REVERSE BAYESIANISM: A COMMENT

CHRISTOPHER P. CHAMBERS AND TAKASHI HAYASHI

ABSTRACT. Karni and Vierø (2013) present an interesting theory of decisions in the presence of new actions and consequences. We establish results on the observable implications of the model. When introducing new consequences, arbitrary preference reversals over feasible actions are permitted. This occurs even if an outside observer can uniquely pin down the decision maker's ordinal ranking over consequences.

1. INTRODUCTION

We intend this note to make a simple point about positive decision theory. Throughout, we take the perspective of an outside observer, or experimenter, wishing to test a decision theory. We understand a positive aspect of the theory to be any aspect relating to (potentially observable) data. Thus, a primitive notion for any such exercise should be the notion of observability. In order to pin down the theory being tested, we understand that the notion of observable data and its interpretation within the theory is fixed. We focus on Karni and Vierø (2013) (hereafter KV), though the issues here are somewhat broader.

Economic theories postulate a relationship between observable (data) and unobservable (theoretical) concepts. Different notions of data result in different testable hypotheses. For example, one interpretation of Savage (1972) imagines that the outside observer is only able to access potential data on choices between acts. In such an exercise, probability itself is just part of a representation, and has little or no meaning outside of the predictions it makes for choice. But nothing precludes an outside observer from considering other types of data. For example, data on historical frequencies might inform an outside observer toward an objective concept of probability, or questionnaire data may be useful in interpreting subjective probability. So long as observable data can be meaningfully interpreted in the context of a model, the model speaks to the relationships between these data.

In this note, we focus on two types of data: choice data, on the one hand, and measurement data, on the other. We hope to be explicit about the (necessarily informal) distinction between the two.

We are grateful to two referees for extremely useful comments and corrections. Also errors remain our own.

2. The model

Let us now set up the mathematical primitives of the model of Karni and Vierø (2013) (hereafter KV).

- A finite set of feasible actions, F, with typical elements $f, g \in F$.
- A finite set of consequences C, with typical elements $c, d \in C$.

We emphasize that it is the nature of a decision that *one and only one element of F must be chosen*. Now, given F and C, an individual is endowed with an observable binary relation (a preference) $\succeq$ over F. The analysis of KV is intended to understand how this preference should adapt to new actions and consequences. The paper is normative, and has a clear message about how an individual should respond to the arrival of information in the form of unanticipated consequences or feasible actions. We illustrate the model and the point we wish to make with the following example, motivated by the treatment effects literature.

Example 1. A college student wants to get a PhD in economics. She needs to determine which school to attend. Suppose there are two schools, say a and b; hence $F = \{a, b\}$. She talks to her advisor and understands that there are two possible consequences of her choices: either getting a job at a research university, or working in a teaching college. Let these consequences be denoted R and T, respectively, so that $C = \{R, T\}$.

One type of data, which we refer to as *normative* or *introspective* data, consists of data accessible to the decision maker herself, but not necessarily data which are revealed by choice behavior.[1] She could ask herself whether she prefers R or T outright. She could also ask herself how likely she would view the outcome R after choosing a, and how likely she would view R after choosing b. She could then synthesize these judgments into an overall judgment about the better choice for her, a or b.

From the perspective of an outside observer, however, these considerations are invisible. Instead, the outside observer sees the decision of a or b. It cannot be inferred from this choice alone whether the decision maker thinks R is better than T, or how likely the outcomes are conditional on the two actions (though we will later show that it does not matter whether the outside observer can observe the preference between R and T).

Now, suppose the decision maker realizes that, as a third consequence, she might end up with a job in the private sector; so that $C = \{R, T, P\}$. KV suggest that the decision maker "update" her beliefs in a natural way upon this realization. Her ranking between R and T should not change, and conditional on P being understood as "impossible" no matter what her action, her beliefs over the relevant uncertainty

[1] The idea that introspective data is itself a type of data is not novel; see Kreps (1990).

should remain the same as in the initial situation. But the decision maker is now permitted to attribute some probability to the outcome P conditional on the two actions, and she also should evaluate how she ranks P with respect to R and T.

As we have suggested, these considerations may be invisible to the outside observer. Our point is that, no matter what the decision maker chose when there were two outcomes (R and T), her choice can be reversed upon the realization that P is also a possibility. And in fact, what we show is broader: the example focuses on two schools, a and b, but the reversal happens for any finite list of schools. And the reversal can be quite arbitrary. Further, this reversal can occur even if the outside observer (say, the advisor) knows the college student's ranking over pure consequences (over R, T, and P).[2] Thus there is no systematic prediction about behavior which can be made upon the introduction of the new consequence.

3. States of the world

In our understanding of the standard conception of the state space, each state is usually understood as an exhaustive answer to a list of questions, which do not reference consequences or actions.[3] Importantly, each of the questions can be answered, and the asking of any question does not have any effect on the answer of any other question. For example, a natural state space would be the outcome of a coin flip, together with the temperature: $S = \{H, T\} \times \mathbb{R}$. Flipping the coin does not change the temperature, and measuring the temperature does not determine the outcome of the coin flip.

KV suggest viewing the problem differently. One of KV's innovations is to endogenize the states of the world, via a classical construction. States are defined implicitly via the evaluation map; that is, the state space is $\Omega(C, F) = C^F$.[4] Each state therefore is a mapping specifying the consequence stemming from choosing each act. For example, if the state s is such that $s(f) = c$, then in this state, when choosing f, c realizes. This conception of the state space permits the decision maker to entertain the possibility that any profile of consequences could be realized by choosing different actions.

---

[2]So long as R and T are not indifferent.

[3]In principle, there are potential difficulties with such a construction. If the number of questions is infinite; then it is reasonable to assume that only a finite number of them could ever be answered. A more natural primitive in this case would be an algebra of events; this is motivated by the famous Stone Representation Theorem, of Stone (1936). A Boolean algebra is a set of propositions, closed under negation, finite disjunction, and finite conjunction, together with the vacuously true proposition. Hence a Boolean algebra allows for the possibility that only a finite number of questions could be asked (since disjunctions and conjunctions are finite). Stone's Theorem asserts that every Boolean algebra is isomorphic to an algebra of events on some state space.

[4]This notation does not appear in KV; however, they use the notation $S(C, F)$ and terminology *feasible states* to indicate non-null states.

4. Conceivable Actions

A secondary object is considered in KV, which are the *conceivable actions*. The set of *conceivable actions* is the set of functions $f : \Omega(C, F) \to \Delta(C)$, where $\Delta(C)$ is the set of lotteries on C. We denote this set by $\hat{F}(C, F)$. Importantly, the conceivable actions depend on both the feasible actions and the consequences. When either of these two sets changes, so does the set of conceivable actions.

4.1. **Feasible Actions as Measurement Devices.** One interpretation of KV imagines that the feasible actions F themselves are choices of measurement; this seems to be a suggestion of Karni (2015). Let us consider the case of one feasible action, say f. Suppose, as is suggested in Karni (2015) that the set of consequences can be written as $C = O \cup X$, where O is a set of observable outcomes of action f. On the other hand, elements of X could not possibly be the outcome of action f. Karni (2015) gives the example where f is "running a horse race," elements of O are the possible configurations according to which the horses may cross the finish line, and X are monetary payments.

Since there is only one feasible action, $\Omega(O \cup X, F)$ can be identified with C; so that C are the states of the world. But, it is absurd to suppose the outcome of a horse race is a monetary payment. Instead, any "reasonable" decision maker would attribute a zero probability to states in X. On the other hand, it is eminently reasonable that one could offer monetary payments (elements of X) *conditional* on the outcome of the horse race. So, in this setting, conceivable actions make perfect sense, and can be identified with actions of the form $g : O \to \Delta(X)$. Thus, the primitive effectively coincides with that of Anscombe and Aumann (1963). And the empirical content of choice over conceivable acts is determined by the axioms of Anscombe and Aumann (1963).

There is question as to how the choice over these conceivable actions changes with the introduction of new outcomes. For example, suppose it is realized that none of the horses may finish the race, so that $O \subseteq \hat{O}$, where $\hat{O}$ is the list of horses *together* with the outcome that none of them finishes. The set of conceivable actions $\hat{F}(f, X \cup \hat{O})$ would then be identified with maps $g : \hat{O} \to \Delta(X)$, and the framework of KV would state that, conditional on all horses finishing the race, preferences coincide with the original preference.[5] On the other hand, when introducing new possible payments, so that $X \subseteq \hat{X}$, we have a new set of conceivable actions, $g : O \to \hat{X}$. The restriction their model would place in this environment is that when restricted to the set of conceivable actions which only pay elements of X, the ranking should coincide with the original one.

[5] "Conditional" would here be interpreted in the usual way: any two conceivable actions g, h which pay the same in $\hat{O} \backslash O$ should be ranked only on their basis of $g|_O, h|_O$.

4.2. **Feasible Actions as Choices.** On the other hand, consider either the example of choosing schools from earlier, or the following example from KV:

"The discovery of the New World ushered in its wake a new consequence of sexual intercourse. The risk of contracting venereal diseases was well known in the Old World. Syphilis, however, was new. Its discovery expanded the universe of the Europeans."

Here, the focus is clearly on explaining the decision whether to have sexual intercourse (the feasible actions presumably being to have and not to have). It is difficult to imagine this decision as being one of measurement alone; one would have to imagine an individual having sexual intercourse purely to see whether they contract syphilis; and taking bets on this. In fact, the consequences in this case seem to be something decision makers would inherently care about; whereas the horse-race example, the outcome of the horse race mattered only inasmuch as it determined a conditional monetary payment. Thus, we imagine the framework of KV as being applied here as an aid to individuals choosing feasible actions where they would potentially care about the consequences.

Now, when feasible actions are supposed to represent choices, the realized state $s \in C^F$ is not observable (Gilboa (2009) and Karni (2015) make this point quite clearly). The reason is the previously mentioned characteristic of a decision: one and only one $f \in F$ must be chosen. So, if $c \in C$ obtains when $f \in F$ was chosen, an outside observe would learn the event: $\{s \in C^F : s(f) = c\}$. But this observer cannot see what would have happened had a different $g \in F$ been chosen. This is why Gibbard and Harper (1978) calls this construction of states "conjunction of counterfactuals." As a consequence, elements of $\hat{F}(C, F)$ are unobservable as soon as $|F| \geq 2$. To quote Gilboa (2009) (p. 116), "But the problem we encounter here is that the choices between elements of $\hat{F}(C, F)$ cannot be observable *in principle*."[6]

We do not claim that feasible actions are the only method of describing choice situations. We agree with Karni (2015) that they can be used to represent measurement devices which can be used to derive the states of the world. Rather, we only have to claim that there are situations of interest in which the set of feasible actions should readily be interpreted as the primitive choices of interest.

[6]The preceding view is controversial. Schipper (2016) argues that certain elements of $\hat{F}(C, F)$, the so-called *coherent* acts, can be offered as choices to a subject. Coherent acts are those $f^* \in \hat{F}(C, F)$ for which for each $s \in \Omega(C, F)$, there is $f \in F$ for which $f(s) = f^*(s)$. Arguably, one could interpret coherent acts such as f^* as written contracts, specifying that if $s \in \Omega(C, F)$ is realized, then $f^*(s)$ is paid. Our take on the issue is that with such a contract, in general, one could never verify the clauses written down, so what is written on the contract and the actual consequence will have no relation. As such, we would argue that the action f^* should also be included in the primitive set of feasible actions. Or, if such contracts were offered to a decision maker, we believe that she would be very unlikely to choose them.

In this note, we therefore investigate the observable restrictions on choice behavior over feasible actions implied by the reverse Bayesian model when consequences can change.

Given a feasible action $f \in F$, its induced conceivable action $f' \in \hat{F}(C, F)$ is the action for which whenever $s(f) = x$, we have $f'(s) = x$. In practice we will usually abuse notation and refer to them both as f.

KV postulate a theory of how preferences over conceivable actions change as new information is developed. The baseline theory postulates Anscombe-Aumann behavior over conceivable actions. The idea is the following. When introducing new consequences, starting from C and going to C' for which $C \subseteq C'$, new states become available (while the old ones remain admissible); so $\Omega(C, F) \subset \Omega(C', F)$. Further, new conceivable actions also become available. Hence, the preference relation is updated to a new preference relation over the new conceivable actions. The theory postulates that for (lotteries of) consequences which are available in both situations, the von Neumann-Morgenstern utility indices coincide, and that the probability associated with the $\Omega(C, F)$ environment is the Bayesian update of the probability associated with the $\Omega(C', F)$ environment.

The following proposition asserts that with this model, observable behavior over feasible actions can change quite arbitrarily as new consequences are introduced. This can happen even if the outside observer is willing to commit to strict assumptions about the nature of the decision maker's ordinal ranking over pure consequences.[7] Hence, in the example of university choice, we allow for the possibility that the outside observer "knows" that the decision maker prefers R to T to P.[8] Still, we establish that arbitrary preference reversals are admissible, and no predictions about the decision of which school to choose can be made on this basis. Knowing the choices made when C are the available consequences gives no information on how the subject would choose when $C' \backslash C$ becomes available. Under the conditions stated in Proposition 1, the theory makes no predictions about behavior when consequences change.[9]

To see this, observe that Proposition 1 claims that no matter what the initial preference over actions, after introducing new consequences, the new preference over actions can be anything. This tells us that simply knowing the initial preference over actions and learning that the set of consequences expands tells us nothing about how behavior will change.

[7]So long as she does not imagine that total indifference prevails across all "initially present" consequences.

[8]Such assumptions are often eminently reasonable. If the consequences are monetary, any reasonable model would specify that the decision maker likes more to less.

[9]Gilboa (2009), p. 115–116 already describes informally describes the lack of predictive power of this model in the static setting, when consequences do not change.

Proposition 1. *Let F, C, and C' be finite sets where $C \subset C'$ and $|C'| > |C| \geq 2$. Let $\succeq^*$ be a weak order over C' such that there exists $x^*, y^* \in C$ for which $x^* \succ^* y^*$. Let $\succeq^C$ and $\succeq^{C'}$ be weak orders over F. Then there exist $u : C' \to \mathbb{R}$, $\pi^C \in \Delta(\Omega(C,F))$, and $\pi^{C'} \in \Delta(\Omega(C',F))$ satisfying the following conditions:*

(1) *For all $x, y \in C'$, $x \succeq^* y$ if and only if $u(x) \geq u(y)$.*
(2) *For all $f, g \in F$, $f \succeq^C g$ if and only if $\sum_{s\in\Omega(C,F)} u(f(s))\pi^C(\{s\}) \geq \sum_{s\in\Omega(C,F)} u(g(s))\pi^C(\{s\})$.*
(3) *For all $f, g \in F$, $f \succeq^{C'} g$ if and only if $\sum_{s\in\Omega(C',F)} u(f(s))\pi^{C'}(\{s\}) \geq \sum_{s\in\Omega(C',F)} u(g(s))\pi^{C'}(\{s\})$.*
(4) *For all $s, s' \in \Omega(C,F)$ for which $\pi^C(\{s\}) > 0$,*

$$\frac{\pi^C(\{s'\})}{\pi^C(\{s\})} = \frac{\pi^{C'}(\{s'\})}{\pi^{C'}(\{s\})}.$$

Proof. Let $x^*, y^* \in C$ as in the statement of the theorem (that is, so that $x^* \succ^* y^*$). We consider two cases.

Case 1: There is $w^* \in C'\backslash C$ for which $x^* \succ^* w^*$.

Let U^C represent $\succeq^C$ for which $U^C(f) > 0$ for all $f \in F$, and $\sum_{f\in F} U^C(f) = 1$. Fix $x^*, y^* \in C$ for which $x^* \succ^* y^*$. Let $u(x^*) = 1$ and $u(y^*) = 0$. For all $z \in C\backslash\{x^*, y^*\}$, let $u(z)$ be chosen so that u otherwise represents $\succeq^*$.

For all $f \in F$, let $s_f \in \Omega(F,C)$ be the state for which $s_f(f) = x^*$ and for all $g \neq f$, $s_f(g) = y^*$. The following table illustrates an environment in which $F = \{f, g, h\}$, and lists the three states s_f, s_g, and s_h.

Action	s_f	s_g	s_h
f	x^*	y^*	y^*
g	y^*	x^*	y^*
h	y^*	y^*	x^*

For all $f \in F$, let $\pi^C(\{s_f\}) = U^C(f)$. All remaining states are assigned probability 0. Observe that conditions 1 and 2 are satisfied, independently of the choice of u on $C\backslash\{x^*, y^*\}$.

Now, we choose $u(w^*) < 1$ so that the relation between $u(w^*)$ and $u(y^*)$ conforms to $\succeq^*$.

Let $\epsilon > 0$ and let $U^{C'} : F \to \mathbb{R}$ be a utility representation of $\succeq^{C'}$ such that $U^{C'}(f) > \epsilon + (1-\epsilon)u(w^*)$ for all $f \in F$ and such that $\sum_{f\in F} U^{C'}(f) = \epsilon + |F|u(w^*)(1-\epsilon) + (1-\epsilon)(1-u(w^*))$.[10]

[10]To see that such objects exist, observe first that $\frac{\epsilon}{|F|} + u(w^*)(1-\epsilon) + \frac{(1-\epsilon)}{|F|}(1-u(w^*)) > \epsilon + (1-\epsilon)u(w^*)$ when $\epsilon = 0$; and therefore, by continuity, there exists $\epsilon^* > 0$ small for which the inequality holds as well. Denote $\gamma(\epsilon) = \frac{\epsilon}{|F|} + u(w^*)(1-\epsilon) + \frac{(1-\epsilon)}{|F|}(1-u(w^*))$. Now, let $v : F \to \mathbb{R}$ be any utility representation of $\succeq^{C'}$ for which $\sum_{f\in F} v(f) = 0$. Observe that for $1 > \alpha > 0$ close to 1, $U^{C'}(f) = \alpha\gamma(\epsilon^*) + (1-\alpha)v(f)$ satisfies the stated properties.

For every $s \in \Omega(F, C)$, let $\pi^{C'}(\{s\}) = \epsilon\pi^{C}(\{s\})$. For all $f \in F$, let $s_f^* \in \Omega(F, C')$ be the state defined by $s_f^*(f) = x^*$ and $s_f^*(g) = w^*$ for all $g \neq f$.

Action	s_f^*	s_g^*	s_h^*
f	x^*	w^*	w^*
g	w^*	x^*	w^*
h	w^*	w^*	x^*

For all $f \in F$, define $\pi^{C'}(\{s_f^*\}) = \frac{U^{C'}(f) - \epsilon\pi^{C}(\{s_f\}) - u(w^*)(1-\epsilon)}{1-u(w^*)}$. Define $\pi^{C'}(\{s\}) = 0$ for all remaining states.

First, observe that $\pi^{C'}(\{s\}) \geq 0$ for all $s \in \Omega(C', F)$. Now, observe that

$$\begin{aligned}
&\sum_{f\in F} \pi^{C'}(\{s_f^*\}) \\
=&\frac{\sum_{f\in F} U^{C'}(f) - \epsilon - |F|u(w^*)(1-\epsilon)}{1-u(w^*)} \\
=&\frac{\epsilon + |F|u(w^*)(1-\epsilon) + (1-\epsilon)(1-u(w^*))}{1-u(w^*)} \\
=&1-\epsilon.
\end{aligned}$$

Consequently, $\sum_{s\in\Omega(C',F)} \pi^{C'}(\{s\}) = \sum_{f\in F}[\pi^{C'}(\{s_f\}) + \pi^{C'}(\{s_f^*\})] = 1$.

Finally, let $g \in F$. Then

$$\begin{aligned}
&\sum_{s\in\Omega(C',F)} u(g(s))\pi^{C'}(\{s\}) \\
=&\pi^{C'}(\{s_g\}) + u(w^*)\sum_{f\in F}\pi^{C'}(\{s_f^*\}) + \pi^{C'}(\{s_g^*\})(1-u(w^*)) \\
=&\pi^{C'}(\{s_g\}) + u(w^*)(1-\epsilon) + U^{C'}(g) - \epsilon\pi^{C}(\{s_g\}) - u(w^*)(1-\epsilon) \\
=&U^{C'}(g).
\end{aligned}$$

So 3 is satisfied. Finally, 4 is satisfied by construction (as for all $s \in \Omega(C, F)$, $\pi^{C'}(\{s\}) = \epsilon\pi^{C}(\{s\})$).

Case 2: For all $w \in C' \backslash C$, $w \succeq^* x^*$.

In this case, consider the relation $\succeq_2^*$ on C' defined by $x \succeq_2^* y$ if and only if $y \succeq^* x$, and the relations $\succeq_2^C$ and $\succeq_2^{C'}$ on F defined by $f \succeq_2^C g$ if and only if $g \succeq^C f$, and $f \succeq_2^{C'} g$ if and only if $g \succeq_2 f$. We can complete the construction of Case 1, reversing the roles of x^* and y^*, obtaining u, π^C, and $\pi^{C'}$. The tuple $(-u, \pi^C, \pi^{C'})$ then completes the construction. □

Remark 1. In Karni and Vierø (2015), a different framework is proposed. Instead of conceivable actions which map $\Omega(C, F)$ into $\Delta(C)$, the authors there consider conceivable actions which map $\Omega(C, F)$ into C (roughly), and then study $\Delta(\hat{F}(C, F))$.

Obviously, a special case of such an object is a lottery over elements of F; thus, this theory makes predictions about choices over $\Delta(F)$. Aside from the fact that the posited relation over $\Delta(F)$ must satisfy the von-Neumann axioms no matter which set of consequences is faced, one can ask if the theory makes any further predictions. Here, we distinguish two cases.

In the first case, preference over C' is observable, but not over $\Delta(C')$. In this case, the answer is no: any such additional implications would be delivered by the cardinal structure of the functions U^C and $U^{C'}$ described in the proof of Proposition 1. However, in the proof of Proposition 1, the cardinal structure of the functions U^C and $U^{C'}$ is not pinned down, and could be chosen to be consistent with any expected utility ranking over $\Delta(F)$ (again subject to the caveat that there are $x^*, y^* \in C$ for which $x^* \succ y^*$).

On the other hand, if preference over $\Delta(C')$ were observable, we believe there would be content to the model. This is also evidenced in the proof of Proposition 1, where we cannot choose the utility index u arbitrarily. So it is likely that pinning down the cardinal structure of u (perhaps by observing risk preferences over $\Delta(C')$) would impose content.

5. Discussion

How general are these issues? There are two possible directions such a discussion can take.

First, an astute referee has pointed out that the specific issue of the relations between states, consequences, and actions may take can be framed quite broadly. The KV setup embeds all possible logical statements in which consequences could logically follow from different actions. In this sense, writing down the model as KV have done identifies some implicit assumptions made in Savage's approach. For example, Savage (1972) assumes that all relevant actions can be broken down into maps from states to consequences, and that no "causal" relationship exists between actions and consequences. Starting with a Savage-style framework, one could perform the exercise of endogenously deriving "new" states of the world, via the procedure sketched above. Some of these "newly derived" states would correspond naturally to Savage's primitive states, and some would not. In particular, Savage gives no particular method for actually choosing the state space, acts, and set of consequences. The construction in KV offers a "canonical" method of deriving such a setup, taking actions and consequences as primitive. In such a way, it reveals implicit, and potentially ad-hoc, assumptions about the relation between states, actions, and consequences in Savage's framework. For example, one possible interpretation of this exercise is that Savage relies on an implicit assumption that many of the states derived by KV are "null."

We agree that the analysis may suggest this interpretation. However, our take on this is that the underpinnings of Savage's framework, while important, are separate from the concerns of actually testing Savage's model. One does not test a "decision theory" per se, but an instance of the theory in which all interpretations of the model are pinned down. An experimenter who identifies actions with states of the world and consequences need not even announce the relationship between states, actions, and consequences to a subject. She can still test her interpretation of the theory by offering appropriate choices, and Savage's theory is well-known to be testable. With some of the caveats we discussed above, KV is not. Nevertheless, the discussion about the distinction between states and consequences, and their relationship to actions, is one well-worth having.

Second, another question of generality relates to the observability of data. Decision theoretic models which provide predictions only on complex domains of objects are ubiquitous. It is not uncommon to see decision theoretic models whose axiomatic derivation requires the use of lotteries over menus over lotteries, and so forth. Often these models seek to explain some kind of psychological phenomenon in terms of choice data, and the objects they study are in principle observable. However, little is usually understood about such a model's predictions on "typical economic data." Our feeling is that as much as possible should be done to understand a decision theoretic model's predictions on data which we would expect to observe in real-world situations.

6. Conclusion

The aim of this note is to focus attention on data, observability, and interpretation of primitives in decision-theoretic models. The standard decision-theoretic paradigm is flexible, and as the results of this note suggest, cannot be tested without pinning down the interpretation of its primitives. To construct empirical tests of such theories, one must have concrete interpretations of all of the primitives *and* agree on what it means to observe them. We suggest that with one such interpretation, featuring feasible actions as choice behavior, the KV model has little predictive power to an outside observer. As Schipper (2016) writes: "Conclusions about choice behavior may depend dramatically on the representation of the decision problem." We fully agree with this assessment, and argue that one cannot test a decision-theoretic paradigm *per se*, but must test an instance of this paradigm, with the real-world interpretation of all mathematical symbols pinned down and immutable after the experiment.[11] Thus, one cannot speak of a refutation of a theory without speaking of how the primitives are interpreted.

References

ANSCOMBE, F. J., AND R. J. AUMANN (1963): "A definition of subjective probability," *The Annals of Mathematical Statistics*, 34(1), 199–205.

GIBBARD, A., AND W. L. HARPER (1978): "Counterfactuals and two kinds of expected utility," *Foundations and Applications of Decision Theory*.

GILBOA, I. (2009): *Theory of decision under uncertainty*. Cambridge University Press.

KARNI, E. (2015): "States of nature and the nature of states," *Economics and Philosophy*, forthcoming.

KARNI, E., AND M.-L. VIERØ (2013): ""Reverse Bayesianism": A choice-based theory of growing awareness," *American Economic Review*, 103(7), 2790–2810.

——— (2015): "Awareness of unawareness: A theory of decision making in the face of ignorance," *Journal of Economic Theory*, forthcoming.

KREPS, D. (1990): *A Course in Microeconomic Theory*. Princeton University Press.

SAVAGE, L. J. (1972): *The foundations of statistics*. Dover Publications.

SCHIPPER, B. (2016): "A note on states and acts in theories of decision under uncertainty," working paper.

STONE, M. H. (1936): "The theory of representation for Boolean algebras," *Transactions of the American Mathematical Society*, 40(1), 37–111.

[11]We emphasize that Schipper (2016) does not share our view on states of the world. Rather, he argues that the state space $\Omega(C, F)$ is a practical and canonical device for representing situations of subjective uncertainty, and that theories based on it can be tested.