $\boldsymbol{AX}$ is said to be *complete*, for the language $\mathcal{L}$ with respect to a class of structures $\mathcal{N}$ if every valid formula in $\mathcal{N}$ is a theorem of $\boldsymbol{AX}$.

[5] Recall, a relation is *Euclidean* if xRy and xRz imply yRz.

[6] To ensure that existence is a statement on which the DM can condition a contract.

time 1. Axioms LU and AU state that the DM's ignorance is diminishing. The first stipulates, semantically, that $R_1 \subseteq R_0$: the set of states the DM considers possible gets smaller over time. In the presence of ***S5*** this means that the DM's period 1 partition of the state space is a (weak) refinement of her period 0 partition. The second requires that $\mathcal{A}_0 \subseteq \mathcal{A}_1$: the DM becomes aware of more objects over time.

Proposition 2.2. ***S5*** $\cup$ LE $\cup$ LU $\cup$ AU *is a sound and complete axiomatization of* $\mathcal{L}$ *with respect to the subclass of models such that* $R_1 \subseteq R_0$ *and* $\mathcal{A}_0 \subseteq \mathcal{A}_1$ *and* X_s *is constant on the partition defined by* R_1 *(denoted* $\mathcal{M}^{LRN}$*).*

Proof. In Appendix D. □

It is immediate that LU and AU imply that the DM's explicit knowledge is increasing over time. Notice, the converse is not true. Increasing explicit preference does imply increasing awareness (which can be seen by examining knowledge of statements $E(a) \vee \neg E(a)$, which the DM alway implicitly knows). However, is possible that the DM's implicit knowledge diminishes even as her explicit knowledge grows. This is possible if she stops being able to distinguish between worlds which differ only regarding statements she is unaware of (at both time 0 and time 1), thereby not changing her explicit preference at all.

Let ***AW★*** $=$ ***S5*** $\cup$ Cmp $\cup$ Trv $\cup$ Bnd $\cup$ LU $\cup$ AU $\cup$ LE, and let $\mathcal{M}^\star$ denote the corresponding class of models. The remainder of the paper will consider only models of $\mathcal{M}^\star$.

2.5. Epistemic preferences

Preferential axioms play the role of traditionally decision theoretic restrictions (i.e., completeness, transitivity, etc.); any (satisfiable) theory including these restrictions will have a model of decision making adhering to the corresponding decision theoretic framework. The importance, therefore, of including preferential statement in our logic is that it provides us a language to make a clean distinction between true preference (a feature of the physical state) and the DM's understanding of her preference (a feature of her epistemic state) and to analyze interplay there between. Specifically, the distinction between some elementary (read, true) preference and the preference the DM *knows* or is *aware of*.

The discrepancy between the DM's "true" preferences, her implicitly known preferences, and, in the presence of unawareness, her explicitly known preferences, can be made formal. To do this, define the following three preference relations:

Definition 1. Let $M \in \mathcal{M}^\star$ be a model of decision making and let $s \in S$ denote some state. For each $(x, y) \in X \times X$, let $\mu_{x,y}$ be an assignment such that $\mu(a) = x$ and $\mu(b) = y$. Define the following relations on $X \times X$

$x \succcurlyeq_s y$ if and only if $(M, s) \models_{\mu_{x,y}} (a \succcurlyeq b)$,
$x \succcurlyeq_{L_t,s} y$ if and only if $(M, s) \models_{\mu_{x,y}} L_t(a \succcurlyeq b)$, and
$x \succcurlyeq_{K_t,s} y$ if and only if $(M, s) \models_{\mu_{x,y}} K_t(a \succcurlyeq b)$.

The following remark lends an equivalent, but purely semantic, definition of these three relations.

Remark 2.3. Let $M \in \mathcal{M}^\star$. For each $s \in S$:

1. $x \succcurlyeq_s y$ if and only if $u_s(x) \geq u_s(y)$,
2. $x v_{L_t,s} y$ if and only if $u_{s'}(x) \geq u_{s'}(y)$ for all $s' \in \mathbb{R}_t(s)$, and
3. $x \succcurlyeq_{L_t,s} y$ if and only if $x \succcurlyeq_{L_t,s} y$ and $x, y \in \mathcal{A}_t(s)$.

The interpretation is as follows: if $(M, s) \models_\mu (a \succcurlyeq b)$ then in state s the DM prefers $\mu(a)$ to $\mu(b)$. However, she may not know this fact. Specifically, if $(M, s) \models_\mu (a \succcurlyeq b) \wedge \neg L_t(a \succcurlyeq b)$, then in state s at time t, the does not (implicitly) know her preference. She considers some state in which her preference is reversed. The preference relations $\succcurlyeq_{L_t,s}$ represents the subset of the DM's true preference which are invariant in all states she considers possible. Further, it is possible that $(M, s) \models_\mu L_t(a \succcurlyeq b) \wedge \neg K_t(a \succcurlyeq b)$, so that the DM implicitly, but not explicitly, knows her preference. In this case, it must be that either $\mu(a)$ or $\mu(b)$ is not in the DM's awareness, $\mathcal{A}_t(s)$. While the implicit preferences (and, over $\mathcal{A}_t(s)$, her explicit preferences) inherit reflexivity and transitivity, the same can not be said about completeness.[7] Even if the DM explicitly knows her preferences are complete, she might not know what her preference is.

[7] As stated, the DM's explicit preference will be, in general, very incomplete. It is entirely possible, however, that the DM *completes* her preferences via some across-state aggregation, for example by making probabilistic judgments. This footnote, as well as footnotes 8 and 9, informally explore this line of reasoning, and argue that the main points of this paper are invariant to such a modeling choice. For a given model, M, let $S \setminus \mathcal{A}$ denote the quotient of S obtained by identifying states which differ only on the truth values assigned to statements about objects outside of $\mathcal{A}$, with $[s]$ being the equivalence class containing s. Let $\Gamma_{s,t} : \mathbb{R}^{S \setminus \mathcal{A}_t(s)} \to \mathbb{R}$ be a weakly increasing aggregator which (i) maps constant functions to their output (i.e, is non-trivial), (ii) is invariant to the value of $f(s)$ for any s such that $[s] \cap R_t(s) = \emptyset$ (i.e., is unaffected by states which are known to not have occurred) and (iii) is constant

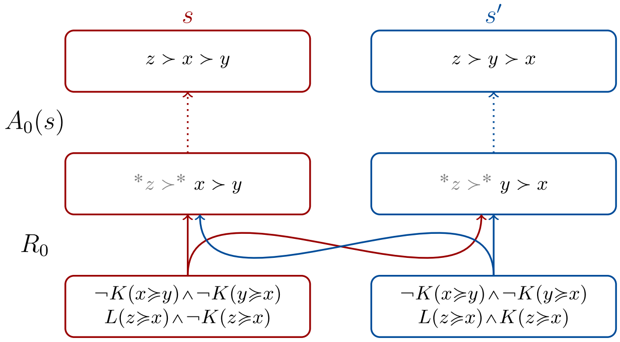

Fig. 1. A visual representation of Example 1. Assessable states are linked by arrows. A statement is implicitly known if it is true in every state linked by an arrow. Awareness structures are indicated by dotted lines—if a statement is true but contains object the DM is unaware of, it is written in gray and delineated with by *.

Example 1. There are two states of the world, $S = \{s, s'\}$ and three elements that can be consumed in each state, $X = X_s = X_{s'} = \{x, y, z\}$. The preference relations in each state are given by $u_s = [1, 0, 2]$ and $u_{s'} = [0, 1, 2]$ and the accessibility relation is the trivial $R_0 = S^2$. $\mathcal{A}_0(s) = \{x, y\}$ and $\mathcal{A}_0(s') = X$. Then, the implicit preferences (at both states) rank z above x and y, but cannot compare x and y and so, are not complete. So, $M \models_\mu L((a \succcurlyeq b) \vee (a \succcurlyeq b)) \wedge \neg L(a \succcurlyeq b) \wedge \neg L(b \succcurlyeq a)$, for any μ such that $\mu(a) = x$ and $\mu(b) = y$. Moreover, the explicit preference at state s is the restriction of implicit preference to x and y (and hence in the minimal reflexive relation), while at state s' it coincides with the implicit relation. (See Fig. 1.)

To be thoroughgoing, the following point needs to be made: I assume that only explicit preferences are observable in any meaningful way. Indeed, the DM can only asses, and therefore only act in accordance with, her explicit knowledge. As such, implicit preference is introduced only to act as a comparison to explicit preference. When the DM is fully aware, the two relations coincide. Therefore, any pattern in preference which systematically differs between $\succcurlyeq_{L_t,s}$ and $\succcurlyeq_{K_t,s}$ is indicative of unawareness. In other words, if we find a domain in which unawareness implies that $\succcurlyeq_{L_t,s}$ and $\succcurlyeq_{K_t,s}$ will impart different behavior, then such a behavior is an observable criterion for unawareness.

Remark 2.4. Let W be a denumerable set and $\geq \subseteq W \times W$. Then, the following are equivalent:

1. $\geq$ is a reflexive and transitive.
2. There exists a model, $M \in M^\star$ with $W = X$, and a state, s, thereof, such that $\geq = \succcurlyeq_{L_0,s}$.
3. There exists a model, $M \in M^\star$ with $W \subsetneq X$, and a state, s, thereof, such that $\geq = \succcurlyeq_{K_0,s}$.

Proof. **[(1) $\Longrightarrow$ (3)]**: Since $\geq$ is a reflexive and transitive it admits a multi-utility representation $\mathcal{U}$ (Evren and Ok, 2011). Let $S = \mathcal{U}$ and let $u_s = s$ for each $s \in S$. Let $\mathcal{A}_0(t) = W \subsetneq X_s = X = W \cup \{z\}$, for some $z \notin W$. Let $R_0 = S \times S$. Then for all $x, y \in W$, $x \geq y$ iff $u(x) \geq u(y)$ for all $u \in \mathcal{U}$ iff $u_{s'}(x) \geq u_{s'}(y)$ for all $s' \in R_0(s)$ and $x, y \in \mathcal{A}_0(t) = W$. By Remark 2.3, this is iff $x \succcurlyeq_{K_0,s} y$. **[(3) $\Longrightarrow$ (2)]**: Let $\langle S, X, \{X_s\}_{s\in S}, \mathcal{V}, \{R_t\}_{t=0,1}, \{\mathcal{A}_t\}_{t=0,1}, \{u_s\}_{s\in S}\rangle$, $s \in S$, be the model and state prescribed by (3). Then it is easily check that $\langle S, W, \{X_s \cap W\}_{s\in S}, \mathcal{V}|_W, \{R_t\}_{t=0,1}, \{\mathcal{A}_t\}_{t=0,1}, \{u_s|_W\}_{s\in S}\rangle$, s, will satisfy (2). **[(2) $\Longrightarrow$ (1)]**: By Remark 2.3, $\geq$ admits a multi-utility representation. By (Evren and Ok, 2011) it is a pre-order. □

The above result exposes the extent to which an outside observer can differentiate a DM who is unaware of some of the available action from a DM who is fully aware. Remark 2.4 demonstrates that, without fixing the underlying model (in particular, if the outside observer is not himself fully aware), the DM's revealed preferences are dually consistent with either epistemic condition. In other words, the *structure* of preference does not change as awareness changes, even if the domain of the preference does. The remainder of this paper qualifies this limitation, showing that the DMs preference over dynamic choices *can* reveal her unawareness.

3. Contingent planning

Within the models of $\mathcal{M}^\star$, the DM becomes more informed over time—both her uncertainty and her unawareness abate. Intuitively, anticipating either a reduction in uncertainty or in unawareness will lead to a preference for delay. In the case of

over the cells of the partition R_t (i.e., the DM implicitly knows how she is aggregating). One such example: $\Gamma_{s,t}$ is the expectation operator with respect to $p \in \Delta(S)$ conditional on $R_t(s)$. Notice that for all $x \in \mathcal{A}_t(s)$, $u_s(x)$ is $S \setminus \mathcal{A}_t(s)$ measurable, so that $\Gamma^*_{s,t} := \Gamma_{s,t} \circ (x \mapsto \{u_s(x)\}_{s \in S \setminus \mathcal{A}_t(s)})$ is a well defined functional $\mathcal{A} \times S \setminus \mathcal{A} \to \mathbb{R}$. For a given functional form of Γ, we consider the alternative model where DM's working preferences are the order generated by $\Gamma^*_{s,t}$. It is easy to see that such an order is a completion of $\succcurlyeq_{K_t,s}$.

uncertainty, the DM would prefer to delay making a choice so that she might condition her action on the information she expects to receive. Correspondingly, if the DM expects to become aware of novel action, she would prefer to delay making a choice to allow her future self the possibility of choosing an action of which she is currently unaware.

This section shows that by focusing on preference over *contingent plans*, the uncertainty-related preference for delay can be completely eliminated without reducing any unawareness-related preference for delay. A contingent plan is a commitment to a particular consumption alternative, the identity of which can depend on aspects of the true world. Over such a domain, a preference for costly delay can arise only as a response to unawareness, and is therefore an observable marker of unawareness, even if the underlying model is not identified by the modeler.

Fix a model M, with a set of consumption alternatives X. First, we must extend the DM's preferences, in particular those embodied by $\succcurlyeq_K$, to her preference over such dynamic objects. The mapping $(\varphi \mapsto b)_\mu$, where $\varphi \in \mathcal{L}$, $x \in \mathcal{X}$, and $\mu : \mathcal{X} \to X$ is the commitment to consume $\mu(b)$ in period 1, if, in the true state she explicitly knows φ is true and the prescribed consumption is feasible. In other words if $(M, s) \models_\mu K_1\varphi$. This is a *partial contingent plan*, since it does not specify what happens in states where φ is not known.

There are three relevant outcomes at time 1: (i) φ is not explicitly known, so her commitment does not bind, (ii) the commitment binds, $\mu(b)$ exists, and she does not explicitly know any object she prefers to $\mu(b)$, or, (iii) the commitment binds and either $\mu(b)$ does not exist or she explicitly knows an object she prefers to $\mu(b)$. So, the DM is willing to commit to $(\varphi \mapsto b)_\mu$ if she believes (iii) will never occur.

Definition 2. Given a model M, a partial contingent plan, $(\varphi \mapsto a)_\mu$, is **acceptable in state** s if

$$(M, s) \models_\mu K_0\Big(K_1\varphi \implies \big(E(b) \wedge \forall a P_1\big(b \succcurlyeq a\big)\big)\Big) \tag{3.1}$$

Parsing this formula, a DM is finds that partial plan acceptable if she knows at time 0, that at time 1, whenever she will explicitly know φ is true $\mu(b)$ will be feasible and she will also consider it possible that $\mu(b)$ is preferred to any other element. In other words, she knows she will not know φ is true and an object y is preferred to $\mu(b)$. Notice, if at time 0 she knows $\neg\varphi$ then $\big(K_1\varphi \implies (E(b) \wedge \forall a P_1(b \succcurlyeq a))\big)$ is always true and the DM finds the partial contingent plan acceptable.

Stringing together partial contingent plans can produce a complete contingent plan, which dictates a unique consumption in each state of the world.

Definition 3. Fix $M \in \mathcal{M}^\star$ and a state $s \in S$. A **contingent plan** is a triple $\langle \Phi, c, \mu\rangle$, such that Φ a finite subset of $\mathcal{L}$, $c : \Phi \to X$, and μ is an assignment, and such that for each $s \in S$ there exists an unique $\varphi \in \Phi$ such that $(M, s) \models_\mu K_1\varphi$.

With the model is fixed, contingent plans will often be written as $(c\colon \Phi \to \mathcal{X})_\mu$ with the understanding that Φ meets the necessary requirements in M. A contingent plan is a collection of partial plans, and, because it dictates a unique consumption in each state, the image of c partitions the state space. Since the relevant statement in Φ must be known at time 1, such a partition is a coarsening of the time 1 accessibility relation.

Remark 3.1. Fix $M \in \mathcal{M}^\star$, then $(c\colon \Phi \to \mathcal{X})_\mu$ is a contingent plan if and only if $\{s \in S \mid (M, s) \models_\mu K_1\varphi\}_{\varphi\in\Phi}$ is a coarsening of R_1.

Then intuition about when a DM would be willing to commit to such a plan extends from the case with partial plans. A contingent plan is acceptable if it provides outcomes that are no worse than what could have been selected by the DM had she waited until time 1 and then made a decision in accordance with her time 1 explicit knowledge.

Definition 4. A contingent plan, $(c\colon \Phi \to \mathcal{X})_\mu$, is **acceptable** to a DM, M, in state s, if

$$(M, s) \models_\mu K_0 \bigwedge_{\varphi\in\Gamma} \Big(K_1\varphi \implies \big(E(c(\varphi)) \wedge \forall a P_1\big(c(\varphi) \succcurlyeq a\big)\big)\Big), \tag{3.2}$$

with $a \notin \textsc{Im}(c)$. It is **unacceptable**, if it is not acceptable.

The DM knows that had she waited, she would not know any outcome dominates the outcome prescribed by the plan.[8] The following remark shows that we could equivalently have begun with intuition regarding the unacceptability of contingent plans, captured by a similar syntactic requirement:

---

[8] Assume the DM has a probability distribution $p \in \Delta(S)$ (as described in Footnote 7) and aggregates her preferences according expected utility. Then, acceptability would be recast as

$$K_0 \bigwedge_{\varphi\in\Gamma} \Big(K_1\varphi \implies \big(E(c(\varphi)) \wedge c(\varphi) \in \operatorname{argmax}\Gamma^*\big)\Big) \tag{3.3}$$

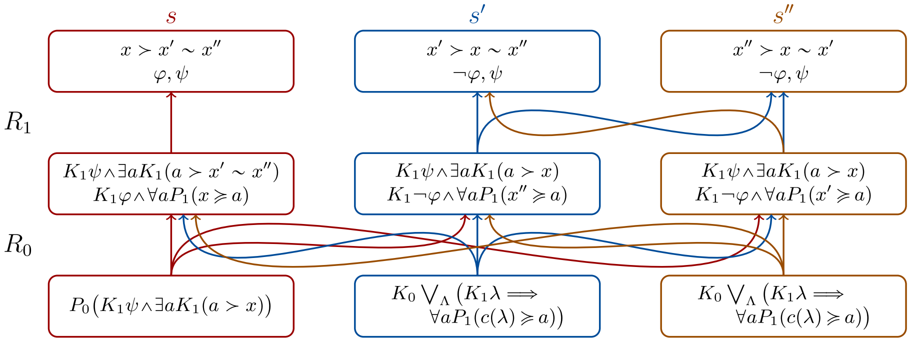

Fig. 2. A visual representation of Example 2. The statements in s indicate that any contract based on ψ is not acceptable, while the statements in s' and s'' related to the acceptability of the contract defined by (3.4).

Remark 3.2. Let $M \in \mathcal{M}^\star$. Then for any contingent plan, $(c\colon \Phi \to \mathcal{X})_\mu$, in each state, c is unacceptable if and only if

$$(M,s) \models_\mu P_0 \bigvee_{\varphi \in \Gamma} \Big(K_1\varphi \wedge \big(\neg E(c(\varphi)) \vee \exists a K_1\big(a \succ c(\varphi)\big)\big)\Big),$$

with $a \notin \mathrm{I{\scriptsize M}}(c)$.

Remark 3.2 follows straightforwardly from several applications of De Morgan's Law, the duality of $\langle\forall,\exists\rangle$ and $\langle K,P\rangle$, and, critically, on the fact that if φ and ψ are materially equivalent, and contain the same free variables, then $K_t\varphi \iff K_t\psi$. This last observation is a direct consequence of K for L_t and A3 for A_t.

There exists one potential issue when considering full awareness. It may be the DM is unaware of some aspect of the contingent plan itself, and therefore could not make reasonable choices regarding it. The following remark placates any such concern, showing that if a contingent plan is inarticulable, that is to say lies outside the DM's awareness, is always unacceptable.

Remark 3.3. Let $M \in \mathcal{M}^\star$ and $(c\colon \Phi \to \mathcal{X})_\mu$ be any contingent plan such that there exists a $\varphi \in \Phi$ and a $b \in \mathcal{X} \setminus A_0(s)$, such that either b is free in φ or $b = c(\varphi)$. Then $(c\colon \Phi \to \mathcal{X})_\mu$ is unacceptable in state s.

Remark 3.3 follows immediately from the fact that the variable b is free in expression (3.2), thereby ensuring the DM is not aware of such an expression, and subsequently that she does not explicitly know it.

3.1. The existence of acceptable contingent plans

A plan is unacceptable if the DM believes it is possible that if she waits until time 1 she will know a strictly better element than the one she is prescribed. Therefore, in the absence of unawareness, a contingent plans is acceptable if and only if it reduces exposure to uncertainty exactly as much as waiting until time 1. The following example (and especially the accompanying Fig. 3) provides intuition as to how the formal, syntactic definition of (un)acceptability is related to the DM's exposure to uncertainty.

Example 2. There are three states of the world, $S = \{s, s', s''\}$ and three elements that can be consumed in each state, $X = X_s = X_{s'} = \{x, x', x''\}$. The preference relations in each state are given by $u_s = [1, 0, 0]$, $u_{s'} = [0, 1, 0]$ and $u_{s''} = [0, 0, 1]$. $\mathcal{A}_0 = \mathcal{A}_1 = X$ (for all $s \in S$). The accessibility relations are given by the partitions $R_0 \cong \{S\}$, and $R_1 \cong \{\{s\}\{s', s''\}\}$.

Let μ be any assignment such that $\mu(b) = x$, $\mu(b') = x'$, $\mu(b'') = x''$. Let $\varphi \in \mathcal{L}$ be such that, under μ, φ is true only at states s. It is easy to verify that

$$c : \begin{cases} \varphi \mapsto b \\ \neg\varphi \mapsto b' \end{cases} \tag{3.4}$$

is an acceptable contingent plan in every state. The same holds true if $c(\neg\varphi) = b''$, but not $c(\neg\varphi) = b$. Contrastingly, if $\psi = \varphi \vee \neg\varphi$ then $(\psi \mapsto b)$, $(\psi \mapsto b')$, and $(\psi \mapsto b'')$ are *all* unacceptable contingent plans. (See Fig. 2.)

---

where $K_t\psi$ can be read as $p(\psi = \mathrm{TRUE} \mid R_t(s)) = 1$. Of course, (3.3) is not actually well defined because Γ^* is a semantic notion. Now if the DM places probability 0 on φ, she will also place probability 0 on $K_1\varphi$ and therefore the prescription $c(\varphi)$ does not affect acceptability. Hence, the existence of measure 0 events do not prohibit the DM from finding a contingent plan acceptable, whereas unawareness does. Intuitively, the DM places positive probability on learning novel outcomes, and so, cares about what happens in such an event.

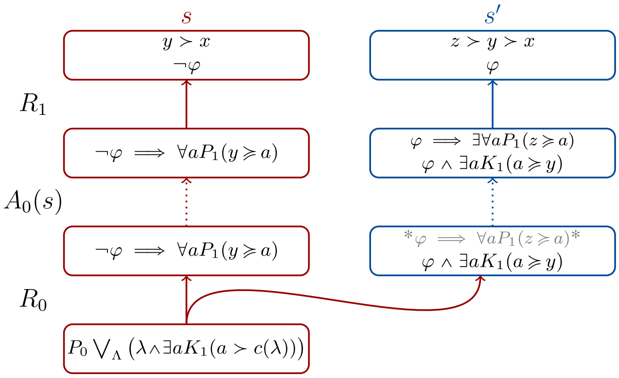

Fig. 3. A visual representation of Example 3.

For a contingent plan to be deemed acceptable, two things must occur: first the set of formulae on which it is based must be rich enough so that it could mimic any decision making process the DM could have implemented without a contingent plan (i.e., by waiting until 1 and making a single decision). In Example 2, any contract based on the tautology ψ is unacceptable, because it does not let the DM capitalize on the information she will receive by time 1. Second, the proscribed outcomes must not be dominated in the event to which they are associated. Returning to the example again, this is why c was acceptable but would have been unacceptable when $c(\neg\varphi) = b$. When the DM explicitly knows that $\neg\varphi$ has obtained—to wit, states s' and s''—$\mu(b)$ is a dominated element.

Theorem 3.4. *Let $M \in \mathcal{M}^\star$ with a finite state space, S, and such that $\mathcal{A}_0(s) = X$ for all $s \in S$. Then, there exists a contingent plan that is acceptable at every $s \in S$.*

Proof. In Appendix E. $\square$

Theorem 3.4 states that, when the DM is fully aware, an acceptable contingent plan always exists. In other words, in the absence of limitations on the DM's understanding of the world, she can always describe, exhaustively, the conditional actions she will take in the future. This result should come as little surprise, as it tacitly assumed in much economic formalism, from the elicitation of conditional preferences via Savage acts to the use of ex-ante strategies in repeated games.

To further interpret Theorem 3.4, assume the DM is tasked with a decision between (i) designing a contingent plan which will act as a commitment device, or (ii) making a consumption choice at time 1, and paying a cost δ. Because an acceptable contingent plan mimetically implements the DM's future choices, the DM who can report an acceptable plan has no desire to delay her choice and would thus report a value of $\delta = 0$. As such, a DM who is fully aware, and who can commit to any contingent plan, will not have a desire to delay.

The above line of reasoning relies on the dictate that the DM explicitly knows *what* it is she might learn. In other words, while a contingent plan allows the DM to specify consumption in the event she learns a particular piece of information, it is requisite she knows (at time the contingent plan is written) every piece of information she might learn. This is markedly impossible in the event she is unaware, and aware of her unawareness! This is, finally, the behavioral implication of unawareness: the unwillingness to commit to *any* contingent plan, even under circumstances that make implicit knowledge very well behaved.

The following example shows even under very well behaved implicit knowledge, the existence of unawareness can render *every* contingent plan unacceptable.

Remark 3.5. There exist models of $\mathcal{M}^\star$, and a state $s \in S$, such every contingent plan is unacceptable at s.

Proof. Example 3. $\square$

Example 3. Let S contain two states, s and s', which the DM can distinguish only in period 1. Let $X_s = \{x, y\}$, $X_{s'} = X = \{x, y, z\}$, with $u_s = u_{s'} = [0, 1, 2]$. $\mathcal{A}_t(s) = X_s$ and $\mathcal{A}_t(s') = X_{s'}$ for $t = 0, 1$.

Let $(c\colon \Phi \to \mathcal{X})_\mu$ be any contingent plan. Let φ be the unique statement that holds at s'. For acceptability, it must be that $(M, s) \models_\mu \forall a P_1\big(c(\varphi) \succcurlyeq a\big)$, which is true if and only if $c(\varphi) = z$. However, by Remark 3.3, c is then unacceptable. (See Fig. 3.)

Example 3 shows that once unawareness is introduced, there is no longer a guarantee of acceptability.[9] The introduction of unawareness has fundamentally changed the behavior of the DM—creating a preference for delay that cannot be assuaged by allowing the DM to make conditional decisions. And, by Proposition 3.4, this behavior, unlike incompleteness or a preference for flexibility, cannot be explained in a framework with full awareness, no matter how much uncertainty exists. It is a behavioral trait that indicates the presence of unawareness.

A qualification: the DM in Example 3 is not *technically* fully aware, she *is* aware of all extant objects. However, even though the DM is aware of every feasible alternative she does not know this fact. The DM believes she might learn about new objects—to which, owing to linguistic constraints, she cannot currently commit—and is therefore disinclined towards commitment. Thus, the possibility of unawareness is sufficient to engender a preference for delay. The next result shows it is also necessary.

Theorem 3.6. *Let $M \in \mathcal{M}^\star$ with a finite state space. If M admits contingent plans, and the DM finds every contingent plan is unacceptable at state s, then the DM explicitly knows it is possible she is unaware. Specifically,*

$$(M, s) \models K_0 P_0 (\exists a (\neg A_0 E(a))) .$$

Proof. In Appendix E. □

A DM cannot be so unaware she is not even aware waiting will afford her a more complete world view. That is, the DM must be introspectively unaware. The intuition of this result is straightforward. The DM, acting on explicit knowledge, must explicitly know all contingent plans are unacceptable and this requires she is aware she will have more choices if she does not commit.

4. Unawareness and contracts

This section contains a simple example to show how the framework presented above could be used in applications. Assume there are two players: a *Principal* (player p), who is offering a take-it-or-leave-it contract to an *Agent* (player a). The model takes place in an interactive awareness structure in which players' knowledge and awareness are defined over atomic statements and both their own and their opponents knowledge and awareness.[10] It is in this framework that I will show the Principal has an incentive to conceal mutually beneficial information. The intuition being that, although certain novel actions are Pareto improving in every ex-post scenario, the Agent will react to the discovery of novel actions by becoming more sensitive to her own unawareness, hence increasing her aversion to commitment. In other words, the display of surprising outcomes indicates to the Agent that the novel outcomes are more valuable than she previously thought; the added value to waiting (and taking an outside option) is greater than the value added by the novel outcome itself. Further, I will show that this incentive can naturally lead to the optimality of incomplete contracts.

The timing is as follows. In period 0, the Principal offers the Agent a continent plan to be executed in period 1 and some monetary transfer, $m : S \to \mathbb{R}$; to make matters as simple as possible, the transfer is in utility terms and the contingent plan is a function from S to X.[11] If the offer is rejected the Agent can take an outside offer, some action in X. When the Agent accepts a contract, c, then her state contingent utility is $u_{s,a}(c(s)) + m(s)$ and the Principal's is $u_{s,p}(c(s)) - m(s)$. If the contract is rejected, the Principal receives a utility of 0.

Of course, to make our problem well defined, we also have quantify the Agent's perceived value of the outside option. In the case of full awareness (or, naive unawareness, as in Auster, 2013), the Agent would have a well defined understanding over the outside option. This is not the case with introspective unawareness, as the Agent is aware of the possibility that waiting will afford novel actions. So consider a mapping $\delta : 2^X \to \mathbb{R}^S$ with the restriction that

$$\delta(\mathcal{A})(s) \geq \max_{x \in X \cap \mathcal{A}} u_{s,i}(x), \tag{4.1}$$

In light of the results in section 3, I further restrict that (4.1) holds with equality if the Agent is fully aware or naively unaware. δ captures the DM's attitude towards unawareness, her perceived value to the objects that she is currently unaware of (also, implicitly, the likelihood of discovering these novel actions in different states). Keeping in line with the formal model, I do not take a stand on how the Agent aggregates across states. Still, we can assume that it is individually rational for each DM to accept any *acceptable* contract (i.e., one that weakly dominates the outside option) and reject any unacceptable contract. What makes this problem distinct from the classical contracting problem (even one with indescribable states/alternatives) is that the Principal, though offering a contract, has the ability to alter the Agent's awareness.

[9] To conclude the discussion from Footnotes 7 and 8, notice that, for any ordering $\Gamma^*_{s,t}$ the alternative notion of acceptability given by (3.3) is strictly stronger than definition in the body of this paper. Therefore, the failure of existence show by Theorem 3.5 persists even when the DM's "explicit" preferences are complete. Finally, notice also, Theorem 3.4 would still hold under the more restrictive definition: since $\Gamma^\star_{s,1}$ is constant across $R_1(s)$, the DM can choose $c(\varphi)$ to be $\Gamma^\star_{s,1}$ maximal in the event that φ obtains. It is in this sense that I mean the main results are agnostic to whether the DM's actions are in accordance with her explicit preferences or a completion of them.

[10] It is clear that the same axiomatization will suffice, simply by adding additional indexes to the modalities.

[11] In this section, I assume $X_s = X$ for all $s \in S$. This is inessential for all results/intuition, but expedient in saving on notation.

4.1. The principal's problem

I focus on the case where actions are verifiable and the Principal is a standard economic DM: a fully aware expected value maximizer. As such, the Principal's problem is simply to offer the acceptable contract that maximizes his expected payoff. That is, maximize his payoff subject to a participation constraint on behalf of the Agent. Moreover, we will assume the Agent explicitly knows the Principal is fully aware. However, and unlike prior application of awareness, the Agent is introspectively unaware.

Let $S = \{s_1, s_2\}$ and $X = \{x, y, z, w\}$. Assume that neither player implicitly knows the state in period 0 (and that the Principal believes the states are equally likely), but both will know the state in period 1. Assume utilities are given by

$u_{s,p}$

	x	y	z	w
s_1	4	1	3	0
s_2	1	3	4	6

$u_{s,a}$

	x	y	z	w
s_1	3	1	2	6
s_2	4	3	5	0

In period 0, let $\delta(\mathcal{A}) = (3, 4)$ if $z, w \notin \mathcal{A}$ and $\delta(\mathcal{A}) = (6, 5)$ otherwise. Consider the case where the Agent's initial awareness is $\mathcal{A}_{0,a} = \{x, y\}$ and, if left unperturbed by the Principal's offer, it remains her awareness in period 1. Alternatively, if the Principal offers the contract c, then the Agent's awareness set becomes $\mathcal{A}_{0,a} = \{x, y\} \cup \textsc{Im}(c)$.

What is the Principal's optimal strategy, given that he is constrained to offer complete contracts? Notice that if he offers the contract $\langle c^\star = (x, y), m^\star = (0, 1)\rangle$, it is accepted—the Agent is indifferent between accepting and rejecting, given that her perception of the value of the outside option is also $(3, 4)$. The Principal gets a utility of $\frac{1}{2}(4 - 0) + \frac{1}{2}(3 - 1) = 3$. It is easy to verify that this is the best the Principal can do. To see this, note that if the Principal offers a contract containing either z or w, he must provide the Agent with a utility of at least $(6, 5)$ (the Agent's new participation constraint). The best way for the Principal provide this is, $\langle c' = (x, z), m'' = (3, 0)\rangle$. But this gives the Principal an expected utility of $\frac{5}{2}$, worse than $\langle c^\star, m^\star\rangle$.

Nonetheless, the contract $\langle c'' = (x, z), m'' = (0, 0)\rangle$ makes both players strictly better off: c'' provides state contingent utilities of $(3, 5)$ and $(4, 4)$ for the Agent and Principal, respectively. Hence, when the Principal is constrained to offer a complete contract, he willingly conceals a Pareto improving action. The intuition is simple: expanding the Agent's awareness makes her more aware of her own awareness, and hence she displays a larger aversion to commitment. This second effect outweighs the first, so the Principal chooses not to disclose the actions.

Now, consider the case where the Principal can offer an incomplete contract. Such a contract does not provide any alternative for a particular state, upon the realization of which the players renegotiate. Now the Principal can offer the contract $c = (x, \cdot)$ (read: x in state 1, re-negotiate in state 2). This is acceptable to the Agent, since $\delta(\{x, y\})(s_1) = 3 = u_{s_1,a}(x)$. In period 1, if state s_2 is realized, the Principal offers the new contract $c = z$. This is again acceptable since $\delta(\{x, y, z\})(s_2) = 5 = u_{s_2,a}(z)$. Therefore by appealing to incomplete contracts, the Principal can implement his unconstrained optimal contract.

One component obviously missing from this example is how the Agent's perception of unawareness reacts to the offer of an incomplete contract. It is reasonable to assume the agent believes when such a contract is offered, it must be due to strategic concerns relating to options outside if her current awareness. Hence the offer of an incomplete contract is itself reason to change her perception of the value of delay. This effect cannot be captured at all by naive unawareness, and highlights the importance of creating a richer epistemic framework. However, because this behavior is complicated, and the agent's shifting perception is likely subject to equilibrium effects, I leave any formal analysis to future work.

It is worth briefly addressing the relation between this environment and previous work connecting awareness with incomplete contracts. There is a large body of literature on incomplete contracts arising from the indescribably of states, leading to the well known discussion of Maskin and Tirole (1999). They show, so long as players understand the utility consequences of states, indescribability should not matter. This paper, on the other hand, allows the players to have asymmetric awareness regarding the set of actions that can be taken. Because the Agent is not fully aware of the set of actions, simply offering a particular contract might alter her awareness state, and therefore, her preferences.

The above example, while highly stylized, is indicative of a general phenomena. Although the effect of unawareness can be quantified via δ, and delay can be calculated, unawareness introduces behavior that intrinsically different than uncertainty. Unlike in the more standard framework, the value of delay (i.e., the outside option) changes with the Agent's epistemic state, and therefore is itself a function of the contract being offered. As such, there may exist feasible contracts which are initially individually rational, but cease to be so when offered. It is this effect, driven by introspective unawareness, that can make incompleteness strictly beneficial.

Remark 4.1. Let $\langle c^\star, m^\star\rangle$ be a contract accepted in equilibrium. If there exists some $\hat{x} \in X$ such that $\hat{x}$ is strictly preferred to $c^\star(s)$ in the same (non-empty) set of states for both players, then the Agent is introspectively unaware.

Proof. We will show there exists some $D \subseteq X$ and $s \in S$ such that $\delta(D)(s) > \max_{x \in X \cap D} \bar{u}_{s,a}(x)$. Since by definition (4.1) holds with equality when the Agent is fully aware or naively unaware, this suffices to prove the claim. Assume no such D existed. Let $\hat{S}$ denote the set of states where $\hat{x}$ is strictly preferred to $c^\star(s)$. Consider a new contract c, whose value equals $c^\star(s)$ for all s except in $\hat{S}$ where $c(s) = \hat{x}$.

We now claim that $\langle c, m^\star\rangle$ would have been accepted and is preferred by the Principal to $\langle c^\star, m^\star\rangle$. Indeed, towards the first claim, notice by construction $u_{s,a}(c(s)) \geq u_{s,a}(c^\star(s))$ for all s. Hence, we need only worry about states such that $\delta(\mathcal{A}_{0,a} \cup \hat{x})(s) > \delta(\mathcal{A}_{0,a})(s)$. But, this, by our assumption of naiveté, can only happen if $\delta(\mathcal{A}_{0,a} \cup \hat{x})(s) = u_{s,a}(\hat{x})$, which only happens for $s \in \hat{S}$. Towards the second claim, notice by construction it is also true that $u_{s,p}(c(s)) \geq u_{s,p}(c^\star(s))$ for all s. □

Naive awareness can also induce the Principal to withhold Pareto improving contracts, so long as the two players' preferences are not aligned with regard to the novel outcome. The Principal may withhold information strategically, as the novel outcomes may be of direct value as an outside option, making the participation constraint harder to satisfy. This is similar in spirit to the arguments put forth in Filiz-Ozbay (2012) and Auster (2013), where the agents are naively unaware.

5. Literature review

This paper is within the context of two distinct, albeit related, literatures: that on epistemic logic and unawareness, and that on unawareness and unforeseen contingencies in decision theory. Unawareness was first formalized within modal logic by Fagin and Halpern (1988), who introduced the modal operator for awareness, A, and explicit knowledge, K. This was extended later by Halpern and Rêgo (2009) to include quantified statements that allow for introspective unawareness, and extended further by Halpern and Rêgo (2013), to allow the agent to be uncertain about whether she has full awareness or not. Quantification, in these logics, is ranges over formulae. Because the foremost concern is over the alternatives that can be consumed (and over which preferences can be defined), I make use of the logic introduced by Board and Sau Chung (2011) and Board et al. (2011), where awareness is based on objects and predicates rather than the formulae themselves. Board and Sau Chung (2011) was the first to point out that fixed domain semantics do not allow the DM to be uncertain about whether she is fully aware.

In economics, *state space models*—the semantic structure that include states, and define knowledge and unawareness as operators thereon, as in this paper—have been of particular interest. Modica and Rustichini (1994) and Dekel et al. (1998) both provide beautiful, albeit negative, results in this domain. They show, under mild conditions, unawareness must be in some sense trivial; the DM is either fully aware or fully unaware. While Modica and Rustichini (1994) consider a specific awareness modality, Dekel et al. (1998) show, under reasonable axioms, state-space models do not allow any non-trivial unawareness operator. As stated, this would be a very damning result for this paper, as it would imply either $\succcurlyeq_K = \succcurlyeq_L$ or $\succcurlyeq_K = \emptyset$, either way, not making for an interesting decision theory. This paper eschews the issue by disentangling explicit and implicit knowledge. Considering these forms of knowledge separately avoids ever simultaneously satisfying the necessary axioms for DLR's negative result. A far more succinct and intuitive discussion than I could hope to achieve is found in Section 4 of Halpern and Rêgo (2013), and so, I refer the reader there.

Beyond the separation of implicit and explicit knowledge, there have been other approaches to the formalization of unawareness that circumvent the problems outlined in the previous paragraph. Modica and Rustichini (1999) propose models in which the DM is aware only of a subset of formulae (necessarily generated by primitive propositions), and entertains a subjective state space (a coarsening of the objective state space) in which the DM cannot distinguish between any two states that differ only by the truth of a proposition of which she is unaware. Heifetz et al. (2006) and Heifetz et al. (2008) consider a lattice of state spaces that are ordered according to their expressiveness. In this way, unawareness is captured by events that are not expressible from different spaces—events that are not contained in the event nor the negation of the DM's knowledge. Li (2009) also provides a model with multiple state spaces, where the DM entertains a subjective state space (similar to the above papers, the set of possible state spaces forms a lattice). This allows the DM to be unaware of events in finer state spaces, while having non-trivial knowledge in coarser state spaces. It is an open problem as to how introspection can be incorporated into multi-state-space models.

The decision theoretic take on unawareness is primarily based on a revealed preference framework, and so, unlike its logical counterpart does not dictate the structure of awareness but rather tries to identify it from observable behavior. The first account of this approach (and which predates the literature by a sizable margin) is Kreps (1979). Kreps considers a DM who ranks menus of alternatives, and whose preferences respect set inclusion. The motivation being larger menus provide the DM with the flexibility to make choices after *unforeseen contingencies*. This interpretation, while not strictly ruled out by the model, is certainly not its most obvious interpretation, especially in light of the titular representation theorem. That Krepsian behavior can always be rationalized in a model without appealing to unawareness is shown formally in Theorem A.1; a longer discussion in relation to this paper is found in Appendix A.

More recently there has been a growing interest in modeling the unaware DM. Kochov (2015) posits a behavioral definition of unforeseen contingencies. He considers the DM's ranking over streams of acts (functions from the state space to consumption). An event, E, is considered foreseen if all bets on E do not distort otherwise prefect hedges. That is to say, an event is unforeseen if the DM cannot "properly forecast the outcomes of an action" contingent on the event. Kochov shows the events a DM is aware of form a coarsening of the modeler's state space. In a similar vein, Minardi and Savochkin (2017) also contemplate a DM who has a coarser view of the world than the modeler. This coarse perception manifests itself via imperfect updating; the DM cannot "correctly" map the true event onto an event in her subjective state space. The events that are inarticulable in the subjective language of the DM can be interpreted as unforeseen. However, in these works, the objects of which the DM is supposedly unaware are encoded objectively into the alternatives she ranks. Because of this, I argue they are behavioral models of *misinterpretation* rather than unawareness.

Karni and Vierø (2016), Grant and Quiggin (2014) are more explicit about modeling unawareness, and, along with their companion papers, are (to my knowledge) the only decision theoretic paper that deals with unawareness of consumption alternatives, rather than contingencies. They examine a DM who evaluates acts which may specify an alternative explicitly demarcated as "something the DM is unaware of," and who can be interpreted as possessing probabilistic belief regarding the likelihood of discovering such an outcome. A key difference between this paper and the above papers, is that the latter require the modeler fully understands the awareness structure of the DM, before proceeding the analysis. Indeed, the construction of primitives in each model rely exactly the set of outcomes (resp., events) about which the DM is aware (resp., can fully describe)—*conceivable acts* (resp., *describable conceivable states*) in Karni and Vierø (2016) and *foreseen-contingencies* (resp., *surprise free acts*) in Grant and Quiggin (2014). Further, the authors' identification rely on significant structural assumptions (e.g., expected utility type tradeoffs) placed on the DM's preferences over acts that provide unforeseen outcomes—contrastingly, this paper does not require the modeler to take a stand on the structure of preferences nor awareness (beyond that it does not diminish over time).

Grant and Quiggin (2012) develop a model to deal with unawareness in games, founded on a modal logic which incorporates unawareness in a similar way to Modica and Rustichini (1994). They show that while this model is rich enough to provide non-trivial unawareness, it fails to allow for introspective unawareness, even when first order quantification is permitted. (This limitation arises because of the desired interplay between the structure of knowledge and the structure of awareness as facilitated by the game theoretic environment.) By relaxing the connection with the modal underpinnings, they then consider possible heuristics that a player might exhibit when she inductively reasons that she is introspectively unaware. In a companion paper, Grant and Quiggin (2014) provide a (decision theoretic) axiomatization of such heuristics.

Morris (1996) works somewhat in the reverse direction of the current paper, providing a characterization of different logical axioms (for example, K, T, 4, etc.) in terms of preferences over bets on the state of the world. Schipper (2014) extends this methodology to include unawareness structures as described in Heifetz et al. (2006). In a similar set up, Schipper (2013) constructs an expected utility framework to elicit (or reveal) a DM's belief regarding the probability of events (when she might be aware of some events). Schipper concludes, the behavioral indication of unawareness of event E is that the DM treats both E *and* its complement as null. The idea that a DM might manifest choice objects via a formal language is not new. Contingent plans, in this paper, share much with mapping between syntactic tests and savage acts in Blume et al. (2009). In both constructs, the DM constructs conditional statements that, under a given semantic interpretation, correspond to a savage act over a state space.

Finally, this characterization is of particular interest in relation to models of subjective learning. To identify what a DM believes she might learn, axiomatizations (Ergin and Sarver, 2010; Riella, 2013; Dillenberger et al., 2014, 2015; Piermont et al., 2016; Piermont and Teper, 2017) often include the requirement that any dynamic choice behavior is indifferent to some contingent plan—in essence, assuming the existence of acceptable, and articulable, plans.[12] As such, the results of this paper mandate that a theory of subjective learning under unawareness cannot be built on the same machinery. Put differently, current models of subjective learning necessarily reduce all learning to resolution of uncertainty rather than from the arrival unanticipated information.

6. Conclusion

This paper contemplates a framework that separates a DM's knowledge from her awareness, in such a way that allows the DM to reason about her own ignorance. Within this environment, I assume the DM has a ranking over consumption alternatives that is informed by her epistemic state (i.e., what she knows and what she is aware of). The main result is a characterization of the effect of unawareness on observable choice, and the provision of the requisite domain for identification. When the DM is introspectively unaware, she will exhibit a preference for costly delay, even when given the flexibility write any contingent plan she wishes.

One last point: In interpersonal contracting environments, where disputes might arise, the DM might not be able to prove her explicit knowledge. Because the conditioning events of a contingent plan are generated by the DM's explicit knowledge, some contracts might not be enforceable, if the relevant events cannot be externally verified. In this case, the DM might prefer to delay choices even when no unawareness is present. While this might complicate identification in interpersonal environments, it also provides an interesting link between two behaviors. When there is a fear that contracts will not be upheld by a court, it has the same effect as introducing introspective unawareness: both situations prohibit the DM from writing a contract that she explicitly knows will implement her future behavior (i.e., her behavior if there had been no contract). Thus, viewed at a higher level, these two phenomena are the same. They both place constraints on which types of contracts can be implemented, and therefore, both drive a wedge between the value of the optimal (constrained) contract and the value of delaying choice.

[12] The first two references do not directly construct such plans; nonetheless, the interpretation of both papers concerns a DM who constructs a contingent plan after observing the choice objects they receive.

Appendix A. A preference for flexibility

One interpretation of Kreps (1979) is the anticipation of learning induces a *preference for flexibility*. That is, the DM's preference over menus (i.e, subsets of X), respects set inclusion: if $m' \subseteq m \subseteq X$ then m is preferred to m'. A DM who expects to learn her true preference, but is currently uncertain, will prefer the flexibility to make choices contingent on the information she learns. In this section, I will show that the Krepsian framework can be faithfully reproduced as a special case of the general model outlined above. In particular, this special case is one of full awareness; as such, the unforeseen contingencies interpretation is not strictly needed, and a preference for flexibility is not alone the behavioral indication of unawareness.

Definition 5. A menu, $m \subseteq X$, **s-dominates** a menu, $m' \subseteq X$, (at period t), if and only if

$$(M, s) \models_\mu K_0 \bigvee_{a \in \mu^{-1}(m)} K_1 \bigwedge_{b \in \mu^{-1}(m')} (a \succcurlyeq b), \tag{A.1}$$

for some surjective assignment μ. Further, m **strictly** s-dominates m', if it dominates m' and m' does not dominate m.

That is, m dominates m' if the DM explicitly knows that irrespective of the state of affairs, she will choose an alternative out of the menu m rather than m'. Of course, she does not need to know which element is the maximal one. For example:

Example 4. Let $S = \{s, s'\}$ and $X = X_s = X_{s'} = \{x, y, z\}$, with $u_s = [0, 1, 2]$ and $u_{s'} = [2, 1, 0]$. Let $R_0 = S^2$ and $R_1 = \{(s, s), (s', s')\}$ and $\mathcal{A} = X$ for both states and time periods. Notice, $\{x, z\}$ strictly s-dominates $\{y\}$. Indeed, $(M, s) \models K_1(x \succcurlyeq y)$ and $(M \succcurlyeq s') \models K_1(z \succcurlyeq y)$. So, $M \models K_0(K_1(x \succcurlyeq y) \vee K_1(z \succcurlyeq y))$. So, the DM knows, in the true state of affairs, either x or z is preferred to y, but does not know which preference is her true preference.

The following result shows that beginning with the s-dominance relation, generated by some epistemic model, and extending it to a weak order captures exactly the "preference for flexibility" described in Kreps (1979). In fact, the converse it also true: every Krepsian DM can be formulated as the extension of an s-dominance relation with respect to some epistemic model.

Theorem A.1. *Let $\geq$ be a weak order over the non-empty subsets of X. The following are equivalent:*

(i) *$\geq$ satisfies Kreps' axioms,*
(ii) *There exists an $M \in \mathcal{M}^\star$ with u_s invariant over the partition induced by R_1, $\mathcal{X}_s = A_0(s) = X$ for all $s \in S$ and such that for some $s' \in S$ we have m s'-dominates m' implies $m \geq m'$, and m strictly s'-dominates m' implies $m > m'$.*

Proof. In Appendix E. □

Notice that the dominance relation, projected onto singleton menus, produces, $x \succcurlyeq_{K_{s,0}} y$ if and only if $\{x\}$ s-dominates $\{y\}$. Hence a preference for flexibility can be seen as a natural extension of multi-utility models.

Notice there is a natural connection between acceptability and the dominance relation over menus. Intuitively, when the menu is thought of as the image of a contingent plan, then the contingent plan specifies how the DM will choose out of the menu. Under this interpretation, Theorem A.2 the provides the connection between having a well defined preference over menus and being willing to accept a contingent plan.

Remark A.2. Let $M \in \mathcal{M}^\star$ satisfies the requirements of Theorem A.1(ii) then a finite menu $m \subseteq X$ is not strictly s-dominated if and only if it is the image of an acceptable contingent plan, $(c\colon \Phi \to \mathcal{X})_\mu$.

If a menu is undominated (according to the definition given by (A.1)), one must be able to construct an acceptable contingent plan from it. In particular, such a menu must not be dominated by $m = X$: for every alternative, $x \in X$, there is a corresponding element of m which will be known to be maximal. Conversely, if a contingent plan is acceptable it specifies each alternative which will be explicitly known to be maximal, and hence, the collection of all proscribed consumption alternatives must form an undominated menu.

As a final remark, notice that the epistemic model specified by A.1 is one of full awareness. It is this observation that qualifies the Krepsian interpretation of flexibility as arising from unforeseen contingencies. There is a bijective relationship between a class of models exhibiting full awareness and with preference relations adhering to the Krepsian paradigm. To be clear, this is not to say a preference for flexibility can*not* arise from the anticipation of unforeseen contingencies, but rather, that a preference for flexibility is not proof of the anticipation of unforeseen contingencies. Every Krepsian preference is wholly consistent with an epistemic model in which the DM is both fully aware and fully rational (i.e., obeys the ***S5*** axioms). This caveat stands in contrast to the results of Section 3: the non-existence of an acceptable contingent plan is *not* consistent with any epistemic model of full awareness.

Appendix B. The AWARE axiom system

Two formulae φ and ψ are called *bound alphabetic variants* of one another if φ and ψ differ only because where φ has well formed sub-formulae of the form $\forall a\zeta$ where ψ has $\forall b\eta$ and ζ has free occurrences of a in exactly the same places as η has free occurrences of b. $\varphi[a/b]$ denotes the formula created first by taking a bound alphabetic variant of φ with no bound occurrences of b, and then replacing every free a with b.

The following axiom schemata,

[Prop] All substitution instances of valid formulae in propositional logic.
[E1] $\forall a E(a)$
[E2] $\forall a\varphi \implies (Eb \implies \varphi[a/b])$
[E3] $\forall a(\varphi \implies \psi) \implies (\forall x\varphi \implies \forall x\psi)$
[E4] $\varphi \implies \forall x\varphi$, provided x is not free in φ
[L] $L_t(\varphi \implies \psi) \implies (L_t\varphi \implies L_t\psi)$
[A0] $K_t\varphi \iff (L_t\varphi \wedge A_t\varphi)$
[A1] $A_t\varphi$, provided φ is a sentence.
[A2] $A_t\varphi \wedge A_t\psi \implies A_t(\varphi \wedge \psi)$
[A3] $A_t\varphi \implies A_t\psi$, provided φ and ψ share the same free variables.

and inference rules,

[MP] From φ and $(\varphi \implies \psi)$ infer ψ.
[LN] From φ infer $L_t\varphi$.
[UG] From φ infer $\forall a\varphi$.
[UGL] From

$$\varphi_1 \implies L_t(\varphi_2 \implies \ldots \implies L_t(\varphi_h \implies L_t\psi)\ldots),$$

where $h \geq 0$, infer

$$\varphi_1 \implies L_t(\varphi_2 \implies \ldots \implies L_t(\varphi_h \implies L_t\forall a\psi)\ldots),$$

provided that a is not free in $\varphi_1 \ldots \varphi_h$.

regarding the language $\mathcal{L}$ were first proposed by Board and Sau Chung (2011). Refer to the above axiom system as ***AWARE***.

Appendix C. Supporting results

Lemma 1. *Let $\Gamma \subset \mathcal{L}$ be a maximally consistent set of formulae containing* Cmp ∪ Trv. *For each $a \in \mathcal{X}$ define $l(a) = \{b \in \mathcal{X} \mid (a \succcurlyeq b) \in \Gamma\}$. Then for any $a, b \in \mathcal{X}$ either $l(a) \subseteq l(b)$ or $l(b) \subseteq l(a)$.*

Proof. Assume $(a \succcurlyeq b) \in \Gamma$. So, let $c \in l(b)$. By definition, $(b \succcurlyeq c) \in \Gamma$, and therefore so to is $(a \succcurlyeq b) \wedge (b \succcurlyeq c)$. Then, since every instance of Trv $\in \Gamma$, $((a \succcurlyeq b) \wedge (b \succcurlyeq c) \implies (a \succcurlyeq c)) \in \Gamma$, so by MP, $(a \succcurlyeq c) \in \Gamma$, implying $c \in l(a)$ and so $l(b) \subseteq l(a)$. Conversely, assume $(a \succcurlyeq b) \notin \Gamma$. Since Γ is maximally consistent, $\neg(a \succcurlyeq b) \in \Gamma$. Moreover, since every instance of Cmp $\in \Gamma$, $((a \succcurlyeq b) \vee (b \succcurlyeq a)) \in \Gamma$. Hence, $(b \succcurlyeq a) \in \Gamma$. Appealing to the argument above, we have that $l(a) \subseteq l(b)$. □

Lemma 2. *Fix a model $M \in \mathcal{M}^\star$ with a finite state space, S, and a assignment μ. For each s set $s^{L_1} = \{\lambda \in \mathcal{L} | (M, s) \models_\mu L_1\lambda\}$. Let $\sim_{L_t}$ be the equivalence relation on S defined by $s \sim_{L_1} s'$ if $s^{L_1} = s'^{L_1}$. Let $S/{\sim_{L_t\mu}}$ denote the resulting quotient space of S, with elements $[s]$. Then, there exists a set $\Phi \subset \mathcal{L}$ such that*

$$h : [s] \mapsto \{\varphi \in \Phi \mid (M, s) \models_\mu L_1\varphi\} \tag{C.1}$$

defines a bijection between $S/{\sim_{L_t}}$ and Φ.

Proof. Let M be such a model. Enumerate the elements of $S/{\sim_{L_1}}$. The proof is by induction on the number of elements in $S/{\sim_{L_1}}$. If $S/{\sim_{L_1}}$ contains a single element, any single tautological statement provides such a set, Φ.

Assume the result holds for n, with the corresponding set $\Phi_n = \{\varphi_{1,n} \ldots \varphi_{n,n}\}$, and let $S/{\sim_{L_1}}$ contain $n+1$ elements. By definition of $S/{\sim_{L_1}}$, it must be that for each $[s_i]$, $i \leq n$, there exists some statement λ_i, such that,

$$\lambda_i \in s_i^{L_1} \setminus s_{n+1}^{L_1}, \text{ or,} \tag{C.2}$$

$$\lambda_i \in s_{n+1}^{L_1} \setminus s_i^{L_1}. \tag{C.3}$$

So, for each $i \leq n$, define $\psi_i = \neg L_t\lambda_i$ if (C.2) holds, and $\psi_i = L_t\lambda_i$ if (C.3) holds. Define,

$$\varphi_{n+1,n+1} = \bigwedge_{i\leq n} \psi_i,$$

$$\varphi_{i,n+1} = \varphi_{i,n} \wedge \neg L_1\varphi_{n+1,n+1},$$

for $i \leq n$. We claim $\Phi_{n+1} = \{\varphi_{1,n+1} \dots \varphi_{n+1,n+1}\}$ defines a bijection between $S/{\sim_{L_t}}$ and Φ defined by (C.1).

Towards surjectivity, let $s \in [s_i]$. We claim $(M,s) \models_\mu L_1\varphi_{i,n+1}$. First, if $k = n+1$ then by construction $(M,s) \models_\mu \psi_i$ for each i, and further each ψ is of the form $L_1\lambda_i$ or $\neg L_1\lambda_i$. Either way, $(M,s) \models_\mu L_1\psi_i$ by the fact that R_1 is an equivalence relation. The claim follows since $\models$ is closed under conjunction. If $k \neq n+1$ then $(M,s) \models_\mu L_1\varphi_{i,n}$, by the inductive hypothesis. Further, by construction $(M,s) \models_\mu \neg\psi_i$ and hence $(M,s) \models_\mu \neg\varphi_{n+1,n+1}$. By the reflexivity of R_1, this implies $(M,s) \models_\mu \neg L_1\varphi_{n+1,n+1}$.

To show injectivity, we will show that $L_1\varphi_{n+1,n+1} \wedge L_1\varphi_{i,n+1}$ is unsatisfiable for all i. This will prove the claim since, appealing to the inductive construction, it implies $L_1\varphi_{i,n+1} \wedge L_1\varphi_{j,n+1}$ is never satisfiable. Indeed, $L_1\varphi_{i,n+1} \implies L_1\neg L_1\varphi_{n+1,n+1} \implies \neg L_1\varphi_{n+1,n+1}$ (where the second implication comes from T). □

Appendix D. Soundness and completeness results

Proof of Proposition 2.1. The proof of soundness of the $\boldsymbol{AWARE}$ axioms can be found in Board and Sau Chung (2011). For the remaining axioms, let $M \in \mathcal{M}$. Let μ be arbitrary. Assume $(M,s) \models_\mu \neg(a \succcurlyeq b)$. By definition $u_s(\mu(a)) \not\geq u_s(\mu(b))$, and therefore $u_s(\mu(b)) > u_s(\mu(a))$, so, $(M,s) \models_\mu (b \succcurlyeq a)$: since μ was arbitrary, Cmp is valid in $\mathcal{M}$. So Cmp is sound with respect to $\mathcal{M}$. Assume $(M,s) \models_\mu (a \succcurlyeq b) \wedge (b \succcurlyeq c)$. By definition, this implies $u_s(\mu(a)) \geq u_s(\mu(b))$ and $u_s(\mu(b)) \geq u_s(\mu(c))$; by the transitivity of $\geq$ this implies $u_s(\mu(a)) \geq u_s(\mu(c))$, so, $(M,s) \models_\mu (a \succcurlyeq c)$. Since μ was arbitrary, Trv is sound with respect to $\mathcal{M}$. Now let $M \in \mathcal{M}^{BND}$, and $s \in S$. Then $u_s|_{X_s}$ attains its supremum, say at $\bar{x}$. Then let μ be such that $\mu(a) = \bar{x}$. Let $\mu' \sim_b \mu$. Then $u_s(\mu'(a)) = u_s(\bar{x}) \geq u_s(\mu'(b))$ which implies $(M,s) \models_{\mu'} a \succcurlyeq b$. Since μ' was an arbitrary b variant of μ, we have $(M,s) \models_\mu \forall b(a \succcurlyeq b)$, and therefore, $(M,s) \models_\mu \exists a\forall b(a \succcurlyeq b)$. Bnd is valid in $\mathcal{M}$; Bnd is sound with respect to $\mathcal{M}^{BND}$.

To show completeness, we construct a canonical structure, M^c. As usual the sets of states will be maximally consistent subsets of $\mathcal{L}$, which obey the '$L\forall$-property.' A set of formulae, Γ, satisfies the $L\forall$-property if:

1. For every formula φ and variable a, there is some variables b such that the formula $E(b) \wedge (\varphi[a/b] \implies \forall a\varphi)$ is in Γ;
2. For any formulas $\varphi_1 \dots \varphi_h$ $(h \geq 0)$ and ψ, and every variable a that is not free in $\varphi_1 \dots \varphi_h$, there is some variable b such that the formula $L_t(\varphi_1 \implies \dots \implies L_t(\varphi_h \implies L_t(E(b) \implies \psi[a/b]))\dots) \implies L_t(\varphi_1 \implies \dots \implies L_t(\varphi_h \implies L_t(\varphi_h \implies \forall a\psi)\dots)$ is in Γ.

To this end, let S^c denote the set of all maximally $\boldsymbol{AWARE} \cup$ Cmp $\cup$ Trv consistent sets of formulae in $\mathcal{L}$ that satisfy the $L\forall$-property. Let n denote an injection from $\mathcal{X}$ to $\mathbb{N}$. For each s, define $s^{L_t} = \{\varphi \in \mathcal{L} \mid L_t\varphi \in s\}$ and $s^{A_t} = \{a \in \mathcal{X} \mid A_t E(a) \in s\}$, and for each $a \in \mathcal{X}$, $l(s,a) = \{b \in \mathcal{X} \mid (a \succcurlyeq b) \in s\}$.

Define the canonical model $M^{c,n} = \langle S^c, X^c, \{X^c_s\}_{s\in S}, \mathcal{V}^c, \{R^c_t\}_{t\leq T}, \{u^{c,n}_s\}_{s\in S^c}, \{\mathcal{A}^c_t\}_{t\leq T}\rangle$, where S^c is defined as above, $X^c = \mathcal{X}$, $X^c_s = \{a \in \mathcal{X} \mid E(a) \in s\}$, $\mathcal{V}^c$ is defined by $(a_1 \dots a_n, s) \in \mathcal{V}^c(\alpha)$ if and only if $\alpha a_1 \dots a_n \in s$, R^c_t is defined by $sR^c_t s'$ if and only if $s^{L_t} \subseteq s'$, $u^c_s = a \mapsto \sum_{b\in l(s,a)} \frac{1}{2^{n(b)}}$, and $\mathcal{A}^c_t(s) = s^{A_t}$ for all t. Finally, define the conical assignment as the identity, $\mu^c : a \mapsto a$.

Board and Sau Chung (2011) show that this canonically structure satisfies all $\boldsymbol{AWARE} \cup$ Cmp $\cup$ Trv-consistent formulae in the fragment of $\mathcal{L}$ not containing $\succcurlyeq$. Thus, we are tasked with showing that $(M^c, s) \models_{\mu^c} (a \succcurlyeq b)$ iff $(a \succcurlyeq b) \in s$. Indeed, $(M^c, s) \models_{\mu^c} (a \succcurlyeq b)$ iff $u^{c,n}(a) \geq u^{c,n}(b)$ iff $\sum_{c\in l(s,a)} \frac{1}{2^{n(c)}} \geq \sum_{c\in l(s,b)} \frac{1}{2^{n(c)}}$, which by Lemma 1 (and the fact that $\frac{1}{2^{n(c)}} > 0$ for all $c \in \mathcal{X}$) is iff $b \in l(s,a)$ iff $(a \succcurlyeq b) \in s$.

Finally, we show that the same canonical construction, with the inclusion of Bnd, yields utility functions and domains such that $u_s|_{X_s}$ attains its supremum for all $s \in S$. Indeed, since each $s \in S$ contains Bnd. Therefore $\forall a\neg\forall b(a \succcurlyeq b) \notin s$. Since s has the $L\forall$-property there is some formula $E(a') \wedge (\neg\forall b(a' \succcurlyeq b) \implies \forall a\neg\forall b(a \succcurlyeq b))$. By the maximal consistency $E(a') \in s$—hence $a' \in X_s$—and $\forall b(a' \succcurlyeq b) \in s$. We claim that $u_s(a') = \sup_{b\in X_s} u_s(b)$. Indeed, let $b' \in X_s$. Then both $E(b') \in s$ and $\forall b(a' \succcurlyeq b) \in s$, so by E2, $(a' \succcurlyeq b') \in s$. Finally, Lemma 1 implies that $l(s,b') \subseteq l(s,a')$ and therefore that $u_s(a) \geq u_s(b)$. □

Proof of Proposition 2.2. Let $M \in \mathcal{M}^{LRN}$, and μ be arbitrary. Assume $(M,s) \models_\mu L_0\varphi$. So, $(M,s') \models \varphi$ for all $s' \in R_0(s)$. In particular, $(M,s'') \models_\mu \varphi$ for all $s'' \in R_1 \subseteq R_0$. By definition, $(M,s) \models_\mu L_1\varphi$. Now assume $(M,s) \models_\mu A_0\varphi$, then for all free a in φ, $\mu(a) \in \mathcal{A}_0(s) \subset \mathcal{A}_1(s)$. Therefore $(M,s) \models_\mu A_1\varphi$. Finally, assume $(M,s) \models_\mu E(a)$, so that $\mu(a) \in X_s$, where X_s is constant on $R_1(s)$. Then $(M,s) \models_\mu L_1E(a)$. We have that LU, AU, and LE are valid in $\mathcal{M}^{LRN}$.

Towards completeness, we construct the canonical structure, $M^{c:LRN}$, as usual. The result follows if $R_1(s) \subseteq \mathcal{A}_0(s)$ and $A_0(s) \subseteq \mathcal{A}_1(s)$ is true for all $S^{c:LRN}$. So fix some $s \in S^{c:LRN}$, and let s' be such that $sR_1^{c:LRN}s'$. By definition this implies $s^{L_1} \subseteq s'$. Now, let $\varphi \in s^{L_0}$: by definition $L_0\varphi \in s$. Since s contains every instance of LU, $(L_0\varphi \implies L_1\varphi) \in s$, and consequently, $L_1\varphi \in s$. By definition $\varphi \in s^{L_1}$ Since φ was arbitrary, $s^{L_0} \subseteq s^{L_1} \subseteq s'$, implying $sR_0^{c:f}s'$, as desired. Now let $a \in s^{A_0}$ so by definition $A_0E(a) \in s$. Since s contains every instance of AU, $(A_0E(a) \implies A_0E(a)) \in s$, and consequently, $A_1E(a) \in s$, so $a \in s^{A_1}$, as desired. Finally, let $a \in X_s^c$, so that $E(a) \in s$. Since s contains every instance of LE, $L_1E(a) \in s$, or, equivalently, $E(a) \in s^{L_1}$. Now let $s' \in R_1(s)$; be definition $E(a) \in s^{L_1} \subseteq s'$. Therefore $a \in X_{s'}^c$ as desired. □

Appendix E. Omitted proofs

Proof of Theorem 3.4. Let μ be surjective. As in Lemma 2, set $[s] = \{s' \in S | s' \sim_{L_1} s\}$. By the conclusion of that lemma, there exists a set $\Phi \subset \mathcal{L}$ be a finite set of formulae such that

$$h : [s] \mapsto \{\varphi \in \Phi \mid (M, s) \models_\mu L_1\varphi\}$$

is a bijection between $S/\sim_{L_t}$ and Φ.

For each state s, set $x(s) \in \text{argmax}_{x \in X_s} u_s(x)$, guaranteed to exists since $M \in M^{BND}$. For each $\varphi \in \Phi$, define $c(\varphi)$ to be any element of $\{a \in \mathcal{X} \mid \mu(a)=x(s), s \in h^{-1}(\varphi)\}$ (which is non-empty by the surjectivity of h and μ). Further, by the injectivity of h (associating $\{\varphi\} \cong \varphi$), we have that for each s, $(M, s) \models_\mu L_1\varphi$ for unique $\varphi \in \Phi$, denoted by the abuse of notation, $h(s)$. Thus, $(c\colon \Phi \to \mathcal{X})_\mu$ is indeed a contingent plan.

Also, notice that for each s, X_s must be constant on $[s]$ by LE. Moreover, for each s, $s \in [s] = h^{-1}(h(s))$, implying that $(M, s) \models_\mu L_1h(s) \wedge E(c(h(s)))$.

It remains to show that it is acceptable. In light of the above observation regarding existance, it suffices to show there no state s' such that, $(M, s') \models_\mu L_1h(s') \wedge \exists a L_1(a \succ c(h(s')))$. Therefore, for some $\mu' \sim_a \mu$,

$$(M, s') \models_{\mu'} L_1(a \succ c(h(s'))). \tag{E.1}$$

Further, by the reflexivity of R_1, $(M, s') \models_{\mu'} (a \succ c(h(s')))$. i.e., $c(h(s'))$ is not $\succcurlyeq_{s'}$ maximal.

By the construction of c, there must be some other state, s'', such that $c(h(s')) = x(s'')$ and $s'' \in [s'] = h^{-1}(h(s'))$. Therefore, $s'^{L_1} = s''^{L_1}$. So, by (E.1), $(M, s'') \models_{\mu'} L_1(a \succ c(h(s')))$, a contradiction, via the reflexivity of R_1, to the $u_{s''}$ maximality of $c(h(s'))$. □

Proof of Theorem 3.6. Assume this was not true for some $M \in \mathcal{M}^\star$ and $s \in S$. This assumption entails that for all $s' \in R_s(0)$, $(M, s) \models \forall a A_0E(a)$, otherwise put, that $\mathcal{A}_0(s') = X_{s'}$ for all $s \in R_0(s)$. Consider the alternative model M' which leaved all aspects of M unchanged, except $\mathcal{A}_0(s) = \mathcal{A}_1(s) = X_s$ for all $s \in S$. Then M' admits an (everywhere) acceptable contingent plan $(c'\colon \Phi \to \mathcal{X})_\mu$. We claim that c' is acceptable in M at s. Indeed, we have not altered the truth value of any formula in any state $s' \in R_0(s)$: $(M', s) \models_\mu \psi$ if and only if $(M, s) \models_\mu \psi$ for all $\psi \in \mathcal{L}$. Applying this observation to the definition of acceptability, (3.2), concludes the proof. □

Proof of Theorem A.1. $[2 \implies 1]$. We will prove that a Krepsian representation exists: a strictly increasing aggregator $\Gamma : \mathbb{R}^{|R_0(s)|} \to \mathbb{R}$ such that

$$m \geq m' \iff \Gamma(\{\max_{x \in m} u_{s'}(x)\}_{s' \in R_0(s)}) \geq \Gamma(\{\max_{x \in m'} u_{s'}(x)\}_{s' \in R_0(s)}) \tag{E.2}$$

represents $\geq$. For each m and s, let $\bar{m}(s) = \text{argmax} u_s$. Since $\geq$ is complete and transitive, there exists some $V : 2^X \to \mathbb{R}$ that represents it. Define $\xi(m) \equiv \{\max_{x \in m} u_{s'}(x)\}_{s' \in R_0(s)}$. So let Γ be any strictly increasing extension of the map: $\xi(m) \mapsto V(m)$. It remains to show that Γ is well defined. Indeed, if $\xi(m) = \xi(m')$, then for all $s' \in R_0(s)$, we have $u_{s'}(\bar{m}(s')) = u_{s'}(\bar{m}'(s'))$, implying (since $u_{s'}$ is constant over $R_0(s)$), $(M, s') \models_\mu L_1 \bigwedge_{a \in \mu^1(m)} (\mu^1(\mu^{-1}(\bar{m}(s')) \succcurlyeq a) \wedge L_1 \bigwedge_{a \in \mu^{-1}(m')} (\mu^{-1}(\bar{m}'(s')) \succcurlyeq a)$. It follows that m s-dominates m' and that m' s-dominates m, $V(m) = V(m')$. Now if $\xi(m) > \xi(m')$ (i.e., component wise inequality with some strict), we have likewise have for all $u_{s'}(\bar{m}(s')) \geq u_{s'}(\bar{m}'(s'))$, (with some strict preference) implying (since $u_{s'}$ is constant over $R_0(s)$), $(M, s') \models L_1 \bigwedge_{a \in \mu^{-1}(m)} (\mu^{-1}(\bar{m}(s')) \succcurlyeq a)$, and for at least one state $s'' \in R_0(s)$, $(M, s'') \models \neg L_1 \bigwedge_{a \in \mu^{-1}(m')} (\mu^{-1}(\bar{m}(s'')) \succcurlyeq a)$. It follows that m strictly s-dominates m', so, $V(m) > V(m')$, as desired.

$[1 \implies 2]$. We will construct the model that generates $\geq$. So let $\geq$ satisfy the axioms of Kreps (1979), and so, the representation therein holds, (i.e., of the form of (E.2), with an arbitrary state space, Ω). It is easy to check the following model suffices, $S \cong \Omega$, $\mathcal{V}$ can be arbitrary, $R_0 = S^2$, $R_1 = \bigcup_{s \in S}(s, s)$, and for each s, let $u_{s_\omega} = u_\omega$. □

References

Auster, Sarah, 2013. Asymmetric awareness and moral hazard. Games Econ. Behav. 82, 503–521.
Blume, Lawrence, Easley, David, Halpern, Joseph Y., 2009. Constructive Decision Theory. Working paper.
Board, Oliver J., Sau Chung, Kim, 2011. Object-Based Unawareness: Axioms. Working paper.
Board, Oliver J., Sau Chung, Kim, Schipper, Burkhard C., 2011. Two models of unawareness: comparing the object-based and the subjective-state-space approaches. Synthese 179 (Suppl. 1), 13–34.
Dekel, Eddie, Lipman, Barton L., Rustichini, Aldo, 1998. Standard state-space models preclude unawareness. Econometrica 66 (1), 159–173.
Dillenberger, David, Lleras, Juan Sebastián, Sadowski, Philipp, Takeoka, Norio, 2014. A theory of subjective learning. J. Econ. Theory, 1–31.
Dillenberger, David, Vijay Krishna, R., Sadowski, Philipp, 2015. Dynamic Rational Inattention. Working paper.

Ergin, Haluk, Sarver, Todd, 2010. A unique costly contemplation representation. Econometrica 78 (4), 1285–1339.
Evren, Özgür, Ok, Efe A., 2011. On the multi-utility representation of preference relations. J. Math. Econ. 47 (4–5), 554–563.
Fagin, Ronald, Halpern, Joseph Y., 1988. Belief, awarness and limited reasoning. Artif. Intell. 34, 491–501.
Fagin, Ronald, Halpern, Joseph Y., Moses, Yoram, Vardi, Moshe Y., 1995. Reasoning About Knowledge, vol. 1. MIT Press.
Filiz-Ozbay, Emel, 2012. Incorporating unawareness into contract theory. Games Econ. Behav. 76 (1), 181–194.
Grant, Simon, Quiggin, John, 2012. Induction reasoning about unawareness. Econ. Theory, 1–39.
Grant, Simon, Quiggin, John, 2014. A preference model for choice subject to surprise. Theory Dec. 79 (2), 167–180.
Halpern, Joseph Y., Rêgo, Leandro C., 2009. Reasoning about knowledge of unawareness. Games Econ. Behav. 67 (2), 503–525.
Halpern, Joseph Y., Rêgo, Leandro C., 2013. Reasoning about knowledge of unawareness revisited. Math. Soc. Sci. 65 (2), 73–84.
Heifetz, Aviad, Meier, Martin, Schipper, Burkhard C., 2006. Interactive unawareness. J. Econ. Theory 130 (1), 78–94.
Heifetz, Aviad, Meier, Martin, Schipper, Burkhard C., 2008. A canonical model for interactive unawareness. Games Econ. Behav. 62, 304–324.
Hughes, G.E., Cresswell, M.J., 1996. A New Introduction to Modal Logic, vol. 107. Routledge.
Karni, Edi, Vierø, Marie-Louise, 2016. Awareness of unawareness: a theory of decision making in the face of ignorance. J. Econ. Theory 168, 301–328.
Kochov, Asen, 2015. A Behavioral Definition of Unforeseen Contingencies. Working paper.
Kreps, David M., 1979. A representation theorem for "preference for flexibility". Econometrica 47 (3), 565–577.
Li, Jing, 2009. Information structures with unawareness. J. Econ. Theory 144 (3), 977–993.
Marschak, Jacob, DeGroot, Morris H., Becker, Gordon M., 1964. Measuring utility by a single response sequential method. Syst. Res. Behav. Sci. 9 (3), 226–232.
Maskin, Eric, Tirole, Jean, 1999. Unforeseen contingencies and incomplete contracts. Rev. Econ. Stud. 66, 83–114.
Minardi, Stefania, Savochkin, Andrei, 2017. Subjective Contingencies and Limited Bayesian Updating (May 1, 2017). HEC Paris Research Paper No. ECO-SCD-2017-1203. Available at SSRN: https://ssrn.com/abstract=2962260.
Modica, Salvatore, Rustichini, Aldo, 1994. Awareness and partitional information structures. Theory Dec. 37 (1), 107–124.
Modica, Salvatore, Rustichini, Aldo, 1999. Unawareness and partitional information structures. Games Econ. Behav. 27 (2), 265–298.
Morris, Stephen, 1996. The logic of belief and belief change: a decision theoretic approach. J. Econ. Theory 69 (35), 1–23.
Piermont, Evan, Teper, Roee, 2017. Plans of Action. Working paper.
Piermont, Evan, Takeoka, Norio, Teper, Roee, 2016. Learning the Krepsian state: exploration through consumption. Games Econ. Behav. 100, 69–94.
Riella, Gil, 2013. Preference for flexibility and dynamic consistency. J. Econ. Theory 148 (6), 2467–2482.
Schipper, Burkhard C., 2013. Awareness-dependent subjective expected utility. Int. J. Game Theory 42 (3), 725–753.
Schipper, Burkhard C., 2014. Preference-based unawareness. Math. Soc. Sci. 70, 34–41.