MULTIPLE PRIORS FOR THE OPEN-MINDED

MARTIN DUMAV, MAXWELL B. STINCHCOMBE

Abstract. A multiple-prior decision maker is *open-minded* if she can describe, as subjective uncertainty, all convex sets of distributions over payoff relevant consequences. Open-mindedness is equivalent to the ability to subjectively describe both the uniform distribution on an interval and the set of all distributions on an interval. Parameterized sets of i.i.d. distributions from classical statistics satisfy these conditions. The use of open-minded sets of priors to model decision makers allows the objective and the subjective approaches to uncertainty to inform each other and changes the implications of previously used axioms for multi-prior preferences. Subjective models with sets of priors that are *not* open-minded yield preferences only over those subjective sets of distributions that are describable. This preference incompleteness always implies the failure to rank elements in a dense class of set, and may rank so few elements that ambiguity attitudes do not affect choices between subjectively uncertain prospects.

Contents

Date: January 23, 2019
Key words and phrases. Ambiguity, decision theory, multiple priors, open-mindedness.
The first author gratefully acknowledges the support from the Ministerio Economia y Competitividad (Spain) through grants ECO2014-55953-P, ECO2017-86261-P and MDM 2014-0431.
Many thanks to Massimiliano Amarante, Svetlana Boyarchenko, Simone Cerreia-Vioglio, William Fuchs, Takashi Hayashi, Urmee Khan, Antoine Loeper, Mark Machina, Diego Moreno, Sujoy Mukerji, Marcin Peski, Vasiliki Skreta, Dale Stahl, Johannes Schneider and Ángel Hernando Veciana for help with this paper. They are blameless for the remaining deficiencies of the paper, they prevented worse.

> ... An open mind is all very well in its way, but it ought not to be so open that there is no keeping anything in or out of it." *The Note-Books of Samuel Butler*: Part 4, Page 631, J. M. Dent & Company, London.

1. Introduction

When someone knows their decision environment well enough to reliably assign probabilities to random but utility relevant outcomes arising from their choices, they are facing a *risky* problem. When they cannot reliably assign such precise probabilities, e.g. when information comes from conflicting sources of different and perhaps unknown qualities, they are facing an *ambiguous* problem. We model choices between measurable functions from a state space to consequences, and we model ambiguity, a separate state of epistemic uncertainty, as involving a *set* of priors on a state space rather than a single prior. This paper gives two results relevant to the properties and structures of sets of priors necessary for adequate modelling of individuals facing both risky and ambiguous problems.

A convex set of priors is *open-minded* if it can induce, via consequence-valued measurable functions, any closed, convex set of distributions on any compact metric space of consequences.

Theorem 1. A convex set of priors is open-minded if and only if it can induce, via a $[0,1]$-valued measurable function, the uniform distribution, and it can induce, via another $[0,1]$-valued measurable function, the set of all distributions on $[0,1]$.

Theorem 2. If a set of priors fails either of the previous two conditions, then the class of sets of distributions that it can *not* induce is dense.

Theorem 1. yields a simple canonical set of open-minded priors — the convex closure of the set of uniform distributions on the line segments $\{x\} \times [0,1]$, $x \in [0,1]$. Another class of open-minded priors are given by classical statistical models — continuously parameterized sets of i.i.d. distributions for which there exists a consistent sequence of estimators. Theorem 2. tells us that working with sets of priors that are *not* open-minded is costly. Some commonly used sets of priors can represent so small subset of sets of distributions that that different attitudes toward ambiguity make no difference to choices over the entire subjectively describable class of sets. Comparative statics analyses are crippled by the need to check ex post and make sure whether the effects the modeler aims to analyzed are indeed in the decision maker's ambit.

These results allow us to bridge the gap between models of ambiguous choices with a state space and those without a state space, and the bridge runs directly parallel to the one for models of risky choice.

1.1. Models of Risky Choice With/Without a State Space.

There are two main complementary models and one hybrid model in the study of risky choice problems. Under the rubric of "objective" risky choice, a state space is not necessary to the model.

Rather, one directly models preferences over distributions on the space of consequences (von Neumann and Morgenstern, 1944, Ch. I, §3.6).

Under the rubric of "subjective" risky choice, one assumes that decision makers can be usefully thought of as having a Kolmogorov-style model for random phenomena. These models invoke a state space in which the stochastic realization happens, and the utility relevant outcomes arise as a (measurable) function of the realization (Savage, 1954, Ch. 5). Here, decision makers are modeled as if they use a single "subjective" prior, and preference for one function over another is identified as preference for one induced distribution over another. Particularly important for this approach is the axiom of *state independence*, i.e., the assumption that the states do not matter for utility but the distributions on the consequences they induce matter.

The hybrid model also uses a state space, but preferences are now defined over the measurable functions that take values in the set of "objective" distributions over consequences. Again, the model uses a single "subjective" prior to evaluate the measurable functions, and again, preference for one measurable function over another depends only on the induced distributions over consequences (Anscombe and Aumann, 1963). The subtlety is that preferences between induced distributions over distributions, by assumption, depend only on comparisons of the average of the induced distribution, the average being taken within the space of distributions using the single subjective prior. The objective model is a special case of the hybrid approach in which the prior is a point mass so that all functions are, with probability 1, constant at some objective probability distributions.

1.1.1. *Synergies Between the Approaches.* In some applications, modeling risky phenomena with a state space is more tractable, in other applications, working without one is more tractable, and in yet others, having both approaches available allows for deeper understandings of the phenomena being analyzed. We illustrate these in turn.

(1) A state space is far more convenient for both the statement and the proof of the strong law of large numbers, and this result is crucial to the entire theory of data-driven inference and estimation necessary to make informed choices. The internal consistency of common prior models in which agents differ by type/information requires a state space for its formulation. Progress in epistemic game theory has depended on identifying an appropriately rich state space (for a survey, see Dekel et al. (2015)).
(2) The lack of a state space foregrounds the preferences over distributions on consequences, lending itself to the study of behavioral regularities and to comparative statics analyses. The regularities include: risk aversion, captured as a negative second derivatives of expected utility functions; 'prudence,' captured as the a positive third derivative of expected utility functions; 'temperance' and 'edginess,' negative fourth and positive fifth derivatives respectively (Eeckhoudt and Schlesinger, 2006; Noussair et al., 2013; Deck and Schlesinger, 2014). The comparative statics analyses concern e.g. portfolio, career, or insurance choices as the risk being faced and/or the degree of risk aversion, or prudence, or temperance, or edginess, changes.
(3) Without a state space, first order stochastic dominance of a distribution p over a distribution q is understood either as the cumulative distribution functions of p shifted weakly to the right from that of q, or, equivalently, as the p-integral of every non-decreasing function being higher than the q-integral. Either formulation has the flavor of p being definitely better than q, and with a state space having a non-atomic prior distribution, this can be made precise — there are measurable

functions f and g on the state space inducing the distributions p and q respectively, and f is, with prior probability 1, weakly better than g.

If estimation of causal structures precedes choice making, the examples of synergies between the two approaches expand. First, one has the entirety of the interplay between Bayesian and classical statistics approaches. Second, the convergence of estimators also has two mutually reinforcing formulations: the central limit theorem is a result about the weak convergence of the distribution of estimators around true values of parameters. Finally, the ability to analyze the convergence of distributions as almost everywhere convergence (see for example, Billingsley (2012, Ch. 24)) often yields much more transparent analysis.

1.1.2. *The Role of Non-Atomicity.* Savage's Postulate P6 guarantees that the subjective prior must be non-atomic, and this is crucial for the subjective approach's *domain equivalence* to the objective approach, i.e., they both model the same set of risky problems. More specifically, Skorohod's representation theorem, Skorohod (1956, Thm. 3.1.1), guarantees that if a prior is non-atomic, then every distribution on consequences can be induced by some measurable function. With a non-atomic prior, the objective and subjective approaches are interchangeable because they have the same domain of problems – the set of all possible distributions over outcomes. By contrast, if the prior has atoms, then there is a dense set of distributions on outcomes that cannot be induced using any measurable function (see below Subsection 2.1 for details). It is worth stressing that we are not arguing that one must assume that priors are non-atomic, one could decide, for very good reasons, that the appropriate model for a decision maker in a particular situation involves atomic priors. Rather, we are pointing out that the costs of *not* assuming non-atomicity involve modeling with incomplete preferences and that such incompleteness can be costly.

By contrast, following the arguments in Anscombe and Aumann (1963, p. 204), the hybrid model can work with any prior, atomic or non-atomic. This is because any distributions over outcomes are available within the model. For specificity, suppose that the prior μ puts mass 0.4 on state s_1 and 0.6 on s_2. Fix a non-extreme distribution p on $[0, M]$ and pick $p_1 \neq p_2$ such that $p = 0.4p_1 + 0.6p_2$. If $f(s_1) = p_1$ and $f(s_2) = p_2$, then $f(\mu)$ is a distribution on the distributions on $[0, M]$, and its average is p.

For ambiguous choice problems, we replace a single prior in the risky problems by a set of priors. The main result in this paper characterizes open-mindedness – the substitute for non-atomicity of a single prior that is appropriate for choice under ambiguity. It fills for ambiguous choice problems the role analogous to non-atomicity of single prior in risky problems, it makes the approaches interchangeable, providing a bridge between the subjective and objective approaches to choice.

1.2. **Models of Ambiguous Choice With/Without a State Space.** In modeling risky problems, the question "What does a choice between two measurable functions represent?" has an immediate answer. It represents a choice between the two induced distributions on the space of consequences. This is due to the state independence assumption, preferences depend not on the state, but on the induced distribution. To put it another way, in the subjective model, the state spaces serve only to model randomness.

After replacing a single prior by a set of priors, we ask the same question, "What does a choice between two measurable functions represent?" Applying state independence in analogous manner as it is applied in risky choice, the answer is that it represents a choice between the two *sets* of induced distributions on the space of consequences. We would like the objective and the subjective approaches to inform each other, and for this we want domain equivalence for ambiguous choice to hold: that is, the objective and the subjective approach covers the same domain of ambiguous choice problems.

In the objective approach to ambiguous choice, one models preferences over *sets* of distributions on the space of consequences. In the subjective, preferences-over-measurable-functions approach to ambiguous problems one substitutes a set of priors for the single prior of the models of risky choice.[1] Two assumptions in the subjective approach are necessary for domain equivalence. First, as mentioned, the preferences over measurable functions must be state independent, only the sets of distributions on consequences induced by the measurable functions can matter. Second, in a direct parallel with the need for the non-atomicity of the single prior in risky models, the set of priors must be open-minded lest the subjective model only capture preferences over a narrow class of sets of distributions on consequences.

1.3. **Open-Mindedness.** The realization that one's knowledge is, sometimes, not sufficient to assign precise probabilities to the utility relevant outcomes is a mild form of intellectual modesty. Open-mindedness, the focus of this paper, is a deeper property. It requires a decision maker's knowledge that their lack of understanding can, in some circumstances, take on quite arbitrary forms.

With any modeling assumption, one should ask about its costs and its benefits. Our main result, Theorem 1, shows that it is easy to satisfy open-mindedness, that the cost is small. Our second result, Theorem 2, as well as a number of examples show that the cost of *not* imposing it are large. One sees the parallel result in the completeness/incompleteness of preferences in risky problems as the prior is nonatomic/atomic.

The subjective, single prior, approach to risky problems only specifies preferences over those distributions that can be induced by some measurable functions, and the subjective, multi-prior, approach to ambiguous problems only specifies preferences over those sets of distributions that can be induced by some measurable functions. Theorem 2 and a number of examples show that this set is small in fashions that impede analysis if multi-prior is not open-minded. More specifically, if one models a decision maker who has a set of priors that fails to be open-minded, and most of those used in the literature are not open-minded, then the set of induced probabilities over outcomes always misses a dense set; often, the resulting preference incompleteness is so drastic that choice behavior does not depend on ambiguity attitudes.

1.4. **Implications for Previous Axioms and Interpretations.** By our lights, the study of multiple prior preferences over measurable functions is central to research looking for and looking into behavioral evidence of what we *start* with in this paper. We start by assuming decision makers' self-awareness of the incompleteness of their knowledge. This awareness is modeled as the decision maker having and using a set of priors, which can be understood as a set of 'scenarios' about what will follow from different courses of action.[2] Adding open-mindedness to this awareness yields complete overlap

[1]The subjective, or neo-Bayesian, multi-prior approach began with Schmeidler (1989); Gilboa and Schmeidler (1989), the relations between the axiomatizations and the consequent functional forms can be seen in the treatments of "invariant, bi-separable preferences" in Ghirardato and Marinacci (2001) and Amarante (2009). Cerreia-Vioglio et al. (2011) contains the corresponding material for state dependent choices, with special emphasis on the role of convexity as the definition of ambiguity aversion, already present in Schmeidler (1989). Variants on the objective approach can be found in Ahn (2008), Olszewski (2007), and Dumav and Stinchcombe (2017).

[2]Vasiliki Skreta suggested scenarios as an interpretation of the priors, variants of this idea are more thoroughly explored in Dumav and Stinchcombe (2017).

between the sets of choices the decision makers can subjectively describe and those that they can objectively describe. This is useful because the objective approach contributes to our understanding of ambiguous choice in ways that the subjective approach has not yet been able to capture.

- Machina (2009) shows that most of the axiomatic approaches to preferences over measurable functions have built in an assumption that implies that decision makers' attitudes toward ambiguity must be wealth independent. An apparent exception to this wealth independence is the work of Ghirardato et al. (2004), who give an α-max-min expected utility representation for measurable functions, a representation in which the parameter α can vary systematically with the measurable function. However, Proposition 1 shows that with their axioms, the parameter α can only vary if the decision maker's set of priors fails to be open-minded. By contrast, the objective approach delivers wealth dependent ambiguity and risk attitudes on sets of induced distributions (Dumav and Stinchcombe, 2017).
- Epstein (1999) and Ghirardato et al. (2004) define comparative ambiguity aversion for preferences over measurable functions. However, the definition only applies to pairs of decision makers having exactly the same risk attitudes. By contrast, in the objective approach, one can combine any risk attitude with any ambiguity attitude (Dumav and Stinchcombe, 2017). This allows one to study choice situations in which risk and ambiguity attitudes reinforce or counter each other.

1.5. **Outline.** The next section contains the main points of the paper in an example-based format. Of particular note is the extreme incompleteness of preferences that are possible when the set of priors fail to be open-minded. The subsequent section gives the formal development, characterizing the classes of open-minded priors, showing that the examples are instantiations of general patterns, and briefly examining open-mindedness when the priors may fail to be countably additive. The next section covers some of the implications of open-mindedness for previous axiomatizations of preferences over measurable functions. The last section provides a summary and a coda. Proofs and some calculations are gathered in the appendix.

2. Open-Mindedness: A Guided Tour

This section contains examples that demonstrate most of the results in the paper.

- §2.1 covers domain equivalence for models of risky choice.
- §2.2 covers domain equivalence for ambiguous models in the subjective approach, with particular attention to instances in which the use of a set of priors that is *not* open-minded can imply that ambiguity attitudes make no difference to choice.
- §2.3 covers the same material for the hybrid rather than the subjective approach, at a lower level of detail.
- §2.4 uses the lens of wealth dependence of risk and ambiguity attitudes to illustrate how the larger domain that comes with the assumption of open-mindedness changes the meaning and interpretation of previous axiomatizations of subjective state space models.

The formal development of the results behind the examples is in the next section.

2.1. **Domain Equivalence for Risky Decision Problems.** The objective and the subjective models of risky choice are related by change of variables. Most modern economics textbooks give the short axiomatic foundation for the von Neumann and Morgenstern (1944) (vNM) preferences over distributions on a space of consequences $\mathbb{X}$. Preferences satisfying their axioms rank distributions p and q by $p \succsim q$ if and only if

$$vNM(p) := \int_{\mathbb{X}} u(x)\, dp(x) \geq vNM(q) := \int_{\mathbb{X}} u(x)\, dq(x). \tag{1}$$

Here, the continuous expected utility function u on $\mathbb{X}$ is derived from properties of the preferences, and is unique up to positive affine transformation.

By contrast, Savage (1954) provides an axiomatic foundation for subjective preferences over measurable functions from a state space S, to the space of consequences $\mathbb{X}$. Preferences satisfying his axioms rank measurable functions $f, g : S \to \mathbb{X}$ by $f \succsim g$ if and only if

$$Sav(f) := \int_S u(f(\omega))\, d\mu(\omega) \geq Sav(g) := \int_S u(g(\omega))\, d\mu(\omega). \tag{2}$$

Here, the prior probability μ and the utility function u are *jointly* derived from properties of the preferences over measurable functions, the prior μ is unique and the utility function $u(\cdot)$ is, again, unique up to positive affine transformation.

The two representations, (1) and (2), are directly related by change of variables. If we take take $p = f(\mu)$ to be the distribution induced on $\mathbb{X}$ by f (defined by $f(\mu)(E) = \mu(f^{-1}(E))$ for $E \subset \mathbb{X}$) and take $q = g(\mu)$, then the integrals on each side of the inequalities (1) and (2) are the same. This equality depends on Savage's assumption of **state independence** — the utility function in Savage's preferences $u(\cdot)$ depends only on $x \in \mathbb{X}$, and not on $s \in S$. Domain equivalence requires more than this change of variables.

The vNM approach specifies complete preferences over all of $\Delta(\mathbb{X})$, the set of distributions on $\mathbb{X}$. However, depending on the prior, μ, the induced preferences over $\Delta(\mathbb{X})$ in the state-space approach may or may not be complete. For completeness in the subjective approach, it is necessary that every $p \in \Delta(\mathbb{X})$ is of the form $f(\mu)$ for some measurable $f : S \to \mathbb{X}$. For this, it is necessary, by Lemma 2 (below), and sufficient, by Skorohod (1956, Thm. 3.1.1), that the prior μ be non-atomic.

To see what is involved, we consider modeling a class of portfolio problems. Suppose there are assets, X_n, $n = 1, \ldots, N$, each of which has an identical distribution that returns $1,000$ euros with probability r and returns 0 with the remaining probability $1-r$. The returns are drawn independently. The portfolio choice problem in particular is how big a share z of the composite asset $X := \sum_n X_n$ to purchase at a price P that is smaller than its expected return $\mathbb{E}X$. With initial wealth W, the expected utility maximization problem is

$$\max_{z \in [0, W/P]} \mathbb{E}u((W - zP) + zX).$$

Notice that each portfolio decision z induces a distribution over the final wealth levels. Therefore, the decision maker's preference over the shares z of the composite portfolios can alternatively be represented as preferences over the distributions on the final wealth levels.

In the vNM approach, the specification of the distribution of X, a Binomial(N, r), is sufficient for the problem to be well-posed, there are no limitations on the possible values of either N or r. By contrast, if in a subjective model the prior μ has an atom, then measurable functions describe only a limited set of distributions of X. A pair of blunt examples make the essential point that such a prior limits modeling flexibility.

Translating this portfolio problem into the subjective framework, suppose that we model with a unit interval $[0,1]$ as a rich enough state space that describe all possible contingencies in the portfolio problem. Let the measurable functions $f_n : [0,1] \mapsto \{0, 1000\}$ represent the return profile of the assets X_n. The composite asset X is then the measurable function $f(s) := \sum_n f_n(s)$ for all $s \in [0,1]$. Suppose also that, for the sake illustration, we model on the state space $[0,1]$ with a prior μ that has an atom.

- If the prior μ has an atom of size $\boldsymbol{a} > 0$, then for any r and for any arbitrary N, some contingencies s in $[0,1]$ receives probability at least $\boldsymbol{a} > 0$. This implies that the distribution induced by the composite asset or equivalently by the measurable function f has an atom of size at least $\boldsymbol{a}$. Notice on the other hand that for a fixed r, for a large enough N, the distributions Binomial(N, r) assign to each outcome probability less than $\boldsymbol{a}$ and hence the induced distributions $f(\mu)$ cannot represent them.
- More generally, if μ is purely atomic, then for fixed N, one cannot vary r continuously and find a measurable function f_r with $f_r(\mu)$ capturing the Binomial(N, r).

In other words, the use of a prior with atoms in the subjective model entails not being able to represent the portfolio problem posed above. It bears repeating — if measurable functions do not represent the objects of choice due to a poor choice of a prior to work with, then the subjective approach does not model the choice problem in a flexible manner. It is the inability of a prior with atoms to describe a wide range of random phenomena that makes the non-atomicity assumption necessary for domain equivalence between the objective and the subjective models of risky choice, and non-atomicity is Postulate P6 of Savage (1954).

For ambiguous problems, the current discussion indicates that domain equivalence will require, at the very least, the non-atomicity of each prior in the set of priors. The main aim of this paper is to develop the replacement for non-atomicity appropriate for the multi-prior approach to choice in the presence of ambiguity.

2.2. **Domain Equivalence for Ambiguity: The Subjective Model.** The neo-Bayesian approach to choice under ambiguity replaces the single non-atomic prior, μ, of risky models with a convex set, Π, of non-atomic priors. Letting $f(\Pi) = \{f(\mu) : \mu \in \Pi\}$ a set Π is **open-minded** if each compact convex set of distributions on every compact metric space $\mathbb{X}$ is of the form $f(\Pi)$ for some measurable f. We begin by showing that there is a simple canonical set of open-minded priors, and then turn to examples that illustrate what can go wrong if one models a decision maker with a set of priors that is not open-minded.

2.2.1. *Open-Mindedness is Easy.* For risky choice, and indeed for most of probability theory, statistics, stochastic process theory, and econometrics, there is a simple canonical state space and prior — the unit interval, $[0,1]$, with the Borel σ-field and the uniform distribution. For multi-prior models of ambiguous choice, there is a comparably simple canonical *set* of priors: for each $x \in [0,1]$, let λ_x be the uniform distribution on the line $\{x\} \times [0,1]$ in $[0,1] \times [0,1]$; Π°, the closed convex hull of $\{\lambda_x : x \in [0,1]\}$, is open-minded. This characterization follows directly from Theorem 1, which shows that a set of priors, Π, on a state space S is open-minded if and only if

(a) there exists a measurable function $f_\lambda : S \to [0,1]$ such that for each $\mu \in \Pi$, $f(\mu) = \lambda$ (where λ is Lebesgue measure/the uniform distribution), and

(b) there exists a measurable function $f_\Delta : S \to [0, M]$ such that $f(\Pi) = \Delta([0, M])$, $M > 0$.

To verify that the canonical set Π° given above is open-minded, the measurable function f_λ can be defined by $f_\lambda(x, y) = y$, and the measurable function f_Δ can be defined by $f_\Delta(x, y) = M \cdot x$.

To interpret these conditions, for open-mindedness, it is necessary and sufficient that the decision maker with a set of priors Π can describe, as subjective uncertainty, the uniform distribution on an interval, and that they can describe, again as subjective uncertainty, knowing only that some distribution on an interval obtains. We believe that this is a minimal amount of sophistication to require. We now turn to examples illustrating what the failures of this minimal amount of sophistication entail for the modeling of decision makers.

2.2.2. *Bounded Radon-Nikodym Derivatives.* For $0 < c < 1 < d < \infty$, let $\Pi_{c,d}$ be the convex set of priors μ having densities satisfying $c \leq \frac{d\mu}{d\eta} \leq d$ where $\eta \in \Pi_{c,d}$ is a non-atomic prior. We examine the sets of distributions induced by measurable functions taking values in the two point set $\{\boldsymbol{x}, \boldsymbol{y}\}$ with $\boldsymbol{x} \prec \boldsymbol{y}$. This set of induced distributions, the **descriptive range** of $\Pi_{c,d}$, is the class of convex sets of distributions on $\{\boldsymbol{x}, \boldsymbol{y}\}$ on which subjective preference relations make pair-wise comparisons.

While it may be sensible to model a decision maker as having such a set of priors as $\Pi_{c,d}$, this choice has several drawbacks: the set of choices that it can describe is a closed set and has no interior; on the choice scenarios that can be described by $\Pi_{c,d}$, different attitudes toward ambiguity can make no difference to subjective choice behavior. The latter implies that ambiguity attitudes cannot be identified from choices between measurable functions — with $\Pi_{c,d}$ as the set of priors, the multi-prior model of preferences between measurable functions cannot be used to investigate the effects of different ambiguity attitudes.

Let $q \in [0, 1]$ denote the probability of the good outcome, $\boldsymbol{y}$. Consider an arbitrary measurable, $\{\boldsymbol{x}, \boldsymbol{y}\}$-valued function, $f(s) = \boldsymbol{y}1_E(s) + \boldsymbol{x}1_{E^c}(s)$, E a measurable set. For each f, the set of induced distributions, $f(\Pi_{c,d})$, is an interval of probabilities of the good outcome, $[a, b] \subset [0, 1]$. Since $a \leq b$, each such an interval can be represented as a point above the diagonal in $[0, 1] \times [0, 1]$. The **descriptive range** of $\Pi_{c,d}$, denoted $\mathcal{R}(\Pi_{c,d})$, is defined as the set of all such intervals, $\mathcal{R}(\Pi_{c,d}) := \{f(\Pi_{c,d}) : f \text{ is measurable}\}$. Figure 1 gives a typical example.

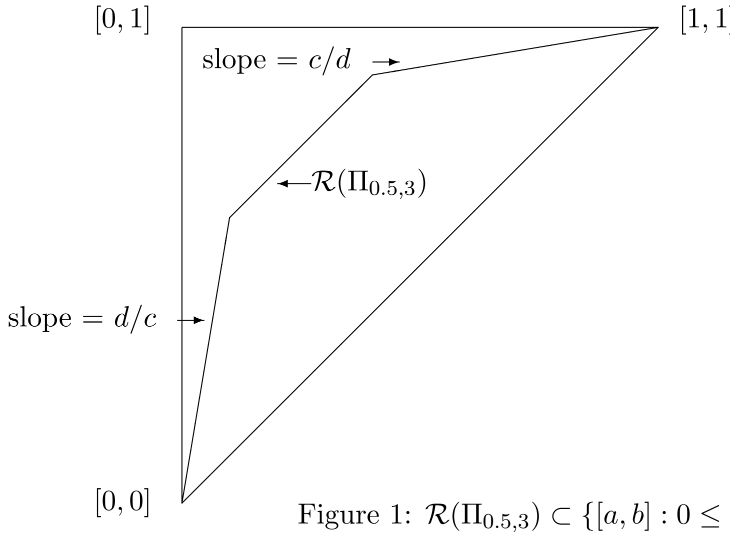


Figure 1: $\mathcal{R}(\Pi_{0.5,3}) \subset \{[a,b] : 0 \leq a \leq b \leq 1\}$

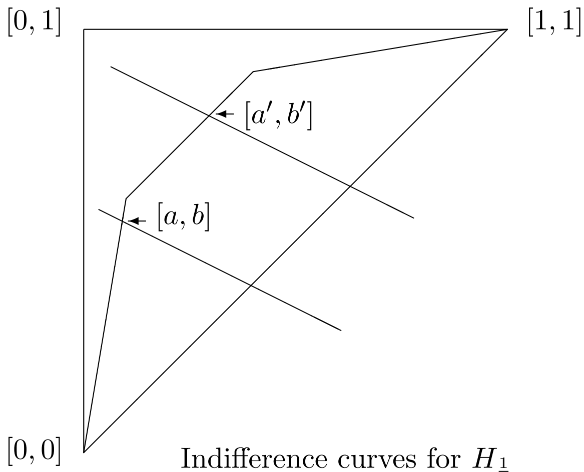
The set of choices that $\Pi_{c,d}$ can describe is the union of three line segments, a very small subset of the two-dimensional set of intervals. The descriptive range contains no risky choices except for certainty of the worst and the best outcomes, $[0,0]$ and $[1,1]$. It has (at least) one other distressing property, ambiguity attitudes make no difference to choices between measurable functions because all monotone preferences give the *same* ranking over $\mathcal{R}(\Pi_{c,d})$ — distinct $[a,b]$ and $[a',b']$ in $\mathcal{R}(\Pi_{c,d})$ either have $a' > a$ and $b' > b$ or else they $a' < a$ and $b' < b$. For the connection with ambiguity aversion, consider the Hurwicz criteria, for $0 \leq \alpha \leq 1$ and $[a,b] \subset [0,1]$,

$$H_\alpha([a,b]) := \alpha \min\{p \in [a,b]\} + (1-\alpha)\max\{q \in [a,b]\} = \alpha a + (1-\alpha)b. \tag{3}$$

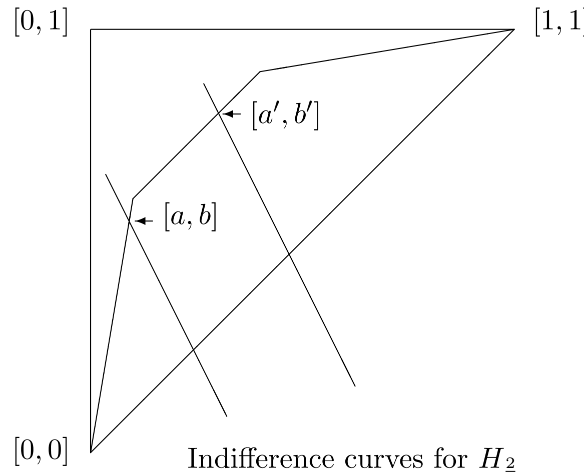
For $\alpha > \frac{1}{2}$, these preferences can be interpreted as demonstrating ambiguity aversion — higher weight on the worst possibility and lower weight on the best possibility. For any two sets $[a',b']$ and $[a,b]$ in the descriptive range and any $\alpha, \beta \in [0,1]$,

$$H_\alpha([a',b']) \geq H_\alpha([a,b]) \text{ if and only if } H_\beta([a',b']) \geq H_\beta([a,b]).$$

The geometry can be seen in the following Figure, which gives the intersection of $\mathcal{R}_{c,d}$ and indifference curves for $H_\alpha(\cdot)$ and $H_\beta(\cdot)$ for $\alpha = \frac{1}{3} < \frac{1}{2} < \beta = \frac{2}{3}$. Notice that two different attitudes $H_\alpha(\cdot)$ and $H_\beta(\cdot)$ rank the pairs of sets in the descriptive range $\mathcal{R}(\Pi_{c,d})$ the same way, and hence are not distinguishable if one uses a multi-prior $\Pi_{c,d}$ that is not open-minded.



Indifference curves for $H_{\frac{1}{3}}$ intersecting $\mathcal{R}_{c,d}$



Indifference curves for $H_{\frac{2}{3}}$ intersecting $\mathcal{R}_{c,d}$

2.2.3. *Larger/Smaller Sets of Priors.* In multiple prior models, larger/smaller set of priors are sometimes identified as representing more/less ambiguity. In some circumstances, this is true in a limited sense. In others, it is not.

- A larger set of priors do not necessarily represent a larger set of choices over outcomes. For $0 < c' < c < 1 < d < d'$, $\Pi_{c',d'}$ strictly contains $\Pi_{c,d}$. At issue is how the use of $\Pi_{c',d'}$ rather than $\Pi_{c,d}$ changes the choices that can be described. For any measurable $f : S \to \{\boldsymbol{x}, \boldsymbol{y}\}$, $f(s) = \boldsymbol{y}1_E(s) + \boldsymbol{x}1_{E^c}(s)$, inspection shows that the interval $f(\Pi_{c',d'})$ is larger than the interval $f(\Pi_{c,d})$. However, for any pair of monotone preferences, $\succ'$ and $\succ$, $f \succ' g$ using $\Pi_{c',d'}$ if and only if $f \succ g$ using $\Pi_{c,d}$, that is, the different sets of priors do not affect preference rankings over measurable functions. Perhaps even worse for the purposes of comparing decision makers, the two sets of priors describe essentially disjoint sets of choices — their only overlaps are certainty of the worst outcome, $[0, 0]$, and certainty of the best outcome, $[1, 1]$.
- A smaller set of priors may model a much larger set of choices. Consider the following one-dimensional subset $\Pi' = \{p_\theta : \theta \in [0, 2]\}$ where each prior $p_\theta \in \Delta([0, 1])$ has a density (with respect to Lebesgue measure λ)

$$h_\theta(s) = (2 - \theta)1_{[0,\frac{1}{2}]}(s) + \theta 1_{(\frac{1}{2},1]}(s).$$

  This set of priors can describe every interval $[a, b]$ in the triangle $\{[a, b] : 0 \leq a \leq b \leq 1\}$.[3] To see why, consider the functions

$$f_{[a,b]}(s) := \boldsymbol{y} \cdot 1_{[0,\frac{a}{2}] \cup (\frac{1}{2},\frac{1}{2}+\frac{b}{2}]}(s) + \boldsymbol{x} \cdot 1_{(\frac{a}{2},\frac{1}{2}] \cup (\frac{1}{2}+\frac{b}{2},1]}(s).$$

  As the variable θ ranges over the interval $[0, 2]$, the corresponding induced distribution $f_{[a,b]}(p_\theta)$ ranges from putting mass a on $\boldsymbol{y}$ to putting mass b on $\boldsymbol{y}$, so that $f_{[a,b]}(\Pi') = [a, b]$. With this smaller set of priors compared to $\Pi_{0,2}$, ambiguity attitude matters — the preferences described by $H_{\frac{1}{3}}$ and $H_{\frac{2}{3}}$ no longer agree in their rankings of pairs of functions f and g taking values in two point sets in the same way.

[3]However, it only describes a negligible subset of the sets of on three point outcome spaces.

In the other direction, the largest set of priors, $\Delta(S)$, has a very limited descriptive range. In what follows ρ_H denotes the Hausdorff distance (its formal definition is contained below in Section 3), and specialized to the two outcome case, $\mathbb{X} = \{x, y\}$, the following result tells us that the descriptive range of $\Pi = \Delta(S)$ contains only three sets, $\{\delta_x\}$, $\{\delta_y\}$, and $\Delta(\{x, y\})$.

Lemma 1. If $f : S \to \mathbb{X}$ is measurable and $F = \text{cl}(f(S))$ denotes the closure of the range of f, then $\rho_H(\Delta(F), f(\Delta(S))) = 0$. In particular, $f(\Delta(S))$ represents a risky choice $p \in \Delta(\mathbb{X})$ if and only if $f(s) \equiv p$.

2.3. **Domain Equivalence for Ambiguity: The Hybrid Model.** In Anscombe and Aumann (1963)'s hybrid model, $f : S \to \Delta(\mathbb{X})$ and each μ in Π induces a distribution on $\Delta(\mathbb{X})$, notationally, $f(\Pi) \subset \Delta(\Delta(\mathbb{X}))$. In this hybrid model, the convention is to identify each element Q of $\Delta(\Delta(\mathbb{X}))$ with its average $C(Q)$ where $C : \Delta(\Delta(\mathbb{X})) \to \Delta(\mathbb{X})$ is defined by $C(Q)(E) = \int \mu(E)\, dQ(\mu)$. As seen above, for risky problems, any prior, atomic or non-atomic, is sufficient to model any problem — simply have $f(s) \equiv p$ for $p \in \Delta(\mathbb{X})$. For ambiguous problems however, one cannot be quite so cavalier about the choice of the set of priors.

2.3.1. *Bounded Radon-Nikodym Derivatives.* With $\Pi_{c,d}$ defined as above and x, y, z being distinct consequences in $\mathbb{X}$, the question we address is how close a measurable $f : S \to \Delta(\mathbb{X})$ can come to having $C(f(\Pi_{c,d})) = \Delta(\{x, y, z\})$, i.e., representing full set of distributions as an element of its descriptive range. The answer turns out to be not very close because descriptive range of $\Pi_{c,d}$ in the hybrid model is limited as in the subjective model. To see this, if $f(s)$ is constant at the point mass δ_x, then $C(f(\Pi_{c,d})) = \delta_x$, but this is the only way in which $\{\delta_x\} \subset \Delta(\{x, y, z\})$ can be in the descriptive range for the hybrid model.

In general, the set of probabilities $(p_x, p_y, p_z) \in \Delta(\{x, y, z\})$ that can be of the form $C(f(\Pi_{c,d}))$ must satisfy the following set of inequalities and equalities,

$$\begin{aligned} &p_x \in [cq_x, dq_x],\ p_y \in [cq_y, dq_y],\ p_z \in [cq_z, dq_z], \\ &p_x + p_y + p_z = 1, \text{ and} \\ &q_x + q_y + q_z = 1. \end{aligned}$$

One arrives at these restrictions on the descriptive range by considering the simple measurable functions of the form $f(s) = \delta_x 1_{E_x} + \delta_y 1_{E_y} + \delta_z 1_{E_z}$ where E_x, E_y, E_z is a measurable partition of S (which spans all measurable functions for the outcome space in question) and noticing $q_x = \eta(E_x)$, $q_y = \eta(E_y)$, and $q_z = \eta(E_z)$.

2.3.2. *Larger/Smaller Sets of Priors.* For $0 < c' < c < 1 < d < d'$, the set of priors $\Pi_{c',d'}$ strictly contains $\Pi_{c,d}$. This increase loosens the set of inequalities given above that defines the descriptive range $C(f(\Pi_{c,d}))$, and $C(f(\Pi_{c',d'}))$ comes closer to $\Delta(\{x, y, z\})$ than $C(f(\Pi_{c,d}))$. The largest set of priors, $\Delta(S)$, works well in this case — for any compact, convex $A \subset \Delta(\mathbb{X})$, if $f(S)$ contains the extreme points of A, then $C(f(\Delta(S)))$ is equal to A. This means that the conditions for descriptive completeness given in Theorem 1 are sufficient for descriptive completeness in the hybrid model, but no longer necessary.

2.4. **On the Wealth Dependence of Choice Behavior.** We have seen that open-mindedness is necessary for subjective models to describe the choice problems encompassed by the objective approach. But the contribution of open-mindedness to the axiomatic theory of subjective choice in the presence of ambiguity is larger than this indicates, it changes the meaning and interpretation of axioms. The reason is that the implications of axioms over preferences on smaller choice domains can be very different than their implications on larger domains: on small domains, the functional forms satisfying the axioms can be quite flexible, and may seem suitable to a rich study of comparative statics; however, on larger domains, the axioms may rule out much of the flexibility. This becomes clear by analyzing in particular how wealth differences affect the choice differences between ambiguous prospects.

Consider an investment problem with varying initial wealth W. Suppose that a decision maker has a set of priors Π, and that the decision maker is choosing between the pair of measurable functions f and g which, respectively, yield the sets of distributions $f(\Pi)$ and $g(\Pi)$ on monetary consequences $[0, M]$. For the Hurwicz preferences with parameter (α), the choice is strictly for f over g if and only if

$$\begin{aligned}&\left[\alpha \cdot \min_{\mu\in\Pi} \textstyle\int_S u(f(s))\, d\mu(s) + (1-\alpha)\cdot \max_{\eta\in\Pi} \int_S u(f(s))\, d\eta(s)\right] > \\ &\quad \left[\alpha \cdot \min_{\mu\in\Pi} \textstyle\int_S u(g(s))\, d\mu(s) + (1-\alpha)\cdot \max_{\eta\in\Pi} \int_S u(g(s))\, d\eta(s)\right]. \end{aligned} \tag{4}$$

For $W > 0$ and each state s, let $f_W(s) = f(s) + W$ and $g_W(s) = g(s) + W$. This has the effect of shifting the sets of distributions $f(\Pi)$ and $g(\Pi)$ to the right by an amount W. For $W_1 < W_2$, one could imagine that the worst distributions in $g_{W_1}(\Pi)$ are very salient, e.g. because they include distributions that reduce the decision maker to penury, so that $f_{W_1} \succ g_{W_1}$. Such considerations matter less when wealth is larger, and this difference might reverse the preference ordering, so that $g_{W_2} \succ f_{W_2}$.

To capture such phenomena, Ghirardato et al. (2004) and Cerreia-Vioglio et al. (2011) provide an axiomatic development of preferences over measurable functions that allow the weight α to depend on the measurable function f. This means that the choice is strictly for f over g if and only if $\mathcal{H}(f) > \mathcal{H}(g)$ where

$$\begin{aligned}\mathcal{H}(f) &= \alpha_f \cdot \min_{\mu\in\Pi} \int_S u(f(s))\, d\mu(s) + (1-\alpha_f)\cdot \max_{\mu\in\Pi} \int_S u(f(s))\, d\mu(s) \quad \text{and} \\ \mathcal{H}(g) &= \alpha_g \cdot \min_{\mu\in\Pi} \int_S u(g(s))\, d\mu(s) + (1-\alpha_g)\cdot \max_{\mu\in\Pi} \int_S u(g(s))\, d\mu(s).\end{aligned}$$

Ghirardato et al. (2004) and Cerreia-Vioglio et al. (2011) differ in the conditions on f and g under which α_f must equal α_g.

- Proposition 1 (below) implies that an open-minded decision maker satisfying the Ghirardato et al. (2004) axioms must have the function $f \mapsto \alpha_f$ constant. In other words, while their axioms might allow for wealth dependent ambiguity attitudes, they can only allow it if the preferences are restricted to the small set of problems that can be modeled with a set of priors that fails to be open-minded.[4]
- Proposition 2 (below) implies that for an open-minded decision maker satisfying the Cerreia-Vioglio et al. (2011) axioms, one cannot change the ambiguity attitudes, which they capture with the weight α_f, without simultaneously changing the risk attitude, captured in the curvature of $u(\cdot)$. In other words, these axioms

[4]Eichberger et al. (2011) show that this class of preferences must have α identically equal to either 0 or 1 when the state space S is finite. They further show that if Π is the set of all countably additive probabilities on $[0, 1]$, then the axioms can be satisfied with a constant α. Our contribution is to show that α must be constant if Π allows for subjective descriptions of all of the uncertainty encompassed by the objective model.

can only allow for independence of risk and ambiguity attitudes if the preferences are restricted to the small set of problems that can be modeled with a set of priors that fails to be open-minded.

In the first case, one cannot examine how wealth changes the salience of worst cases when examining choice between ambiguous investment choices. In the second case, one cannot examine the effects of wealth changes on ambiguous investments without also changing the attitudes toward risk. The objective approach need have neither limitation, see Olszewski (2007); Ahn (2008); Dumav and Stinchcombe (2017).

3. Open-Mindedness: The Formal Development

We assume that our decision makers can be thought of as having Kolmogorov's model for random phenomena. There is a measure space,[5] $(S, \mathcal{S})$ in which randomness occurs. The utility relevant outcomes arise as measurable functions, typically denoted f or g, from the state space to the *compact* metric space of utility relevant outcomes, $\mathbb{X}$.

- §3.1 sets the notation and assumptions.
- §3.2 has three results: Theorem 1 characterizes open-minded sets of priors; Corollary 1.1 shows that classical statistical models are open-minded; and Corollary 1.2 gives a useful almost everywhere continuity result.
- §3.3 covers Theorem 2, showing that a set of priors that fails to be open-minded must fail to describe a dense class of sets.
- §3.4 gives two sufficient conditions for the open-mindedness of a sets of priors that fail to be countably additive.

3.1. Notation and Assumptions.

We assume throughout that singleton sets are measurable, that is $\{s\} \in \mathcal{S}$ for all $s \in S$. Priors on $(S, \mathcal{S})$ are countably additive probabilities. The set of all possible priors is denoted by $\Delta(S)$ and endowed with the minimal σ-field containing the sets $\{\mu : \mu(E) \leq x\}$, $E \in \mathcal{S}$ and $x \in [0, 1]$. Throughout, the set Π denotes a convex set of priors that is closed in the weakest topology making the mappings $E \mapsto \mu(E)$, $E \in \mathcal{S}$, continuous.

The compact metric space of utility relevant consequences is denoted $\mathbb{X}$. The set of countably additive probabilities on the Borel σ-field of $\mathbb{X}$, denoted $\Delta(\mathbb{X})$, is a compact, metric space when given the Prohorov metric,

$$\rho(p, q) := \inf\{\epsilon \geq 0 : \text{ for all closed } F,\ p(F) \leq q(F^\epsilon) + \epsilon \text{ and } q(F) \leq p(F^\epsilon) + \epsilon\}.$$

This metric has the property that $\rho(p^n, p) \to 0$ if and only if $\int v\, dp^n \to \int v\, dp$ for all continuous $v : \mathbb{X} \to \mathbb{R}$.

A prior μ and a measurable function $f : S \to \mathbb{X}$ induce a distribution (also referred to as image law) in $\Delta(\mathbb{X})$. The induced distribution is denoted by $f(\mu)$ and defined as $f(\mu)(E) = \mu(\{s \in S : f(s) \in E\})$. The **descriptive range** of a set of priors Π is the class of sets $\{f(\mu) : \mu \in \Pi, f \text{ is measurable}\}$, denoted compactly by $f(\Pi)$. We assume that preferences over the measurable functions are **state independent**, that is, if $f(\Pi) = g(\Pi)$, then f and g are indifferent. In other words, the state space serves *only* to model randomness.

The class of non-empty, closed convex subsets of $\Delta(\mathbb{X})$ is denoted by $\mathcal{K}_{\Delta(\mathbb{X})}$ and metrized by the Hausdorff metric, $\rho_H(A, B) := \max(\max_{p \in A} \rho(p, B), \max_{q \in B} \rho(q, A))$.

[5] A non-empty set and a σ-field of subsets.

3.2. **Characterization of Open-Mindedness.** The objective approach to ambiguous choice problems starts from preferences on $\mathcal{K}_{\Delta(\mathbb{X})}$. In order for the subjective approach to cover the same set of problems with a set of priors, it is necessary that the set of priors can describe, via some $\mathbb{X}$-valued measurable function, every element of $\mathcal{K}_{\Delta(\mathbb{X})}$.

Definition 1. A measurable, convex set of probabilities, Π, on a measure space $(S, \mathcal{S})$ is **open-minded** if for any compact metric space $\mathbb{X}$ and any $A \in \mathcal{K}_{\Delta(\mathbb{X})}$, there exists a measurable function $f_A : S \to \mathbb{X}$ such that $\{f_A(p) : p \in \Pi\} = A$.

Recall that λ denotes the uniform distribution on the Borel σ-field for $[0, 1]$.

Theorem 1. Π is open-minded if and only if

(a) there exist a measurable function $f_\lambda : S \to [0, 1]$ such that for each $\mu \in \Pi$, $f(\mu) = \lambda$, and
(b) there exists a measurable function $f_\Delta : S \to [0, 1]$ such that $f_\Delta(\Pi) = \Delta([0, 1])$.

Let Π° denote the following canonical set of priors: for each $x \in [0, 1]$, λ_x is the uniform distribution on $\{x\} \times [0, 1] \subset [0, 1] \times [0, 1]$, and Π° is the closed convex hull of $\{\lambda_x : x \in [0, 1]\}$ in $\Delta(\mathbb{X})$. To verify the conditions of Theorem 1, define $f_\lambda(x, y) = y$ and $f_\Delta(x, y) = x$.

The conditions in Theorem 1 are minimum needed to guarantee that Π models risky problems and problems in which the decision maker chooses between different sets $\Delta(F)$. Fix an arbitrary compact metric space (M, d).

- The first condition guarantees that Π models all risky problems with M as the space of consequences. For any Borel probability p on M, there exits a measurable $g : [0, 1] \to M$ such that $g(\lambda) = p$. Therefore $\{p\} = g(f_\lambda(\Pi))$.
- The second condition guarantees that Π models $\Delta(F)$ for any closed $F \subset M$. By the Borel isomorphism theorem (e.g. Dellacherie and Meyer (1978, Theorem III.20) or Dudley (2002, Theorem 13.1.1)), there exists a measurable $h : [0, 1] \to F$ that is onto (and h can be taken to be one-to-one with a measurable inverse if F is uncountable). Therefore $h(f_\Delta(\Pi)) = \Delta(F)$.

It is striking that these minimal conditions are also sufficient.

3.2.1. *The Classical Statistical Models.* Breiman et al. (1964) characterize statistical models of sequences of observations for which there exist consistent estimators. Let Π be a set of probabilities on $\mathbb{R}^{\mathbb{N}}$ representing the distributions of sequences of observations taking values in $\mathbb{R}$.

From Dellacherie and Meyer (1978, Dfn. III.16), $(\Omega, \mathcal{F})$ is a **Lusin** measurable space if it is measurably isomorphic to a measurable subset of a complete separable metric space. The set of countably additive probabilities on a Lusin space is itself a Lusin space with the Prohorov metric.

Definition 2. A set of probabilities Π on a Lusin space $(\Omega, \mathcal{F})$ is **strongly zero-one** if there exists a measurable $\Omega' \subset \Omega$ and an onto measurable mapping $\varphi : \Omega' \to \Pi$ such that for all $\mu \in \Pi$, $\mu(\varphi^{-1}(\mu)) = 1$.

Breiman et al. (1964) show that a set of probabilities is strongly zero-one if and only if there exist consistent estimators for the probabilities in Π. The classical statistical models take Π to be a smoothly parameterized set $\{\mu_\theta : \theta \in \Theta\}$, of i.i.d. sequences where Θ is an open subset of $\mathbb{R}^p$ and maximum likelihood estimators are known to be consistent.

Corollary 1.1. The closed convex hull of a strongly zero-one set of non-atomic priors is open-minded.

The "closed convex hull" in the previous arises because we here define open-mindedness only for closed convex sets. The following are examples of such sets.

- For each $x \in (0,1)$, let μ_r be the necessarily non-atomic distribution of an i.i.d. sequence of Bernoulli(r) random variables in the sequence space $\Omega = \{0,1\}^{\mathbb{N}}$. To verify that this is a strongly zero-one set, define $\psi(\omega) = \liminf_T \frac{1}{T}\sum_{t\leq T}\omega_t$, set $\Omega' = \psi^{-1}((0,1))$, and define $\varphi(\omega) = \mu_{\psi(\omega)}$.
- More generally, let Θ be an uncountable, measurable subset of a Lusin space, and let $\theta \mapsto p_\theta$ be a continuous, one-to-one mapping from Θ to the set of non-degenerate distributions on a complete separable metric space (M,d). For each θ, let μ_θ be the corresponding distribution on $M^{\mathbb{N}}$ of an i.i.d. sequence of random variables having distribution p_θ. Let Ω' denote the set of 'ergodic' sequences in Ω with long run distribution equal to one of the P_θ, that is, $\lim_T \frac{1}{T}\sum_{t\leq T} 1_E(\omega_t) = P_\theta(E)$, $E \subset M$ a measurable set. For each $\omega' \in \Omega'$, define $\varphi(\omega')$ equal to the corresponding μ_θ.

3.2.2. *Almost Everywhere Continuity.* For non-atomic probabilities, Skorohod (1956, Thm. 3.1.1) shows that one can, without loss, replace convergence in distribution with almost everywhere convergence — if a probability μ is non-atomic and $\rho(p_n, p) \to 0$, then there exist measurable functions f, f_n such that $f(\mu) = p$, $f_n(\mu) = p_n$ and $\mu(\{s : f_n(s) \to f(s)\}) = 1$. This replacement makes many arguments more transparent. For open-minded sets of probabilities, one has the analogous continuity result for sets of distributions. The following shows that if the set of induced distributions, A_n, converge to a set A, and Π is open-minded, then it is without loss to assume that the measurable functions converge μ-almost everywhere for each $\mu \in \Pi$.

Corollary 1.2. If Π is open-minded and $\rho_H(A_n, A) \to 0$ in $\mathcal{K}_{\Delta(\mathbb{X})}$, then there exist a sequence of measurable functions f, f_n with $f(\Pi) = A$, $f_n(\Pi) = A_n$, and $\mu(\{s : f_n(s) \to f(s)\}) = 1$ for each $\mu \in \Pi$.

Comments. Dumav and Stinchcombe (2016) gives an alternate characterization of open-mindedness, "measurably mutually orthogonal and simultaneously Skorohod," that is intermediate between being strongly zero-one and the simpler conditions in Theorem 1, and it proves the continuity result, Corollary 1.2. As well as being more clearly interpretable, the simplicity of the conditions in Theorem 1 greatly simplify the proofs.

3.3. **Failures of Open-Mindedness.** We begin with the general results about the small descriptive range for single priors having an atom and for sets of priors that fail to be open-minded. We then turn to the parallel result for sets of priors that fail to be open-minded.

For subjectively risky choice models, non-atomicity is crucial for domain equivalence. If a prior μ has an atom of size $\boldsymbol{a}$, then $f(\mu)$ has an atom of size $\boldsymbol{a}$ or greater. The following tells us that the descriptive range of a single prior with an atom must miss (more than) a dense set when the utility relevant outcomes belong to $[0, M]$.

Lemma 2. For any $\boldsymbol{a} > 0$ the set of probabilities $\{p \in \Delta([0,M]) : \text{there exists } x_0 \in [0,M] \text{ s.t. } p(x_0) \geq \boldsymbol{a}\}$ is closed and nowhere dense.

A similar miss-a-dense-set result holds for sets of priors that fail to be open-minded, though the descriptive range need not be closed.

Theorem 2. If every open set in $\mathbb{X}$ is uncountable and $\mathbb{F}$ is a closed, convex set of priors that fails to be open-minded, then there is a dense subset of $\mathcal{K}_{\Delta(\mathbb{X})}$ that does not belong to the descriptive range of $\mathbb{F}$.

The examples in the previous section showed that for several previously suggested sets of priors, the closure of the descriptive range has empty interior. This is much stronger than missing a dense set.

3.4. **Open-Mindedness without Countable Additivity.** It is sometimes convenient to work with priors that are finitely but not necessarily countably additive. In this case, the induced measures, say on $[0, M]$, may not be countably additive. We give two sufficient conditions for the open-mindedness of a set of priors that may not satisfy countable additivity. We do not have necessary conditions.

For this subsection, priors on $(S, \mathcal{S})$ are non-negative, *finitely* additive set functions $\mu : \mathcal{S} \to [0, 1]$ satisfying $\mu(S) = 1$. The set of priors is $\Delta^{fa}(S)$, and this set is given the weak*-topology — a net $\mu^i \to \mu$ iff $\int f\, d\mu^i \to \int f\, d\mu$ for all bounded measurable $f : S \to \mathbb{R}$, equivalently, iff $\mu^i(E) \to \mu(E)$ for all measurable E.

Finitely additive probabilities p and q on the Borel σ-field of $\mathbb{X}$ are **continuously equivalent** if $\int v\, dp = \int v\, dq$ for all continuous $v : \mathbb{X} \to \mathbb{R}$. If p and q are countably additive, continuous equivalence implies equality, but this is not true if p or q fail to be countably additive (Corbae et al., 2009, §9.8.a). The set of equivalence classes of probabilities on the Borel σ-field of $\mathbb{X}$, denoted $\Delta^{fa}(\mathbb{X})$, is a compact, pseudo-metric space with the Prohorov metric, $\rho(\cdot, \cdot)$. The non-empty, ρ-compact, convex subsets of $\Delta^{fa}(\mathbb{X})$ are denoted $\mathcal{K}_{\Delta^{fa}(\mathbb{X})}$, and meterized by the Hausdorff pseudo-metric, $\rho_H(A, B) := \max\left(\max_{p\in A} \rho(p, B), \max_{q\in B} \rho(q, A)\right)$.

In the weak* topology, $\Delta^{fa}(S)$ is compact, and for any measurable $f : S \to \mathbb{X}$, the mapping $\mu \mapsto f(\mu)$ to the pseudo-metric space $\Delta^{fa}(\mathbb{X})$ is continuous. Hence, for any measurable f, $f(\Pi)$ is a closed, hence compact, convex subset of $\Delta^{fa}(\mathbb{X})$.

For risky choice, non-atomicity is still necessary and sufficient for domain equivalence. For necessity, note that Lemma 2 still applies in the pseudo-metric space $\Delta^{fa}(\mathbb{X})$. For sufficiency we have the following.

Lemma 3. If $\mu \in \Delta(S)$ is non-atomic and $p \in \Delta^{fa}(\mathbb{X})$, then there exists a measurable $f : S \to \mathbb{X}$ such that $\rho(f(\mu), p) = 0$.

The following gives sufficient conditions for open-mindedness.

Theorem 3. If Π is a convex, weak* compact subset of priors, then either of the following conditions are sufficient for Π to be open-minded.

(a) Π is the weak*-closed, convex hull of a countably infinite set of disjointly supported, non-atomic probabilities.
(b) Π is the weak*-closed, convex hull of an open-minded set of countably additive probabilities, Π^{ca}.

Comments. Theorem 3(b) shows that one can take the weak* closure — in the class of finitely additive probabilities — of any open-minded sets of countably additive probabilities and retain open-mindedness. Theorem 3(a) seems to involve only countable sets of probabilities, but taking the weak* closure adds a set of mutually orthogonal (a generalization of disjointly supported) probabilities having cardinality larger than the continuum. Stinchcombe (1997) gives several general techniques for expanding state space choice models so as to retain the structure of the problems being modeled while avoiding the paradoxes that can arise with failures of countable additivity. Unfortunately, none of these approaches has yielded a characterization of open-mindedness for finitely additive probabilities.

4. Axiomatics in the Presence of Open-Mindedness

Open-mindedness, through the larger domain of problems that it entails, changes the meanings, implications and interpretations of previously developed axiomatic formulations of subjective preferences in the presence of ambiguity. We show this by examining the additional restrictions that open-mindedness implies for the functional forms that these preferences take.

4.1. **The $\alpha-$MEU model.** For fixed Π, continuous vNM expected utility function $v : \mathbb{X} \to \mathbb{R}$, and measurable $f, g : S \to \mathbb{X}$, define the corresponding mappings $v_f, v_g : \Pi \to \mathbb{R}$ by $v_f(\mu) = \int v \circ f \, d\mu$ and $v_g(\mu) = \int v \circ g \, d\mu$. Ghirardato et al. (2004) require that $\alpha_f = \alpha_g$ in their modified Hurwicz criterion if $v_f(\cdot)$ and $v_g(\cdot)$ are positive affine transformations of each other on the set Π.

Proposition 1. If Π is open-minded, $f, g : S \to \mathbb{X}$ are measurable, $v : \mathbb{X} \to \mathbb{R}$ is continuous and neither $v_f(\cdot)$ nor $v_g(\cdot)$ is constant, then there exists $f', g' : S \to \mathbb{X}$ such that $f'(\Pi) = f(\Pi)$, $g'(\Pi) = g(\Pi)$, and the mappings $v_f(\cdot)$ and $v_g(\cdot)$ are positive affine transformations of each other.

Comments.

- With state independence, $[f(\Pi) = g(\Pi)] \Rightarrow [f \sim g]$, this proposition yields the constancy of $f \mapsto \alpha_f$: state independence and $f(\Pi) = f'(\Pi)$ yield $\alpha_f = \alpha_{f'}$; $g(\Pi) = g'(\Pi)$ yields $\alpha_g = \alpha_{g'}$; since $v_{f'}(\cdot)$ and $v_{g'}(\cdot)$ are affine transformations of each other, $\alpha_{f'} = \alpha_{g'}$; combining, $\alpha_f = \alpha_g$.
- Much less than the descriptive completeness of Π can force the $f \mapsto \alpha_f$ mapping to be constant. For example, if $\Pi = \{\beta\mu + (1-\beta)\nu : c \leq \beta \leq d\}$, $0 \leq c < d \leq 1$, then for any pair f and g not delivering constant utilities on Π, $v_f(\cdot)$ and $v_g(\cdot)$ are positive affine transformations of each other on Π.

4.2. **Open-Mindedness and MBA Preferences.** If one can loosen the conditions under which one must have $\alpha_f = \alpha_g$ for preferences with a variable α-Hurwicz representation, then one can potentially encompass more variability in e.g. the wealth dependence of attitudes toward ambiguity. The axiomatization of monotonic Bernoullian and Archimedean (MBA) preferences in Cerreia-Vioglio et al. (2011) requires that $\alpha_f = \alpha_g$ only in the case that the mappings $v_f(\cdot)$ and $v_g(\cdot)$ are equal, that is, only in the case that for each $\mu \in \Pi$, $\int u(f(s)) \, d\mu(s) = \int u(g(s)) \, d\mu(s)$. This seems a very stringent condition, which, hopefully means that there is a great deal of flexibility in how α_f can depend on f. However, in the presence of state independence and an open-minded set of priors, the following implies that $\alpha_f = \alpha_g$ under the much weaker condition that the ranges of $v_f(\cdot)$ and $v_g(\cdot)$ are equal.

Proposition 2. Suppose that Π is open-minded, that $f, g : S \to \mathbb{X}$ are measurable, and that $v : \mathbb{X} \to \mathbb{R}$ is continuous. If the sets $v_f(\Pi)$ and $v_g(\Pi)$ are equal, then there exist measurable $f', g' : S \to \mathbb{X}$ such that $f'(\Pi) = f(\Pi)$, $g'(\Pi) = g(\Pi)$ and for all $\mu \in \Pi$, $v_{f'}(\mu) = v_{g'}(\mu)$.

4.3. **The Smooth Ambiguity Model.** The Klibanoff et al. (2005) smooth model of ambiguity aversion also specifies preferences over measurable functions using sets of priors, Π. However, instead of preferences depending on comparisons between $f(\Pi)$ and $g(\Pi)$, they use a distribution, Q, on Π, and compare the distributions on $f(\Pi)$ and $g(\Pi)$ that are induced. Here too, the open-mindedness of Π allows the subjective and the objective versions of the model to cover the same set of choices.

For measurable functions, $f \succ g$ in the smooth ambiguity approach if and only if

$$\text{Smooth}(f) := \int_\Pi \varphi(\langle u, f(\mu)\rangle)\, dQ(\mu) > \text{Smooth}(g) := \int_\Pi \varphi(\langle u, g(\mu)\rangle)\, dQ(\mu).$$

There are several parts to this representation: Π is a set of priors and $Q \in \Delta(\Pi)$ is a distribution on Π; $u(\cdot)$ is the von Neumann-Morgenstern utility function for risky choice; $f(\mu), g(\mu) \in \Delta(\mathbb{X})$; $\langle u, f(\mu)\rangle$ is the integral $\int_{\mathbb{X}} u(x)\, df(\mu)(x)$ (with the same definition for g); and $\varphi : [0,1] \to [0,1]$ is increasing, concave and onto.

Let $A = f(\Pi)$, $B = g(\Pi)$, and let $\mathfrak{p}$ and $\mathfrak{q}$ in $\Delta(\Delta(\mathbb{X}))$ denote the distributions on A and B that arise when picking μ according to the distribution Q. The objective version of these preferences are represented by $\mathfrak{p} \succ \mathfrak{q}$ if and only if

$$\text{Smooth}_{cov}(\mathfrak{p}) := \int_A \varphi(\langle u, r\rangle)\, d\mathfrak{p}(r) > \text{Smooth}_{cov}(\mathfrak{q}) := \int_A \varphi(\langle u, r\rangle)\, d\mathfrak{q}(r).$$

The question is whether every $\mathfrak{p}$ and $\mathfrak{q}$ can be induced by some measurable f and g for some Π and some $Q \in \Delta(\Pi)$.

Proposition 3. There exists an open-minded Π and a non-atomic $Q \in \Delta(\Pi)$ that allows for representation of every $\mathfrak{p}$ in $\Delta(\Delta(\mathbb{X}))$, that is, there exists an $f : S \to \mathbb{X}$ such that $\mathfrak{p}$ is the distribution of $f(\mu)$ when $\mu \in \Pi$ is distributed according to Q.

Comments.

- If Q is a non-atomic distribution on a measurable space $(Z, \mathcal{Z})$, then the Skorohod representation theorem tells us that the measurable functions $f : Z \to \mathbb{X}$ can induce any distribution on $\mathbb{X}$. Here Q is a distribution on $\Delta(Z)$ but the measurable functions still map Z to $\mathbb{X}$ and one wishes to induce a distribution on $\Delta(\mathbb{X})$.
- One can construct non-atomic distributions Q on an open-minded set of priors Π for which the representation result just given does not hold, but we do not have a characterization of the set of Q for which it does hold.
- As with the previously studied preferences, open-mindedness allows for tractable comparative statics analyses — one does not need to check that the choices are describable as the necessary first step. For example, Proposition 3 guarantees that different choices of f or g allow one to e.g. keep A fixed and vary $\mathfrak{p} \in \Delta(A)$ to any $\mathfrak{q} \in \Delta(A)$. This would make it possible to have $\mathfrak{q}$ be a mean preserving spread of $\mathfrak{p}$, which captures a sense of there being more ambiguity.

5. Summary and a Coda

There are two main results in this paper. The first is a characterization theorem, a closed convex set of priors can induce the uniform distribution on an interval and it can induce the set of all distributions on an interval if and only if it can induce any convex set of distributions on any compact metric space. The second gives the pertinent implication of not meeting the given conditions — if a closed convex set of priors does *not* meet these criteria, then it must necessarily fail to describe a dense class of convex sets. We view our results as showing that the costs of assuming open-mindedness are small and the costs of *not* assuming it large. We are *not* arguing that modelers should avoid sets of priors that fail open-mindedness if they think the set of priors is appropriate. Rather, we are arguing that second result and the examples demonstrate that the costs of using sets that fail to be open-minded are larger than was known.[6]

If a subjective state space model cannot describe a class of options, then it does not specify preferences over that class. The incompleteness of preferences can be drastic, many of the commonly used sets of priors can describe so little as subjective uncertainty that they cannot be used to pose questions about about the effects of ambiguity aversion on choice. A central theoretical requirement in economic modeling is the ability to answer questions about how decisions change as the situation faced by the decision maker changes. The ability to conduct such comparative statics analyses with multiple prior models requires that the changes in the situations can be described inside the model. With an open-minded set of priors, it is not necessary to check this before beginning the analysis.

As to other benefits of assuming open-mindedness, the canonical set of priors is as easy to work with as the unit interval in probability theory. It provides an easy way to check intuitions and to visualize proofs. Other familiar examples of open-minded sets of priors include the (closed convex hull of) parametrized sets of distributions of sequences of i.i.d. observations used in classical statistical models provide open-minded sets of priors. Whichever open-minded set of priors one chooses to work with, we reap the benefit of having the objective and subjective approaches to ambiguous choice complementing and informing each other.

Beyond tying together disparate approaches to choice theory, open-mindedness provides a window on the implications of axiomatic approaches to preferences over measurable functions. The extant work has shown that preferences satisfying very weak sets of axioms have representations involving a convex sets of priors Π and various integral formulations of the utility functions over induced sets of distributions. It is the combination of Π with the class of measurable functions that determines what choices can be represented, and what can be represented matters for the meaning and interpretation of the axioms.

One can see this in the allowable patterns of wealth dependence of ambiguity attitudes for both the Ghirardato et al. (2004) and the Cerreia-Vioglio et al. (2011) axioms. The functional forms are variants of the Hurwicz criterion in which the weight on the worst expected utility and the best expected utility can, in principle, vary with the measurable function inducing the set of distributions. If the set of priors can describe the full range of choices, i.e. if it is open-minded, then the weight on the worst utility outcome cannot vary at all with the first set of axioms, and one cannot independently change risk attitudes and the weights with the second set. Flexibility similar to having have ambiguity attitudes be wealth dependent are also available if one uses an open-minded set of priors in the smooth ambiguity model.

The smooth ambiguity model of Klibanoff et al. (2005) offers a flexible subjective framework in which a decision maker has ambiguous beliefs about the lotteries their choice yields. In this approach, the functional form involves a distribution over the set of priors, and this distribution in turn induces a distribution over distributions on the space of utility relevant consequences. Here again, comparative statics are not available

[6]Our thanks to an anonymous referee who made it very clear that we were not at all clear about this point in the previous version of this paper.

until one has checked what can be described, and here again, open-mindedness can be used to render this first step superfluous.

A coda. Halmos (1960, p. vi) wrote, in his textbook about "naive" set theory, that a prospective mathematician should "read it, absorb it, and forget it." It is our hope that choice theorists can treat open-minded sets of priors in much the same way. That they can get on with modeling regularities in choice behavior in the face of ambiguity, knowing, even if only in the back of their minds, that the people in their models can represent as much subjective uncertainty as the theory requires of them.

References

D. S. Ahn. Ambiguity without a state space. *Rev. Econom. Stud.*, 75(1):3–28, 2008.

M. Amarante. Foundations of neo-Bayesian statistics. *J. Econom. Theory*, 144(5): 2146–2173, 2009.

F. J. Anscombe and R. J. Aumann. A definition of subjective probability. *Ann. Math. Statist.*, 34:199–205, 1963.

P. Billingsley. *Probability and Measure*. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., Hoboken, NJ, 2012.

D. Blackwell and L. E. Dubins. An extension of Skorohod's almost sure representation theorem. *Proc. Amer. Math. Soc.*, 89(4):691–692, 1983.

L. Breiman, L. Le Cam, and L. Schwartz. Consistent estimates and zero-one sets. *Ann. Math. Statist.*, 35:157–161, 1964.

S. Cerreia-Vioglio, P. Ghirardato, F. Maccheroni, M. Marinacci, and M. Siniscalchi. Rational preferences under ambiguity. *Econom. Theory*, 48(2-3):341–375, 2011.

D. Corbae, M. B. Stinchcombe, and J. Zeman. *An Introduction to Mathematical Analysis for Economic Theory and Econometrics*. Princeton University Press, Princeton, NJ, 2009.

C. Deck and H. Schlesinger. Consistency of higher order risk preferences. *Econometrica*, 82(5):1913–1943, 2014.

E. Dekel, M. Siniscalchi, et al. Epistemic game theory. In H. P. Young and S. Zamir, editors, *Handbook of Game Theory with Economic Applications*, volume 4, chapter 12, pages 619–702. Elsevier, Amsterdam, 1 edition, 2015.

C. Dellacherie and P.-A. Meyer. *Probabilities and Potential*, volume 29 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1978.

R. M. Dudley. *Real Analysis and Probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.

M. Dumav and M. Stinchcombe. Skorohod's representation theorem for sets of probabilities. *Proc. Amer. Math. Soc.*, 144(7):3123–3133, 2016.

M. Dumav and M. Stinchcombe. Ambiguity aversion and the interpretation of analogies. *Working paper, Department of Economics, Carlos III de Madrid*, 2017.

L. Eeckhoudt and H. Schlesinger. Putting risk in its proper place. *The American Economic Review*, 96(1):280–289, 2006.

J. Eichberger, S. Grant, D. Kelsey, and G. A. Koshevoy. The α-meu model: a comment. *Journal of Economic Theory*, 146(4):1684–1698, 2011.

L. G. Epstein. A definition of uncertainty aversion. *Rev. Econom. Stud.*, 66(3):579–608, 1999.

P. Ghirardato and M. Marinacci. Risk, ambiguity, and the separation of utility and beliefs. *Mathematics of Operations Research*, 26(4):864–890, 2001.

P. Ghirardato, F. Maccheroni, and M. Marinacci. Differentiating ambiguity and ambiguity attitude. *J. Econom. Theory*, 118(2):133–173, 2004.

I. Gilboa and D. Schmeidler. Maxmin expected utility with nonunique prior. *J. Math. Econom.*, 18(2):141–153, 1989.

P. R. Halmos. *Naive set theory*. The University Series in Undergraduate Mathematics. D. Van Nostrand Co., Princeton, N.J.-Toronto-London-New York, 1960.

P. Klibanoff, M. Marinacci, and S. Mukerji. A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892, 2005.

M. J. Machina. Risk, ambiguity, and the rank-dependence axioms. *The American Economic Review*, 99(1):385–392, 2009.

C. N. Noussair, S. T. Trautmann, and G. Van de Kuilen. Higher order risk attitudes, demographics, and financial decisions. *Rev. Econom. Stud.*, 81(1):325–355, 2013.

W. Olszewski. Preferences over sets of lotteries. *Rev. Econom. Stud.*, 74(2):567–595, 2007.

L. J. Savage. *The Foundations of Statistics*. John Wiley & Sons, Inc., New York; Chapman & Hill, Ltd., London, 1954.

D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.

A. V. Skorohod. Limit theorems for stochastic processes. *Teor. Veroyatnost. i Primenen.*, 1:289–319, 1956.

M. B. Stinchcombe. Countably additive subjective probabilities. *Rev. Econom. Stud.*, 64(1):125–146, 1997.

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, New Jersey, 1944.

Proofs

Throughout, we use two results. The first is the Borel isomorphism theorem (Dellacherie and Meyer, 1978, Theorem III.20) or (Dudley, 2002, Theorem 13.1.1). A bijection between measurable spaces is a **measurable isomorphism** if it is measurable and its inverse is also measurable. Let B be a measurable subset of a complete separable metric space (M, d) and B' a measurable subset of a complete separable metric space (M', d'). The Borel isomorphism theorem says that for B and B' are measurably isomorphic if and only if they have the same cardinality.

The second result provides an extension of Skorohod (1956, Thm. 3.1.1). Blackwell and Dubins (1983) show that for any complete separable metric space (M, d) there exists a jointly measurable $\mathbf{b}_M : \Delta(M) \times [0,1] \to M$ such that

(i) for all $p \in \Delta(M)$, the image measure of λ under the function $\mathbf{b}_M(p, \cdot)$ is equal to p, that is, $\mathbf{b}_M(p, \lambda) = p$, and
(ii) if $\rho(p_n, p) \to 0$, then $\mathbf{b}_M(p_n, \cdot) \to \mathbf{b}_M(p, \cdot)$ almost everywhere λ, that is, $\lambda(\{x \in [0,1] : \mathbf{b}_M(p_n, x) \to \mathbf{b}_M(p, x)\}) = 1$.

We will also use the canonical set of priors, Π°, defined as the closed convex hull of the set of uniform distributions, λ_x, on the line segments $\{x\} \times [0,1] \subset [0,1] \times [0,1]$. Every element of Π° can be expressed as an integral of the λ_x with respect to some countably additive probability η on the set of λ_x, $x \in [0,1]$.

Proof of Theorem 1. Fix $A \in \mathcal{K}_{\Delta(\mathbb{X})}$ and let $\psi : [0,1] \leftrightarrow A$ be a measurable isomorphism. Define $f_A(s) = \mathbf{b}_{\mathbb{X}}(\varphi(f_\Delta(s)), f_\lambda(s))$ where $\mathbf{b}_{\mathbb{X}}(\cdot,\cdot)$ is the Blackwell-Dubins function. We verify that $f_A(\Pi) = A$ by showing the containment in two directions.

The first containment: $f_A(\Pi) \subset A$. For each $\mu \in \Pi$, $f_A(\mu) \in A$ because $\psi([0,1]) = A$, A is a closed convex set, and $f_\lambda(\mu) = \lambda$.

The other direction: $A \subset f_A(\Pi)$. Fix $p \in A$ and let $x = \psi^{-1}(p)$. By definition, there exists $\mu_x \in \Pi$ such that $f_\Delta(\mu_x)$ is the point mass δ_x. Since $f_\lambda(\mu_x) = \lambda$, $f_A(\mu_x) = p$. □

Proof of Corollary 1.1. Let Π' be a strongly zero-one set of non-atomic priors and let Π be its closed convex hull. Recall that for a random variable X taking values in $[0,1]$ and having continuous cdf $F_X(\cdot)$, the random variable $Y = F_X(X)$ has the uniform distribution.

From the characterization of strongly zero-one sets in Dumav and Stinchcombe (2016)[Theorem 1 and §3.2] we obtain simultaneously two measurable isomorphisms $\varphi : \Pi' \leftrightarrow [0,1]$ and $g : \Omega' \leftrightarrow [0,1]$ such that for all $r \in [0,1]$ and for all $\mu \in \varphi^{-1}(r)$, $\mu(g^{-1}(r)) = 1$. Consider the mapping $\mu \mapsto F_\mu(t) := \mu(g^{-1}([0,t])$ from Π' to the cdf of the distribution $g(\mu)$. Because μ is non-atomic and $g(\cdot)$ is a measurable isomorphism, the distribution $g(\mu)$ has a continuous cdf. The function $f_\lambda(\omega') := F_{\varphi^{-1}(g(\omega'))}(g(\omega'))$ is measurable, as it is a composite of two measurable functions, and for each $\mu \in \Pi'$, $f_\lambda(\mu) = \lambda$.

By the Borel isomorphism theorem, there exists a measurable bijection $\psi : \Pi \leftrightarrow [0,1]$. Define $f_\Delta(\omega) = \psi(\varphi(\omega))$. For each $\mu \in \Pi'$, $f_\Delta(\mu)$ is point mass on some $x \in [0,M]$. Since Π is the closed convex hull of Π', $f_\Delta(\Pi) = \Delta([0,1])$. □

Proof of Corollary 1.2. Let A, A_n be a sequence in $\mathcal{K}_{\Delta(\mathbb{X})}$. We will show the existence of a sequence of measurable functions, $h, h_n : [0,1] \to \Delta(\mathbb{X})$, with $h([0,1]) = A$, $h_n([0,1]) = A_n$ and $h_n(r) \to h(r)$ for each $x \in [0,1]$. Given such a sequence of functions, define $f(s) = \mathbf{b}_{\mathbb{X}}(h(f_\Delta(s)), f_\lambda(s))$ and $f_n(s) = \mathbf{b}_{\mathbb{X}}(h_n(f_\Delta(s)), f_\lambda(s))$, where $\mathbf{b}_{\mathbb{X}}(\cdot,\cdot)$ is the Blackwell-Dubins function. Because $[\rho(p_n, p) \to 0]$ implies that $\mathbf{b}_{\mathbb{X}}(p_n,\cdot)$ converges to $\mathbf{b}_{\mathbb{X}}(p,\cdot)$ almost everywhere λ and $\rho_H(A_n, A) \to 0$, the result follows.

To define the sequence h, h_n with the requisite properties, let $\psi : [0,1] \leftrightarrow [0,1]^{\mathbb{N}}$ be a measurable isomorphism, and let $\{v_n : n \in \mathbb{N}\}$ be a countable sup-norm dense set of continuous functions mapping $\mathbb{X}$ to $[0,1]$. The identification $p \leftrightarrow \{\int v_n\, dp : n \in \mathbb{N}\}$ defines a linear homeomorphism between $\Delta(\mathbb{X})$ and a compact convex $K \subset [0,1]^{\mathbb{N}}$. Meterize the product topology on $[0,1]^{\mathbb{N}}$ with a strictly convex metric (e.g. the ℓ_2 norm). Define the mapping $n(A, v)$ as the nearest point in A to v. Because the metric is strictly convex, this mapping is jointly continuous. Finally, define $h(r) = n(A, \psi(r))$ and $h_n(r) = n(A_n, \psi(r))$. □

Proof of Lemma 2. Let p_n be a sequence in A with $\rho(p_n, p) \to 0$. We show that $p \in A$. Let $x_n \in [0,M]$ satisfy $p_n(\{x_n\}) \geq \boldsymbol{a}$. Any subsequence of p_n has a further subsequence, still converging to p, for which $x_n \to x$ for some $x \in [0,M]$. By the definition of $\rho(\cdot,\cdot)$, $p(\{x\}) \geq \boldsymbol{a}$. If the closed set A had an interior, it would have to contain a probability having a density with respect to Lebesgue measure, contradicting the existence of an atom. □

Proof of Theorem 2. Because the composition of measurable functions is measurable and $\mathbb{F}$ fails to be open-minded, no open-minded subset of $\mathcal{K}_{\Delta(\mathbb{X})}$ can belong to $\mathcal{R}(\mathbb{F})$. It is therefore sufficient to show that the class of open-minded sets is dense in $\mathcal{K}_{\Delta(\mathbb{X})}$. To this end, pick arbitrary $A \in \mathcal{K}_{\Delta(\mathbb{X})}$ and $\epsilon > 0$. Let A_f denote a finite set of extreme points for A such that $d(A, \mathrm{co}(A_f)) < \epsilon/2$. For any $\delta > 0$ and $x \in \mathbb{X}$, let $B_\delta(x)$

denote the necessarily uncountable, open ball with radius $\delta > 0$ around $x \in \mathbb{X}$. By the Borel isomorphism theorem, there exists $\varphi_{x,\delta} : [0,1] \times [0,1] \leftrightarrow B_\delta(x)$ where $\varphi_{x,\delta}$ is a measurable bijection with measurable inverse. Let $\Pi_{x,\delta}$ denote the open-minded set $\varphi_{x,\delta}(\Pi^\circ)$ where Π° is the canonical open-minded set of priors given above. Pick $\delta < \epsilon/2$ such that the points in A_f are at least 2δ from each other. Since the support sets are disjoint, the closed convex hull of the set $\cup_{x\in A_f}\Pi_{x,\delta}$ is open-minded and within ϵ of A. □

The following argument closely parallels the proof for countably additive priors, but unlike the countably additive case, it does not extend to complete separable metric spaces.

Proof of Lemma 3. Fix $p \in \Delta^{fa}(\mathbb{X})$. Let $\mathcal{E}_n$ be a nested sequence of measurable partitions of $\mathbb{X}$ into elements having maximal diameter less than $1/2^n$. Construct a corresponding sequence of nested partitions of S, $\mathcal{A}_n$ having $\mu(A_n) = p(E_n)$ for each $E_n \in \mathcal{E}_n$. Pick $x_{k,n} \in E_{k,n} \in \mathcal{E}_n$ and define $f_n = \sum_k x_{k,n}1_{E_{k,n}}$. For each s, $f_n(s)$ is a Cauchy sequence, hence converges. The function $f(s) := \lim_n f_n(s)$ is measurable, we must show that $f(\mu)$ is continuously equivalent to p. Let $v : \mathbb{X} \to \mathbb{R}$ be a continuous, hence bounded, function. Being the uniform limit of the f_n, f satisfies $\int_S v(f(s))\, d\mu(s) = \int_{\mathbb{X}} v(x)\, dp(x)$ for all continuous $v : \mathbb{X} \to \mathbb{R}$. □

Proof of Theorem 3. Let $\{\mu_n : n \in \mathbb{N}\}$ be the countably infinite set of non-atomic probabilities supported on with disjoint supports, let $A \in \mathcal{K}_{\Delta^{fa}(\mathbb{X})}$, and let $\{p_n : n \in \mathbb{N}\}$ be a countable dense subset of A. From Lemma 3, there exists a measurable $f : S \to \mathbb{X}$ with $f(\mu_n) = p_n$. Taking weak* closure in $\Delta^{fa}(\mathbb{X})$ and using the continuity of the $\mu \mapsto f(\mu)$ mapping, $\rho_H(f(\Pi), A) = 0$. For the second statement, let A' be the set of countably additive probabilities on $\mathbb{X}$ at distance 0 from A. Pick a measurable $f : S \to \mathbb{X}$ such that $f(\Pi^{ca}) = A'$. Again, taking weak* closure in $\Delta^{fa}(\mathbb{X})$ and using the continuity of the $\mu \mapsto f(\mu)$ mapping, $\rho_H(f(\Pi), A) = 0$. □

Proof of Proposition 1. Let $A = f(S)$ and $B = g(S)$. Define the non-degenerate interval $[a_f, b_f] = \{\int v\, dp : p \in A\}$ and $[a_g, b_g] = \{\int v\, dq : q \in B\}$. We striate A and B as follows: for $u \in [a_f, b_f]$, define $A_u = \{p \in A : \int v\, dp = u$; for $u \in [a_g, b_g]$, define $B_u = \{q \in B : \int v\, dq = u$. It is easy to show that there exists a jointly measurable $m_f : [0,1]\times[0,1] \to A$ such that for all $x \in [0,1]$, $m_f(x, \cdot)$ is a measurable isomorphism between $[0,1]$ and $A_{a_f+(b_f-a_f)x}$. In a similar fashion, there exists a jointly measurable $m_g : [0,1]\times[0,1] \to B$ such that $m_g(x, \cdot)$ is a measurable isomorphism between $[0,1]$ and $B_{a_g+(b_g-a_g)x}$. For later purposes, note that $x \mapsto a_f + (b_f - a_f)x$ and $x \mapsto a_g + (b_g - a_g)x$ are positive affine transformations of each other.

Because Π is open-minded, there exists $h_\lambda : S \to [0,1]$ with $h_\lambda(\mu) = \lambda$ for each $\mu \in \Pi$, and there exists $h_D : S \to [0,1] \times [0,1]$ such that $h_D(\Pi) = \Delta([0,1] \times [0,1])$. Define

$$f'(s) = \mathbf{b}_{\mathbb{X}}(m_f(h_D(s)), h_\lambda(s)) \text{ and } g'(s) = \mathbf{b}_{\mathbb{X}}(m_g(h_D(s)), h_\lambda(s)).$$

For each $\mu \in \Pi$, $f'(\mu)$ is a convex combination of probabilities in A and $g'(\mu)$ is a convex combination of probabilities in B. Since A and B are convex, $f'(\Pi) \subset A$ and $g'(\Pi) \subset B$. For any $(x, y) \in [0,1] \times [0,1]$, let $\mu_{(x,y)} \in \Pi$ put mass 1 on $h_D^{-1}(x, y)$ (because $\delta_{(x,y)} \in \Delta([0,1] \times [0,1])$, there exists such a $\mu_{(x,y)}$). Each $p \in A$ and $q \in B$ is of the form $f'(\mu_{(x,y)})$ and $g'(\mu_{(x,y)})$ respectively, so $A \subset f'(\Pi)$ and $B \subset g'(\Pi)$. Finally $v_{f'}(\cdot)$ is a positive affine transformation of $v_{g'}(\cdot)$ because $x \mapsto a_f + (b_f - a_f)x$ and $x \mapsto a_g + (b_g - a_g)x$ are positive affine transformations of each other. □

Proof of Proposition 2. Let $[a,b] = \{\int v \circ f\, d\mu : \mu \in \Pi\} = \{\int v \circ g\, d\mu : \mu \in \Pi\}$. The only increasing affine transformation of $[a,b]$ with itself is the identity. Apply Proposition 1. $\square$

Proof of Lemma 1. F contains $f(S)$ so that $\Delta(F)$ contains $f(\Delta(S))$. The extreme points in $\Delta(F)$ are the point masses δ_x, $x \in F$. Because both $\Delta(F)$ and $f(\Delta(S))$ are convex, it is sufficient to show that for every $\epsilon > 0$, there exists a probability in $f(\Delta(S))$ within ρ-distance of δ_x. Since $x \in F$, there exists a sequence s_n in S with $f(s_n) \to x$. Because singletons are measurable in S, $\Delta(S)$ contains δ_{s_n}, point mass on each s_n. Finally, $\rho(f(\delta_{s_n}), \delta_x) = \rho(\delta_{x_n}, \delta_x) \to 0$. The second assertion is immediate. $\square$

Proof of Proposition 3. Let Π be the canonical set of priors $\Pi^\circ \subset \Delta([0,1] \times [0,1])$ given above, and let $Q \in \Delta(\Pi)$ be the uniform distribution on $\{\lambda_x : x \in [0,1]\}$. Fix a $\mathfrak{p} \in \Delta(\Delta(\mathbb{X}))$ and let $\varphi : [0,1] \to \Delta(\mathbb{X})$ have the property that $\varphi(\lambda) = \mathfrak{p}$. Define $f : [0,1] \times [0,1] \to \mathbb{X}$ by $f(x,y) = \mathbf{b}_{\mathbb{X}}(\varphi(x), y)$ where, again, $\mathbf{b}_{\mathbb{X}} : [0,1] \times \Delta(\mathbb{X}) \to \mathbb{X}$ is the Blackwell-Dubins function. When λ_x is picked, $f(\lambda_x) = \varphi(x)$. Since Q picks the x according to λ, $\mathfrak{p}$ is the distribution of $f(\mu)$ when μ is distributed according to Q. $\square$

Department of Economics, Universidad Carlos III de Madrid, e-mail: mdumav@eco.uc3m.es, and Department of Economics, University of Texas, Austin, TX 78712-0301 USA, e-mail: max.stinchcombe@gmail.com