What is Ambiguity?

Massimiliano Amarante*

Université de Montréal et CIREQ

March 11, 2014

Abstract

The concept of Ambiguity designates those situations where the information available to the decision maker is insufficient to form a probabilistic view of the world. Thus, it has provided the motivation for departing from the Subjective Expected Utility (SEU) paradigm. Yet, the formalization of the concept is missing. This is a grave omission as it leaves non-expected utility models hanging on a shaky ground. In particular, it leaves unanswered basic questions such as: (1) Does Ambiguity exist?; (2) If so, which situations should be labeled as "ambiguous"?; (3) Why should one depart from Subjective Expected Utility (SEU) in the presence of Ambiguity?; and (4) If so, what kind of behavior should emerge in the presence of Ambiguity? The present paper fills these gaps. Specifically, it identifies those information structures that are incompatible with SEU theory, and shows that their mathematical properties are the formal counterpart of the intuitive idea of insufficient information. These are used to give a formal definition of Ambiguity and, consequently, to distinguish between ambiguous and unambiguous situations. Finally, the paper shows that behavior not conforming to SEU theory must emerge in correspondence of insufficient information and identifies the class of non-EU models that emerge in the face of Ambiguity. The paper also proposes a new comparative definition of Ambiguity, and discusses its relation with some of the existing literature.

Keywords: Non-Expected Utility, Information, Ambiguity

JEL classification: D81

*I benefited from comments and suggestions from Tzachi Gilboa, Fabio Maccheroni, Massimo Marinacci, Rich McLean and David Schmeidler. Financial support from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

1 Introduction

The past few years have witnessed an ever increasing number of applications of non-Expected Utility (non-EU) theories. This has been favored both by the recent theoretical developments and by the general motivation underlying most non-EU theories. On the one hand, the structure of many non-EU theories has been greatly clarified (see [12] for a recent, comprehensive survey), appropriate statistical tools have become available (see, for instance, [16], [18], [28]) and extensions to dynamic settings have been successfully pursued ([8], [14], [26]). On the other hand, non-EU theories seem to provide the right framework for dealing with problems that lie at the core of Economics and Finance such as Macroeconomics Policy, Investment choice, Entrepreneurship and Innovation, Portfolio choice etc. The common thread linking all these problems is the high level of uncertainty surrounding them and, consequently, they all reduce to choice problems in the presence of very limited information.

The idea that behavior in situations of limited information should be qualitatively different from behavior in situations where information abounds is certainly not a new one (dating back at least to F. Knight), and has been the main inspirational motive behind the theoretical work in decision making under uncertainty. For instance, Marinacci [19, p. 1] motivates his work on the multiple-prior model by saying "The basic idea ... is simple and appealing: since the decision maker has not enough information to form a meaningful simple prior, he uses a set of priors, consisting of all priors compatible with his limited information".

Intuitive arguments of this sort have not found, however, a formal counterpart. This is a grave omission because it leaves non-EU models hanging on shaky ground. In particular, it leaves unanswered basic questions such as: (1) Does Ambiguity exist?; (2) If so, which situations should be labeled as "ambiguous"?; (3) Why should one depart from Subjective Expected Utility (SEU) in the presence of Ambiguity?; and (4) If so, what kind of behavior should emerge in the presence of Ambiguity? The present paper fills these gaps. We will begin by studying the set of all possible information structures that might be available to a decision maker. We will show that this consists of two types: those that are compatible with SEU theory and those that are incompatible with it. As we shall see, the mathematical properties of the latter provide the formal counterpart of the intuitive idea of insufficient information. We will use these properties to give a formal definition of Ambiguity and, consequently, to distinguish between ambiguous situations and unambiguous ones. We will, then, show that behavior not conforming to SEU theory emerges in correspondence of insufficient information, thus effectively providing the sought after informational foundation for non-EU theories. Finally, we will identify the class of non-EU models that emerge in the face of Ambiguity.

1.1 Paper outline

Our first step consists of characterizing the environment within which our inquiry will take place. We do so in Section 2 by determining, figuratively speaking, the point where non-EU theories depart from SEU theory. Starting with Section 3, we begin to explore the idea that this departure might be due to the poor quality of the information available to the decision maker. The study of Information is the subject matter of Sections 3 to 5. The main result is Theorem 8, which states that in correspondence of certain information structures the decision maker's behavior cannot conform to SEU theory. In Section 6, we give a few examples of such information structures. In Section 7, we discuss the intuition behind their mathematical properties, thus showing that they are the formal counterpart of the intuitive idea of insufficient information. This section completes the first part of our program, that of showing that Ambiguity exists and that one must depart from SEU theory in the presence of Ambiguity. These findings are summarized in Section 8 by giving a formal definition of Ambiguity. With the subsequent section, we move to the second part of our program, that of identifying the types of non-EU behavior that emerge as a response to Ambiguity. In Section 9, we isolate the set of all acts that the decision maker can evaluate on the basis of his information. We call this set the set of subjectively measurable acts. The problem of identifying the types of non-EU behavior that emerge in the face of Ambiguity takes the form of extending the decision maker's preference functional from the set subjectively measurable acts to the set of all acts *in a way that respects the decision maker's information.* The main result is Theorem 27 of Section 10 which identifies the class of these non-EU behaviors. In Section 11, we study the set of "predictives" in non-EU theories and its relation with the indicators of the Ambiguity perceived by the decision maker. In the process, we propose a new comparative definition of Ambiguity, and study its relation with that proposed by Ghirardato et al. in [10].

2 Background and notation

The intuitive idea behind the study of non-EU models is that the decision maker would conform to SEU theory if the information is, in some sense, good and would depart from it if the information is not good. Our strategy to pursue this idea is as follows. If this intuition is correct, we should be able to find, imaginatively speaking, a point in the theory where the determination of whether or not to conform to SEU theory has not been made yet. At that point, we would then plug in the information available to the decision maker, and obtain SEU when we plug in good information and non-EU when we plug in information that is not good. In this section, we begin our inquiry by looking for this point.

Recall that, following Savage [24], the alternatives available to the decision maker are modeled as mappings $(S, \Sigma) \longrightarrow X$, where (S, Σ) is a measurable space of states of the world and X is a space of consequences. Let $\mathcal{A}$ denote the set of all alternatives, and let $\mathcal{A}_c$ be that of constant alternatives, that is of constant mappings $(S, \Sigma) \longrightarrow X$. Assume that X is a mixture space (see [3] and [11]), and let $\succsim$ denote the decision maker's preference relation over $\mathcal{A}$. In [10], Ghirardato, Maccheroni and Marinacci isolated a core common to several theories of decision making. This consists of the five axioms listed below.

A1 $\succsim$ is complete and transitive.

A2 (C-independence) For all $f, g \in \mathcal{A}$ and $h \in \mathcal{A}_c$ and for all $\alpha \in (0, 1)$

$$f \succ g \qquad \Longleftrightarrow \qquad \alpha f + (1-\alpha)h \succ \alpha g + (1-\alpha)h$$

A3 (Archimedean property) For all $f, g, h \in \mathcal{A}$, if $f \succ g$ and $g \succ h$ then $\exists \alpha, \beta \in (0, 1)$ such that $\alpha f + (1-\alpha)h \succ g$ and $g \succ \beta f + (1-\beta)h$.

A4 (Monotonicity) For all $f, g \in \mathcal{A}$, $f(s) \succsim g(s)$ for any $s \in S \Longrightarrow f \succsim g$.

A5 (Non-degeneracy) $\exists x, y \in X$ such that $x \succ y$.

Then, Ghirardato, Maccheroni and Marinacci observed that alternative sixth axioms correspond to alternative theories of decision making. For instance, one obtains SEU, CEU and MEU as follows:

A6 (a) (SEU, Anscombe and Aumann [3]) For all $f, g \in \mathcal{A}$ such that $f \sim g$, $\frac{1}{2}f + \frac{1}{2}g \sim f$;

A6 (b) (CEU, Schmeidler [25]) For all $f, g \in \mathcal{A}$ such that $f \sim g$, $\frac{1}{2}f + \frac{1}{2}g \sim f$ if f and g are comonotonic;

A6 (c) (MEU, Gilboa and Schmeidler [13]) For all $f, g \in \mathcal{A}$ such that $f \sim g$, $\frac{1}{2}f + \frac{1}{2}f \succsim f$.

Thus, it appears that the point we are looking for lies at the end of the 5th axiom and before the 6th is imposed. The mathematical environment associated with the first five axioms is completely characterized by Theorem 1 below, which was proved in [2, Theorems 1 and 2]. In order to state it, we need to introduce the notation that we will be using throughout the paper.

Notation: The set of bounded, Σ-measurable functions $(S, \Sigma) \longrightarrow \mathbb{R}$ equipped with the sup-norm is denoted by $B(\Sigma)$. Its dual $ba(\Sigma)$, the space of bounded charges on Σ, is always endowed with the weak*-topology produced by the duality $(ba(\Sigma), B(\Sigma))$. The subset of $ba(\Sigma)$ consisting of the finitely additive probability measures on Σ is denoted by $ba_1^+(\Sigma)$. For $\mathcal{C}$ a weak*-compact, convex subset of $ba_1^+(\Sigma)$, a weak*-continuous affine function $\mathcal{C} \longrightarrow \mathbb{R}$ is of the form $\psi_f(P) = \int\limits_S f dp$, $P \in \mathcal{C}$, for some $f \in B(\Sigma)$. The space of all weak*-continuous affine functions on $\mathcal{C}$ equipped with the sup-norm is denoted by $A(\mathcal{C})$. The mapping $\kappa : f \longmapsto \psi_f$ is the canonical linear mapping $\kappa : B(\Sigma) \longrightarrow A(\mathcal{C})$. The Borel σ-algebra on $\mathcal{C}$ is denoted by $\mathcal{B}$, and $B(\mathcal{B})$ denotes the space of bounded, $\mathcal{B}$-measurable functions $\mathcal{C} \longrightarrow \mathbb{R}$ equipped with the sup-norm. Finally, the set of regular Borel measures on $\mathcal{C}$ is denoted by $\mathcal{P}(\mathcal{C})$.

We can now state the theorem characterizing Axioms 1 to 5. Recall that Axioms 1 to 5 imply that there exist a utility function $u : X \longrightarrow \mathbb{R}$ and a functional $I : B(\Sigma) \longrightarrow \mathbb{R}$ such that for $\tilde{f}, \tilde{g} \in \mathcal{A}$ (see [13] and [10], for details)

$$\tilde{f} \succsim \tilde{g} \qquad iff \qquad I(u \circ \tilde{f}) \geq I(u \circ \tilde{g})$$

For notational simplicity, throughout the paper we are going to identify an act $\tilde{f} \in \mathcal{A}$ with the corresponding function $u \circ \tilde{f} = f \in B(\Sigma)$.

Theorem 1 (Amarante [2]) *A preference relation $\succsim$ on $\mathcal{A}$ satisfies Axioms 1 to 5 iff for any $f \in B(\Sigma)$ the functional I representing it can be written as*

$$I(f) = \int\limits_{\mathcal{C}} \kappa(f) d\nu = \int\limits_{\mathcal{C}} \int\limits_{S} f dP d\nu(P)$$

where $\mathcal{C}$ is a convex, weak-compact subset of $ba_1^+(\Sigma)$ and ν is a capacity on the Borel subsets of $\mathcal{C}$.*

Intuitively, Theorem 1 tells us that any behavior satisfying Axioms 1 to 5 corresponds to an integration over priors when this operation is performed in the sense of Choquet.

3 Information

In line with Bayesian statistics, a decision maker who follows Axioms 1 to 5 entertains several probabilistic descriptions of the world, each represented by a probability in the set $\mathcal{C}$. He, then, weights these probabilistic descriptions by using a possibly non-additive set function ν, thus obtaining a possibly non-probabilistic criterion for evaluating his alternatives. As we have seen, it is at this point that we have to plug in information, and determine whether or not the decision maker can be Bayesian. As it is customary, we are going to model information as a partition of the set $\mathcal{C}$. The partition and the associated sub σ-field (that is, the σ-field generated by the partition) convey that the decision maker has only partial information about $\mathcal{C}$. This corresponds to the following situation (see Billingsley [5], pp. 57-58 and pp. 427-29): on the basis of his information, the decision maker can construct a statistical experiment whose outcome would tell him (in a statistical sense) in which element of the partition the true probabilistic description lies. He would not be able, however, to construct *on the basis of his information* an experiment capable of distinguishing among probabilistic descriptions lying in the same cell of the partition. While this is a valuable interpretation, the reader should be cautioned that it is subject to the qualifications that we will discuss in Sections 5 and 7. Aside from information about the true probabilistic description of the world, the decision maker may also have information about the true state of the world. Clearly, this type of information (that is, information of the form "the true state belongs to the set $A \subset S$") can always be trivially expressed as information about the set $\mathcal{C}$.[1] The following definition records formally the concept of information structure.

Definition 2 *An information structure on $(\mathcal{C}, \mathcal{B})$ is a triple $\{(\mathcal{C}, \mathcal{B}), \mathcal{I}, \mathcal{B}_{\mathcal{I}}\}$, where $\mathcal{I}$ is a partition of $\mathcal{C}$ and $\mathcal{B}_{\mathcal{I}}$ is the sub-field of $\mathcal{B}$ generated by $\mathcal{I}$.*

[1] It is interesting to notice that departures from SEU are often associated with the decision maker being given information explicitly about the set $\mathcal{C}$ of probabilities on (S, Σ). This is the case, for instance, in both Ellsberg's experiments (the configurations of the urns) as well as in those of Gardenfors and Sahlin [9] (the ability of the players).

Before we begin our inquiry into the properties of possible information structures, a refinement of this definition is needed. We have been pursuing the idea that at the point we are at - the one ideally lying between Axiom 5 and Axiom 6 - the decision maker tries to be Bayesian: it is only the type of information that he gets at this point that might prevent him from being so. But, if the information is good enough he will adhere to SEU. It is clear that for this to happen, the decision maker must have a probability on the set $\mathcal{C}$. Informally, this corresponds to the integration over priors argument that in traditional Bayesian statistics leads to the determination of the "predictive" (see [20] and [2, Sec. 3.1]): Given his prior μ on $\mathcal{C}$ and the partition $\mathcal{I}$, the decision maker computes a collection of conditional probabilities, one for each element of the partition; then, he averages these conditionals with the weights that μ gives to the corresponding elements of the partition, and SEU obtains. Thus, the existence of the prior μ is a necessary condition for the decision maker to be Bayesian. Since this prior also contains some form of information available to the decision maker, it must be explicitly encoded in the formal definition of information structure:

Definition 3 *A Bayesian information structure on* $(\mathcal{C}, \mathcal{B})$ *is a quadruple* $\{(\mathcal{C}, \mathcal{B}), \mu, \mathcal{I}, \mathcal{B}_{\mathcal{I}}\}$*, where* μ *is a regular Borel measure on* $(\mathcal{C}, \mathcal{B})$*,* $\mathcal{I}$ *is a partition of* $\mathcal{C}$ *and* $\mathcal{B}_{\mathcal{I}}$ *is the sub-field of* $\mathcal{B}$ *generated by* $\mathcal{I}$*.*

Possibly, if the Ambiguity idea holds, the existence of a prior on $\mathcal{C}$ is not sufficient for SEU to obtain, but that remains to be determined. In the meantime, let it be clear that no extra assumption has been made with Definition 3. In particular, it has not been assumed that the capacity in Theorem 1 is actually a measure. As said, the introduction of Bayesian information structures as in Definition 3 is necessary to guarantee that SEU would obtain whenever the information is good. In fact, by using Definition 3 and noticing that no information explicitly appears either in Theorem 1 or in any of the classical representation theorems in decision making under uncertainty, we can now (informally) re-formulate our working hypothesis as *If the Ambiguity idea is correct, then Bayesian integration over priors with bad information must be representable by a non-additive integral.* In other words, the poor quality of the information must reveal itself into the non-additivity of the integral representing the decision maker's preference.

Remark 4 (★) *Later in the paper, Corollary 30 will formally prove that the intuition*

Lebesgue integration w/ bad information = non-additive integration

is, indeed, correct.

Another brief comment concerns the requirement that the measure in Definition 3 be regular Borel. This is motivated by the structure of the problem and is without loss of generality.[2] In fact, by allowing for a wider class of measures, we would only strengthen our results (see footnote 5, Sec. 8).

4 Sometimes, the information is ....

Given a Bayesian information structure $\{(\mathcal{C}, \mathcal{B}), \mu, \mathcal{I}, \mathcal{B}_{\mathcal{I}}\}$, the integration over priors procedure supposedly leads to behavior conforming to SEU theory. For this to be true, two conditions must be met. Let $\{\mu_i\}_{i \in \mathcal{I}}$ be a system of probability measures, one for each element of the partition $\mathcal{I}$. Then, $\forall \varphi \in B(\mathcal{B})$ we must have:

1. The function $\psi(i) : \mathcal{C}/\mathcal{I} \longrightarrow \mathbb{R}$ defined by

$$\psi(i) = \int_{\mathcal{C} \cap i} \varphi \mid_{\iota} d\mu_{\iota}$$

is a measurable function with respect to the canonical σ-field on $\mathcal{C}/\mathcal{I}$ (see Appendix A); and

2.

$$\int_{\mathcal{C}} \varphi d\mu = \int_{\mathcal{C}/\mathcal{I}} \int_{\iota} \varphi \mid_{\iota} d\mu_{\iota} d\mu'$$

where μ' is the pushforward of μ under the canonical projection $\pi : \mathcal{C} \longrightarrow \mathcal{C}/\mathcal{I}$ (see Appendix A).

[2] When the preference is SEU, the functional on $B(\Sigma)$ which represents it is linear and sup-norm continuous. Thus, the integration over priors argument must define a sup-norm continuous, linear functional defined on the space $\mathcal{A}(\mathcal{C})$ of continuous affine affine functions on $\mathcal{C}$. By Hahn-Banach, this functional can be extended to a sup-norm continuous linear functional on $C(\mathcal{C})$, the Banach space of all continuous functions on $\mathcal{C}$ equipped with the sup-norm, and (via the Riesz representation theorem) there exists a unique regular Borel measure representing it, that is $\nu \in \mathcal{P}(\mathcal{C})$.

The exact formal meaning of these conditions is fully spelled out in Appendix A. They are, however, nothing other than the formal transposition of the integration over priors argument. To see this, let us start from condition 2. The part stating that μ' is the pushforward of μ under π simply means that the weights assigned to the cells of the partitions are those determined by μ. Next, let us consider the inner integral on the RHS of condition 2. This is exactly the function ψ that appears in Condition 1. If the family $\{\mu_i\}_{i\in\mathcal{I}}$ has to represent a family of conditional probabilities, then the function $\psi(i) : \mathcal{C}/\mathcal{I} \longrightarrow \mathbb{R}$ describes precisely the association *cell of the partition* $\longrightarrow$ *conditional evaluation of* φ *at that cell of the partition.* Then, the equality in Condition 2. states that things "add up" properly: that is, if for any $\varphi \in B(\mathcal{B})$ we take all of its conditional evaluations and average them with the weights determined by μ, we obtain the unconditional evaluation of φ. This condition is even more transparent when we take $\varphi = \chi_A$, the indicator function of a set $A \in \mathcal{B}$. In such a case, the condition states that if we take A, "cut" it by using the elements of the partition, measure the pieces separately and then add them up, we obtain the original measure $\mu(A)$. To complete the assessment of Conditions 1. and 2., notice that Condition 1. is necessary to even state Condition 2. as in order to take the integral on the RHS of 2., the function that is being integrated has, obviously, to be measurable. Condition 1., however, has a very substantial meaning as well: *it expresses that the type of information available to the decision maker is sufficient both to evaluate his options in the various contingencies he can distinguish (the cells of the partition) and to understand how these evaluations relate to one another.* We conclude this section by giving one example of a partition (hence, of an information structure) for which the two conditions hold.

Example 5 *Let $\mathcal{I}$ be the partition of $\mathcal{C}$ generated by the equivalence relation*

$$P \sim Q \qquad iff \qquad P = Q$$

and let $\{(\mathcal{C}, \mathcal{B}), \mu, \mathcal{I}, \mathcal{B}_{\mathcal{I}}\}$ be the corresponding information structure. In such a case, the RHS of 2. is given by $\int_{\mathcal{C}} \varphi d\mu$ and thus 2. becomes a tautology. Moreover, for this partition, we have $\mathcal{C} = \mathcal{C}/\mathcal{I}$ and $\varphi \in B(\mathcal{B})$ is the same as saying that 1. is satisfied. Thus, each $\varphi \in B(\mathcal{B})$ is evaluated by

$$\int_{\mathcal{C}} \varphi d\mu$$

Finally, the evaluation of each $f \in B(\Sigma)$ is achieved by integrating over priors, that is by means of the functional

$$I(f) = \int_{\mathcal{C}} \kappa(f) d\mu = \int_{\mathcal{C}} \int_{S} f dP d\mu(P)$$

On the weak-compact, convex set $\mathcal{C}$, each μ has a unique barycenter [22, Proposition 1.1] $P^* \in \mathcal{C}$, and we have that*

$$I(f) = \int_{\mathcal{C}} \kappa(f) d\mu = \kappa(f)(P^*) = \int_{S} f dP^*$$

for every $f \in B(\Sigma)$, which is the SEU functional.

Obviously, this conclusion is far from being surprising. The partition in the example is the finest possible partition of $\mathcal{C}$, hence it represents the best possible information for the decision maker. We should have expected the Bayesian integration over priors argument to go through at least in this case. For future reference, we record this formally in the corollary below.

Corollary 6 *If the information available to a Bayesian decision maker is the best possible, then behavior conforming to SEU theory obtains.*

5 ... not enough to rely on a single probability

Now, we have to address the question of whether or not there exist information structures in correspondence of which the integration over priors procedure fails to lead to behavior conforming to SEU theory. Theorem 8 of this section answers the question in the affirmative. In what follows, we are going to make an assumption, which greatly simplifies the exposition. We are going to assume that the set of priors $\mathcal{C}$ of Theorem 1 is a Polish space. This assumption is fairly minor (for instance, it is automatically satisfied any time that $\mathcal{C}$ is finite dimensional) and it is possible to dispense with it but, as said, at the price of a cumbersome exposition. At any rate, in order to dissipate any doubt about the axiomatic foundations of our work, in Appendix C we show that the assumption is satisfied whenever the decision maker's preference relation satisfies the axiom of Monotone Continuity (see Appendix C).[3]

[3] The original monotone continuity axiom was introduced by Arrow in [4], who comments "the assumption of Monotone Continuity seems, I believe correctly, to be the harmless simplification almost inevitable in the formalization of any real-life problem"

The key concept in Theorem 8 below is that of Rokhlin's non-measurable partition [23]. We stress that the concept refers to a property of the partition as a whole and not to a property of the sets making up the partition which might as well be measurable sets.

Definition 7 (Rokhlin [23]) *Let (L, Λ, λ) be a Lebesgue space (see Appendix B), and let $\mathcal{I}$ be a partition of L. Let the quotient $L/\mathcal{I}$ be endowed with the measure structure induced by the canonical projection (see Appendix A). The quotient $L/\mathcal{I}$ is said to be countably separated if there exists a countable family of measurable subsets of $L/\mathcal{I}$ which separates points. The partition $\mathcal{I}$ is called measurable if $L/\mathcal{I}$ is countably separated.*

Theorem 8 *Let $\{(\mathcal{C}, \mathcal{B}), \mu, \mathcal{I}, \mathcal{B}_{\mathcal{I}}\}$ be a Bayesian information structure on $(\mathcal{C}, \mathcal{B})$. Assume that the prior μ is not purely atomic and that μ is not supported by a single cell of the partition. Then, SEU obtains if and only if the partition $\mathcal{I}$ of $\mathcal{C}$ is measurable.*

Notice that the assumption that μ is not supported by a single cell of the partition is clearly necessary for the conclusion in the theorem. In fact, if μ is supported by a single cell, then (modulo sets of μ-measure 0) the partition consists of a single element, and we are back to the case examined at the end of the previous section where SEU obtains trivially.

Proof. By the assumption that μ is not purely atomic, μ can be expressed as the product of a purely atomic measure and a non-atomic one. Since a system of conditional measures of a purely atomic measure always exists, we can assume without loss that μ is non-atomic. If $\mathcal{I}$ is measurable, then by Rokhlin's Theorem [23] there exists a canonical system of conditional probabilities $\{\mu_\iota\}_{\iota \in \mathcal{I}}$. By using Definition 42 (Appendix A), it is straightforward to check that for every $\varphi \in B(\mathcal{B})$ both Conditions 1. and 2. of Section 4 are satisfied. Thus, every $f \in B(\Sigma)$ is evaluated by $\int\limits_{\mathcal{C}} \kappa(f) d\mu$ and the SEU functional obtains by means of the barycenter argument exactly as in the example in Section 4.

Conversely, let $\mathcal{I}$ be a nonmeasurable partition, and let $\{\mu_\iota\}_{\iota \in \mathcal{I}}$ be a system of conditional probabilities, with each μ_ι a non-atomic measure on ι. By Rokhlin's theorem, $\{\mu_\iota\}_{\iota \in \mathcal{I}}$ cannot be canonical. Hence, $\exists \varphi \in B(\mathcal{B})$ such that at least one of Conditions 1. and 2. of Section 4 is violated. If such a φ belongs to *range* $\kappa(B(\Sigma))$, then we are done for in such a case there exists at least one $f \in B(\Sigma)$ that cannot be evaluated by a SEU functional. Now, we are going to show that *range* $\kappa(B(\Sigma))$ necessarily contains at least one such φ.

To begin, observe that the (non-canonical) system of conditional probabilities $\{\mu_\iota\}_{\iota\in\mathcal{I}}$ defines an operator $\tilde{T} : B(\mathcal{B}) \longrightarrow \mathbb{R}^{\mathcal{C}/\mathcal{I}}$ by

$$\psi \mapsto \tilde{T}(\psi) \qquad \textit{where} \qquad \tilde{T}(\psi)(\iota) = \int_{\mathcal{C}} \psi d\mu_\iota$$

Also, observe that $\mathrm{supp}\mu_\iota \subset \iota$. Let

$$\Theta = \left\{ \psi \in B(\mathcal{B}) \mid (a)\ \tilde{T}(\psi) \in B(\mathcal{B}/\mathcal{I});\ (b) \int_{\mathcal{C}} \psi d\mu = \int_{\mathcal{C}/\mathcal{I}} \int_{\iota} \psi \mid_\iota d\mu_\iota d\mu' \right\}$$

By using standard arguments, it is easily checked (see for instance [1], Ch. 13) that Θ is a linear subspace and a lattice. Now, let $\{\psi_n\}_{n\in\mathbb{N}}$ be a sequence in Θ;

CLAIM: If either $\psi_n \nearrow \psi \in B(\mathcal{B})$ or $\psi_n \searrow \psi \in B(\mathcal{B})$, then $\psi \in \Theta$.

Proof of the claim: Let $\psi_n \nearrow \psi \in B(\mathcal{B})$.

(a) By the Dominated Convergence Theorem (DCT), for every μ_ι we have $\int_{\mathcal{C}} \psi_n d\mu_\iota \nearrow \int_{\mathcal{C}} \psi d\mu_\iota$, that is $\tilde{T}(\psi_n) \nearrow \tilde{T}(\psi)$. Hence, $\tilde{T}(\psi)$ is a pointwise limit of measurable functions, and hence measurable. Moreover, since $\psi \in B(\mathcal{B})$, $\tilde{T}(\psi)$ is bounded, i.e. $\tilde{T}(\psi) \in B(\mathcal{B}/\mathcal{I})$.

(b) Observe that

$$\begin{aligned}
\int_{\mathcal{C}} \psi d\mu &= \lim_{n\to\infty} \int_{\mathcal{C}} \psi_n d\mu && \textit{(by the DCT and } \psi \in B(\mathcal{B})) \\
&= \lim_{n\to\infty} \int_{\mathcal{C}/\mathcal{I}} \int_{\iota} \psi_n d\mu_\iota d\mu' && \text{(because } \psi_n \in \Theta) \\
&= \lim_{n\to\infty} \int_{\mathcal{C}/\mathcal{I}} \tilde{T}(\psi_n) d\mu' \\
&= \int_{\mathcal{C}/\mathcal{I}} \tilde{T}(\psi) d\mu' && \textit{(by (a) and the DCT)}
\end{aligned}$$

$$= \int_{\mathcal{C}/\mathcal{I}} \int_{\iota} \psi \mid_\iota d\mu_\iota d\mu'$$

which completes the proof for the case $\psi_n \nearrow \psi$. The other case is similar.

Now suppose, by the way of contradiction, that *range* $\kappa(B(\Sigma)) \subset \Theta$. Let K denote the set of continuous, convex functions on $\mathcal{C}$. Then, if $\gamma \in K$ there exists ([22], p. 19) $\{\zeta_m\}_{m\in\mathbb{N}} \subset A(\mathcal{C}) \subset$ *range* $\kappa(B(\Sigma))$ and a sequence $\{\eta_n\}_{n\in\mathbb{N}}$, with $\eta_n = \wedge\{\zeta_i\}_{i=1}^k$, such that $\eta_n \nearrow \gamma$. The sequence $\{\eta_n\} \subset \Theta$ because Θ is a lattice. Then, by the above claim, $\gamma \in \Theta$, that is $K \subset \Theta$. Since Θ is a linear space, it follows that $K - K \subset \Theta$. By the Stone-Weierstrass theorem, $K - K$ is uniformly dense in $C(\mathcal{C})$, the set of continuous functions on $\mathcal{C}$. Since $\mathcal{C}$ is a metric space, for any closed set $A \subset \mathcal{C}$, there exists ([1], Corollary 3.14) $\{\lambda_n\}_{n\in\mathbb{N}} \subset C(\mathcal{C})$ such that $\lambda_n \searrow \chi_A$, where χ_A denotes the indicator function of A. Since $K - K$ is uniformly dense in $C(\mathcal{C})$, for each $n \in \mathbb{N}$, there exists $\{h_{n_k}\}_{k\in\mathbb{N}} \subset K - K$ such that $h_{n_k} \to \lambda_n$ uniformly as $k \to \infty$. Now, let $k_0 \in \mathbb{N}$ be such that

$$\lambda_0(P) - 1 < h_{0_{k0}}(P) < \lambda_0(P) + 1 \qquad , \qquad \forall P \in \mathcal{C}$$

Then, the function

$$g_0 = h_{0_{k0}} + 2$$

is in Θ because Θ is a linear space, and satisfies

$$\lambda_0(P) + 1 < g_0(P) < \lambda_0(P) + 3 \qquad , \qquad \forall P \in \mathcal{C}$$

Next, let $k_1 \in \mathbb{N}$ be such that

$$\lambda_1(P) - \frac{1}{3} < h_{1_{k1}}(P) < \lambda_1(P) + \frac{1}{3} \qquad , \qquad \forall P \in \mathcal{C}$$

Then, $g_1 = h_{1_{k1}} + \frac{2}{3} \in \Theta$ and satisfies

$$\lambda_1(P) + \frac{1}{3} < g_1(P) < \lambda_1(P) + 1 \qquad \forall P \in \mathcal{C}$$

Moreover, for every $P \in \mathcal{C}$, we have

$$g_1(P) < \lambda_1(P) + 1 \leq \lambda_0(P) + 1 < g_0(P)$$

Inductively, define

$$g_n = h_{n_{kn}} + \frac{2}{3^n}$$

Then, $\{g_n\}_{n\in\mathbb{N}} \subset \Theta$, $g_{n+1}(P) < g_n(P)$ $\forall P \in \mathcal{C}$, and

$$\sup_{P\in\mathcal{C}} |g_n(P) - \lambda_n(P)| < \frac{1}{3^{n-1}}$$

Now, the inequality

$$|g_n(P) - \chi_A(P)| \leq |g_n(P) - \lambda_n(P)| + |\lambda_n(P) - \chi_A(P)|$$

shows that $g_n \searrow \chi_A$. [Notice that $g_n(P) > \lambda_n(P) + \frac{1}{3^n} \geq \chi_A(P)$]

By the above claim, we then have $\chi_A \in \Theta$ for any closed set $A \subset \mathcal{C}$. Next, observe that:

(i) $\chi_{\mathcal{C}} \in \Theta$ because the function $1 \in A(\mathcal{C}) \subset \Theta$;

(ii) if $\chi_A, \chi_B \in \Theta$ and $A \subset B$, then $\chi_{B\setminus A} = \chi_B - \chi_A \in \Theta$ because Θ is a linear space;

(iii) if $A_n \nearrow A$ and $\{\chi_{A_n}\} \subset \Theta$, then $\chi_{A_n} \nearrow \chi_A$ and $\chi_A \in \Theta$ by the claim above.

Hence, we conclude that $\mathcal{D} = \{A \in \mathcal{C} \mid \chi_A \in \Theta\}$ is a Dynkin system, which contains all closed sets. Hence, $\mathcal{D} = \mathcal{B}$ (the Borel σ-algebra generated by the topology on $\mathcal{C}$). But now, it follows that Θ contains all the simple functions (because Θ is a linear space) and since $\{\psi_n\} \subset \Theta$ and $\psi_n \nearrow \psi$ imply $\psi \in \Theta$, we conclude that $\Theta = B(\mathcal{B})$, a contradiction. ■

The proof of Theorem 8 shows that the failure SEU theory manifests itself into possible ways: either for some $f \in B(\Sigma)$ the function $\tilde{T}(\kappa(f))(\iota) = \int\limits_{\mathcal{C}} \kappa(f) d\mu_\iota$ is not measurable, which indicates that the decision maker has not enough information to evaluate $f \in B(\Sigma)$; or $\int\limits_{\mathcal{C}} \kappa(f) d\mu \neq \int\limits_{\mathcal{C}/\mathcal{I}} \int\limits_{\iota} \kappa(f) \mid_\iota d\mu_\iota d\mu'$, which indicates that the decision maker cannot come up with a consistent evaluation of that f. By Theorem 8, this is going to happen whenever the prior μ has a nonatomic part and the partition $\mathcal{I}$ is nonmeasurable. We will comment on these features extensively in Section 7.

6 Examples

This section contains some examples of information structures for which SEU fails. As the interpretation of these examples and of their mathematical properties requires a thorough discussion, some explanations will be postponed until the next section. Those readers mainly interested in the emergence of non-EU theories may jump to Section 8, and possibly come back to these sections later.

6.1 Uncertainty on the class of measure zero events

Let (S, Σ) be a Standard Borel Space (see Appendix B) and let $ca_1^+(\Sigma)$ denote the set of countably additive probability measures on Σ. The first information structure that we consider has an obvious relevance to any theory of decision making. It consists of partitioning the set of measures on Σ so that two measures are in the same cell of the partition if and only if they are associated to the same collection of measure zero events in S. For $P, Q \in ca_1^+(\Sigma)$ this partition is defined by the equivalence relation

$$P \mathcal{Z} Q \qquad iff \qquad P \ll Q \qquad and \qquad Q \ll P$$

where $\ll$ stands for absolute continuity, and two measures are equivalent if and only if they are mutually absolutely continuous. Notice that all cells in this partition are measurable (wrt the Borel σ-algebra generated by the weak*-topology on $ca_1^+(\Sigma)$). Informally, the information described by the partition $\mathcal{Z}$ corresponds to statements like "The class of measure zero events in Σ is either Φ or Ψ", etc..

Theorem 9 (see Kechris and Sofronidis [17]) *The partition $\mathcal{Z}$ is nonmeasurable.*

As an immediate consequence, we have

Corollary 10 *Let the decision maker's information be given by the quadruple $\{(\mathcal{C}, \mathcal{B}), \mu, \mathcal{Z}, \mathcal{B}_{\mathcal{Z}}\}$, where $\mathcal{Z}$ is the partition produced by the measure equivalence relation. Assume that μ contains a non-atomic part. Then, SEU obtains if and only if the decision maker is a priori certain about the class of measure zero events of S.*

In other words, if the only information available to the decision maker regards the class of measure zero events, and if the decision maker is uncertain about this class (his prior on $\mathcal{C}$ is not concentrated on a single equivalence class), then the decision maker cannot be Bayesian.

6.2 Ellsberg's three-color urn experiment

In this subsection, we consider Ellsberg's three-color urn experiment. Ellsberg's two-urn experiment is suitable of similar considerations. In the three-color urn experiment, a decision maker faces bets whose domain is an urn containing 90 balls. He is told that 30 of those are red (R) while the remaining are either black (B) or yellow (Y) in unknown proportions. The following violation of the SEU paradigm is often observed

$$R \succ B$$

but

$$R \cup Y \prec B \cup Y$$

That is, the decision maker prefers betting on red rather than black, but he prefers betting on "black or yellow" rather than "red or yellow". Two aspects of the experiment are worth stressing: First, the decision maker explicitly receives information about the set of possible configurations of the urn; second, the information he receives is symmetric with respect to the labels B and Y. What is mostly interesting about the experiment is that, in correspondence of the symmetry in the information, one typically observes a strong symmetry in the decision maker's table of preferences: one can replace B with Y (and vice versa) in the table of preferences without changing the table itself. We believe that this could hardly be considered an accident. In order to follow up on this idea, we must find a way of properly modeling the notion of symmetry encoded in the information as well as that of symmetry in the corresponding behavior.

6.2.1 Modeling the symmetry in the information

The set of possible configurations of the urn corresponds to the set $\mathcal{C}$ in Theorems 1 and 8. In this section, we are going to look for an alternative, yet equivalent, representation of that set, one that would allow us to clearly express the symmetry encoded in the information given to the decision maker. A configuration of the urn can be thought of as a measure on a set S of 90 points, that is a vector with 90 coordinates. S is partitioned into three subsets called R, B and Y. Let us fix an arbitrary configuration p_0. An arbitrary configuration represented by some vector p_i can be expressed in terms of p_0 as there exists a matrix A_i such that $p_i = A_i p_0$. Thus, the configurations of the urn can be identified to a set of (stochastic) matrices, with p_0 being associated to the identity matrix. We are interested in the relation existing among matrices (i.e., configurations) that can be obtained from one another by means of relabeling of the underlying set S. Here, the idea is that since there is nothing substantial about the labels (the information is exactly the same if we replace B with Y and vice versa), it is impossible to distinguish among these matrices.

A relabeling of S is a bijection $t : S \longrightarrow S$. This is evidently associated with the matrix P_t which changes the probabilities according to the relabeling t. So, when we apply the relabeling t, we transform the configuration p_0 into the configuration $P_t p_0$; and, by applying the matrix A to this, we obtain the configuration $AP_t p_0$. Consider now another configuration, say Bp_0, and let us apply the same relabeling t, thus obtaining the configuration $P_t B p_0$. If it turns out that $P_t B = AP_t$, we conclude that the two configurations represented by A and B are effectively indistinguishable because one can be obtained from the other by means of a relabeling of the underlying space. Summing up,

Definition 11 *Two configurations of the urn are the same up to a relabeling of the underlying set S if the corresponding matrices, A and B, are permutation-similar, that is if there is a permutation matrix P_t such that $B = P_t^{-1} A P_t$.*

Remark 12 *Possibly, the finer relation of unitary equivalence (obtained by requiring that the matrix P_t in definition be a unitary matrix) is more appropriate since it preserves also the structure of the underlying space of bets, which is what the decision maker ultimately cares about. This issue, however, is inconsequential to the remainder of the argument and, therefore, we leave it as is.*

In the next subsection, we are going to study a continuous version of Ellsberg's experiment. We will go back to the finite version in Subsection 6.4.

6.3 A continuous version

Here, the urn is the interval $[0, 1]$, which we should think of as partitioned into three subsets, labeled R, B and Y. The set of bets is the set of all indicator functions χ_E, where $E \in \Lambda$ and Λ is the usual Borel σ-algebra. The set of possible configurations of the urn is the set of non-atomic measures on $([0, 1], \Lambda)$, which we denote by $\mathcal{N}([0, 1])$. Thus, a configuration $P \in \mathcal{N}([0, 1])$ corresponds to the measure space $([0, 1], \Lambda, P)$, which under our assumptions is a Lebesgue space (Appendix B). By fixing a possible configuration as a reference point, say $([0, 1], \Lambda, P_0)$, the Isomorphism Theorem for Lebesgue Spaces (see Appendix B) allows us to identify each configuration $([0, 1], \Lambda, P_i)$ with an invertible measure preserving transformation $g_i : ([0, 1], \Lambda, P_0) \longrightarrow ([0, 1], \Lambda, P_i)$. Thus, the set of all possible configurations of the urn can be identified to the group $G = Aut(P_0)$ of invertible measure preserving transformations of $([0,1],\Lambda,P_0)$. The notion of symmetry of two configurations of the urn is expressed by the following definition.

Definition 13 *Two configurations, g_1 and g_2 in G, are the same up to a relabeling of the underlying space, and we write $g_1 \sim g_2$, if there exists a $t \in G$ such that $g_1 = tg_2t^{-1}$.*

Thus the decision maker's information consists of the partition generated by this equivalence relation along with a nonatomic prior on G, which specifies that only those measure spaces $([0,1],\Lambda,P_i)$ such that $P_i(R) = 1/3$ should be considered. Notice that all the cells of the partition are measurable sets (for the measurable structure on G induced by the mapping $P_i \longmapsto g_i$).

Theorem 14 (see Hjorth [15, Theorem 1.2]) *The partition associated to the equivalence relation in Definition 13 is nonmeasurable.*

From this, just like in the previous subsection, it follows that

Corollary 15 *If the decision maker's prior over G contains a non-atomic part and if the prior is not concentrated on a single equivalence class, then SEU fails.*

6.4 The finite version

As we saw above, the finite case is similar to the continuous one. Some extra consideration is needed, nonetheless. In the finite case, a measure is a vector in $\mathbb{R}^n$ and the usual measurable structure (Borel) on the set of measures is the one generated by the Euclidean topology on $\mathbb{R}^n$. This has the inconvenient feature of producing "oversized" information. There is no reasonable presumption, however, that the Borel structure on $\mathbb{R}^n$ be representative of the decision maker's information in Ellsberg's experiment. In fact, the opposite is true as we shall argue below. For the time being, Corollary 16 will give us some sufficient conditions for the failure of SEU theory.

Corollary 16 *Let $\mathcal{C}$ be endowed with a σ-algebra such that (at least one of) the cells of the partition produced by the relation of permutation similarity of stochastic matrices are not measurable sets. Then, the partition is nonmeasurable. If the decision maker's prior contains a non-atomic part and if the prior is not concentrated on a single equivalence class, then SEU fails.*

A basic result in Linear Algebra tells us that in order to show that two matrices are not similar, we need enough information to be able to show that they have different Frobenius normal forms. Consequently, we need even more information to distinguish between permutation non-similar configurations in Ellsberg's experiment. To see what this entails, let us consider a simpler problem, that of distinguishing between different points in $\mathbb{R}^n$. If we had this ability, we could, for instance, determine that certain matrices are not similar because they have different sets of eigenvalues. It is clear that it is always possible to distinguish between different points in $\mathbb{R}^n$ if we know the Borel σ-algebra on $\mathbb{R}^n$ as, for any given two points in $\mathbb{R}^n$, this contains a set which contains one point but not the other. The Borel σ-algebra is generated by the cylinder sets

$$\{x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n \mid x_i = y_i, 1 \leq i \leq \kappa\}$$

That is, a cylinder set is obtained by fixing the first κ coordinates, and the class of all cylinder sets obtains as κ varies between 1 and n. To endow the decision maker in Ellsberg's experiment with this σ-algebra would imply, for instance, the following. Let P_1 and P_2 be two possible configurations of the urn, that is two measures on S, both supported by 89 points but with *supp* $(P_1) \neq$*supp* (P_2). Then, the decision maker would always be able to distinguish between P_1 and P_2, as he would be able (by using a a property of a set which contains P_1 but not P_2) to construct a statistical experiment whose outcome would tell him which was the true measure. Yet, in the actual experiment there is nothing suggesting that he would be able to do so. That is, his information must be coarser than that represented by the class generated by all cylinder sets. Consider now a subclass of that class, which is obtained by allowing κ to vary only between 1 and m, where $m < n$. Let $\mathcal{B}_m$ denote the σ-algebra on $\mathcal{C}$ which is generated by this subclass. It is clear that this σ-algebra does not separate points (that is, there exist two points in $\mathbb{R}^n$ such that no set in the σ-algebra contains one but not the other). Thus, if all we know about $\mathbb{R}^n$ is this σ-algebra, we cannot distinguish between points and, hence, between non-similar matrices. In fact, for all choices of $m < n$, these σ-algebras satisfy the condition in Corollary 16. In the next section, we will discuss extensively the meaning of the σ-algebra, its interpretation and the role it plays in the failure of SEU theory. We will revisit all three examples of this section in Section 9 and in Section 10.3.

7 Comments: the idea of insufficient information

"A person behind a door slips a blank paper through the door. Is that person a male or a female?". "The eye witness of a robbery describes the perpetrator as a male, wearing a mask and gloves, medium built, medium height. Is he 5'6" tall or 5'9"? Does he weight 140 lbs or 160 lbs? Does he have a beard? Is he white, black, Asian?". These are examples of insufficient information: on the basis of the information that you have, you can make certain distinctions but you still lack information on so many fundamental aspects of the problem that you can only come up with very coarse, almost useless, distinctions. In the second example, for instance, you know that the perpetrator is not a female, that is neither very tall nor very short, neither very heavy nor very light, but clearly there is a lot more you need to know to form an even remote idea of where to begin the search for the perpetrator. In the first example, you do know that the person is either a male or a female, but you will never be able to either check that or to form meaningful probabilistic assessments about that unless some other information is revealed (for instance, you write questions on that piece of paper and slip it back through the door, the person behind the door starts talking etc.). As another example, suppose that I am subject to a technological constraints that allows me to check fractional numbers only up to the tenth digit. Once I know that the number is 3.1415926535, do I conclude that the number is rational or that it is irrational? This is what insufficient information means in practice, and this is precisely what the mathematical concept of nonmeasurable partition conveys: *if my information is described by a nonmeasurable partition I do not know enough to tell things apart*.

In order to get a thorough understanding of why the concept of nonmeasurable partition is the mathematical translation of the intuitive idea of insufficient information, we must somehow return to the basics and keep in mind that the words *subset* and *property* are interchangeable (as, by the definition, a subset is the collection of all points having a certain property). For $\mathcal{C}$ the set of measures, let $\mathcal{I}$ be a partition of $\mathcal{C}$ and let $\pi : \mathcal{C} \longrightarrow \mathcal{C}/\mathcal{I}$ be the canonical projection. Let us begin with an extreme case, which will make certain features more transparent. Let us suppose that the decision maker understands nothing about the set of measures. This situation of absolute ignorance is represented by the decision maker having the trivial σ-algebra $\{\varnothing, \mathcal{C}\}$ on $\mathcal{C}$ (that is, the decision maker knows no properties). In such a case, the σ-algebra on the quotient that makes the canonical projection measurable is again the trivial σ-algebra $\{\varnothing, \mathcal{C}/\mathcal{I}\}$ and the only measurable functions on the quotient are the constant functions. It follows that any non-trivial partition on $\mathcal{C}$ will generate, given any system of conditional probabilities, many nonmeasurable functions on the quotient (Condition 1. in Sec. 4 is violated) and SEU theory must fail. In a manner of speaking, given his (lack of) understanding of $\mathcal{C}$, the decision maker cannot handle partitions other than the trivial one. It is clear that the finer the algebra on $\mathcal{C}$, the more partitions the decision maker will be able to handle. This is only one part of the story, though. As we have seen in the first two examples of the previous section, even when the algebra on $\mathcal{C}$ is fine enough to guarantee that the partition is made only of measurable sets, it is still possible that SEU would fail. This has to do with the fact that while the sets making up the partition are measurable (and, hence, "simple" enough when considered in isolation), the partition as a whole is "too complicated" for the decision maker to be able to assign (non-trivial) measures to those sets. In other words, the nonmeasurability of a partition is either a statement that the decision maker does not fully understand the objects making up the partition or that, while he understands those objects one by one, he is unable to come up with coherent assessments on the whole. Intuitively, whether or not a partition is measurable depends on the comparison between "how complicated the partition is" relative to the decision maker's knowledge ("how many properties he knows of"). Thus, nonmeasurability of a partition can be achieved by either complicating the partition or by reducing the amount of properties that the decision maker knows of.

When we want to distinguish a point x from a point y, short of knowing x with absolute certainty (which corresponds to the atomic case, hence the nonatomicity condition encountered in Theorem 8), the minimal condition is that of knowing an open set that contains x and does not contain y. This is the same as saying that x is different from y because x has a property that y does not have (the one associated with that open set). When x and y are points in $\mathbb{R}^n$, we can distinguish between them when we know the Euclidean topology of $\mathbb{R}^n$, but if we are limited only to the projections on the first $n-1$ coordinates, this is no longer possible. This is what is going on in the third example of the previous section as well as in the rational vs irrational number question above. The explanation is a bit different in the first two examples of the previous section, but the spirit is similar. As said above, those examples express situations where the partition is, intuitively speaking, too complicated. Formally, what is going on is as follows: for any point $x \in \mathcal{C}$, *any* open set around x intersects all equivalence classes. In other words, there is no property that the decision maker knows of that would allow him to distinguish among equivalence classes. Thus, either he is able to distinguish among points at the outset (which is the purely atomic case) or he does not need to distinguish between equivalence classes (because his prior is concentrated on a single equivalence class) or he would not be able to assign weights to the equivalence classes (because as soon as he assigns a weight to a certain property he would assign that weight to all the equivalence classes; ultimately this would result in inconsistent evaluations).[4]

We conclude this discussion by stressing, once more, that the failure of SEU is determined by the interplay between the σ-algebra (= the decision maker's knowledge) and the partition (= the decision maker's information, given his knowledge). This is a necessary feature which accounts for two important aspects of the intuitive idea of ambiguity; namely, (a) the ambiguity that is perceived in a decision problem may vary across different individuals; and (b) for the same individual, the perceived ambiguity may vary across decision problems. This is intuitively clear. A question of the type "Does there exist an extension of this functional satisfying such and such property?" may bear no ambiguity to a trained mathematician while appearing utterly obscure to the untrained person. At the same time, the very same mathematician might find himself/herself at a loss when facing the statement "for this type of shot, this type of club is better than that other type".

8 Ambiguity: a formal definition

We can now summarize our findings by means of the following definition. Let $\mathcal{C}$ be a set of measures on (S, Σ), $\mathcal{T}$ a σ-algebra of subsets of $\mathcal{C}$, $\mathcal{I}$ a partition of $\mathcal{C}$ and let $\mathcal{T}_{\mathcal{I}}$ be the σ-algebra on the quotient induced by the canonical projection.

Definition 17 *A decision maker faces Ambiguity (or Knightian Uncertainty) whenever his information about the set $\mathcal{C}$ is described by a quadruple $\{(\mathcal{C}, \mathcal{T}), \mu, \mathcal{I}, \mathcal{T}_{\mathcal{I}}\}$ (see Definition 3) with the following properties:*

(i) μ contains a non-atomic part;

[4] Hopefully, the discussion in this and the previous section has not generated the erroneous idea that nonmeasurable partitions made exclusively of measurable sets do not exist in finite dimension. In fact, these partitions exist even in the one-dimensional case. An example is given by the unit interval equipped with its usual Lebesgue structure and by the partition of the unit interval produced by the equivalence relation $x \sim y$ if and only if $y = x + \alpha \pmod 1$, where α is a fixed irrational number (see [6]).

(ii) μ is not concentrated on a single equivalence class;

(iii) the partition $\mathcal{I}$ (modulo μ-measure zero events) is non-measurable.

In correspondence to all other information structures, the decision maker faces (Knightian) Risk.

The necessity and the meaning of all three conditions was explained in the previous section. Theorem 8 can now be reformulated as follows: *If the decision maker faces Ambiguity, then he cannot obey the SEU paradigm.*[5]

By means of Definition 17, we can partition the set of all (Bayesian) information structures into two subsets: one represents those information structures which describe good information, and corresponds to situations of Risk; the other represents those information structures which describe information which is not good, and corresponds to situations of Ambiguity. A Bayesian decision maker obeys SEU theory when he faces Risk, and departs from it when he faces Ambiguity.

With this, we have completed the first part of our program: we have shown that Ambiguity exists, we have formally identified those situations that should be deemed as ambiguous, and we have shown that departures from SEU must be observed in these situations. In the next section, we move to the second part of our program, that of showing which non-EU theories might emerge in situations of Ambiguity.

9 Subjectively measurable acts

In this section, we are going to isolate a subset of the acts: those that the decision maker is able to evaluate on the basis of his information and for which an expected utility functional (determined by the decision maker's prior μ) can be meaningfully defined. We call them *subjectively measurable acts*. In the next section, we will study the problem of extending the expected utility functional on the subjectively measurable acts to the set of all acts in a way that respects the decision maker's information. It will be at that point that non-EU behavior will emerge.

[5] As anticipated at the end of Section 3, the considerations of finitely additive priors only strenghtens our findings. In fact, if a prior μ over $\mathcal{C}$ is finitely additive but not countably additive, then by definition there exists a partition $\mathcal{I}$ and a function $\psi \in B(\mathcal{B})$ such that $\int_{\mathcal{C}} \psi d\mu \neq \int_{\mathcal{C}/\mathcal{I}} \int_{\iota} \psi \mid_{\iota} d\mu_{\iota} d\mu'$. It suffices to consider a partition $\mathcal{I} = \{A, \{B_i\}_{i\in\mathbb{N}}\}$ where $\{B_i\}_{i\in\mathbb{N}}$ is a family of disjoint sets for which μ fails countable additivity. Thus, existence of non-measurable partitions is a rather easy matter when we allow for measures that are only finitely additive.

Let $\{(\mathcal{C}, \mathcal{B}), \mu, \mathcal{I}, \mathcal{B}_{\mathcal{I}}\}$ be an ambiguous information structure (that is, the properties in Definition 17 are satisfied) and let $cp = \{\mu_\iota\}_{\iota \in \mathcal{I}}$ be a collection of probability measures with μ_ι supported by ι (viewed as a subset of $\mathcal{C}$). As we have seen in the proof of Theorem 8, the family $cp = \{\mu_\iota\}_{\iota \in \mathcal{I}}$ cannot be a system of canonical condition probabilities and, therefore, cannot be used to evaluate all acts. Yet, the family cp may display a partial compatibility (possibly trivial) with the decision maker's information, and can be used to form at least some conditional evaluations. Precisely, given an information structure and a family of probabilities $cp = \{\mu_\iota\}_{\iota \in \mathcal{I}}$, a subset of the acts may satisfy the two conditions of Section 4. This is the set of *subjectively measurable acts determined by* cp, which we denote by $\Sigma MA(cp)$. Recall that, as in the proof of Theorem 8, an act $f \in B(\Sigma)$ and a system of probabilities $cp = \{\mu_\iota\}_{\iota \in \mathcal{I}}$ induce a function $\psi_f : \mathcal{C}/\mathcal{I} \longrightarrow \mathbb{R}$ which is defined by $\psi_f(\iota) = \int_{\mathcal{C}} \kappa(f) d\mu_\iota$, where as usual κ is the canonical mapping $B(\Sigma) \longrightarrow A(\mathcal{C})$. Thus (see Sec. 5), the *set of subjectively measurable acts* is given by

$$\Sigma MA(cp) = \left\{ f \in B(\Sigma) \mid (a)\ \psi_f \text{ is measurable; } (b) \int_{\mathcal{C}} \kappa(f) d\mu = \int_{\mathcal{C}/\mathcal{I}} \int_{\iota} \kappa(f) \mid_\iota d\mu_\iota d\mu' \right\} \quad (1)$$

When the information structure is ambiguous, the complement of this set in $B(\Sigma)$ is non empty by Theorem 8, and represents those acts that the decision maker cannot evaluate on the basis of his information.

A few comments are in order. Firstly, the set of subjectively measurable acts depends, in principle, on the system of probabilities cp. This is so because the choice of cp may affect the determination of which functions of the type ψ_f are indeed measurable. Because of this feature, it makes sense to consider, alongside with $\Sigma MA(cp)$, another class of acts, which we call the class of *unambiguously measurable acts* and denote by UMA. This is defined as

$$UMA = \left\{ f \in B(\Sigma) \mid \int_{\mathcal{C}} f dP = \int f dP' \text{ for all } P, P' \in \mathcal{C} \right\}$$

In Proposition 18 below, we shall see that these acts are precisely those that are subjectively measurable irrespective of the choice of the system of conditional probabilities cp. Moreover, in some important cases, for instance in the examples of Sec. 6, the properties of the nonmeasurable partition completely pin down the class of subjectively measurable acts and render it equal to UMA. Part (6) of Proposition 18 below gives a condition for this to be the case. Secondly, it is important to notice that if $f \in \Sigma MA(cp)$, its evaluation does not depend on the choice of the system cp but on the decision maker's prior μ only by virtue of condition (b) in the definition (the same goes for the natural measure in the Proposition below). Finally, we should like to observe that it is possible to allow for greater generality than we have done so far. In fact, one could replace the family $cp = \{\mu_\iota\}_{\iota \in \mathcal{I}}$ with a family of functionals $\{V_\iota\}$ which are to be interpreted as conditional evaluations functionals. In such a case, the function ψ_f above would be replaced by the function $\tilde{\psi}_f$ defined by $\tilde{\psi}_f(\iota) = V_\iota(\kappa(f))$, condition (a) by the condition that this function be measurable and condition (b) by the condition $\int\limits_{\mathcal{C}} \kappa(f) d\mu = \int\limits_{\mathcal{C}/\mathcal{I}} \tilde{\psi}_f d\mu'$. The study of this case can be conducted along the same lines. In particular, parts (3) to (6) of Proposition 18 below hold unchanged.

The subset of $\Sigma MA(cp)$ defined by

$$\Sigma ME(cp) = \{\chi_E \in B(\Sigma) \mid \chi_E \in \Sigma MA\}$$

is of special importance as it describes all the events in Σ to which the decision maker can assign probabilities (Proposition 18, part (3) below). We call its elements the *subjectively measurable events determined by* cp. Just like we did above, in parallel to $\Sigma ME(cp)$, we also introduce the class of *unambiguously measurable events*, which is defined by

$$UME = \{\chi_E \in B(\Sigma) \mid \chi_E \in UMA\}$$

The basic properties of the classes $\Sigma MA(cp)$ and $\Sigma ME(cp)$ are stated in the next proposition.

Proposition 18 *The following holds:*

(1) $\Sigma MA(cp)$ is a linear space. Moreover, condition (b) in (1) defines the expected utility functional $\mathcal{E} : \Sigma MA(cp) \longrightarrow \mathbb{R}$ by

$$\mathcal{E}(f) = \int\limits_{\mathcal{C}} \kappa(f) d\mu$$

(2) The class $\Sigma ME(cp)$ is non-empty and is a finite λ-system (i.e., is closed under complementation and finite disjoint unions);

(3) There exists a natural measure N on $\Sigma ME(cp)$, defined by

$$N(E) = \int_{\mathcal{C}} \kappa(\chi_E) d\mu \qquad , \qquad E \in \Sigma ME(cp)$$

where μ is the decision maker's prior on $\mathcal{C}$. In fact, N is the restriction of $\mathcal{E}$ to $\Sigma ME(cp)$ (via the identification $E \mapsto \chi_E$)

(4) $UME \subset \Sigma ME(cp)$ for every system cp; that is, $UE \subset \bigcap_{cp} \Sigma ME(cp)$

(5) $UMA \subset \bigcap_{cp} \Sigma MA(cp)$

(6) If there exists an ergodic nonatomic measure on the quotient $\mathcal{C}/\mathcal{I}$, then for every system of probabilities cp we have

$$UMA = \Sigma MA(cp) \qquad \textit{and} \qquad UME = \Sigma ME(cp)$$

Notice that the Proposition still holds in the Risk case, that is when the partition $\mathcal{I}$ is measurable. In such a case, if cp is canonical then SEU obtains, every event in Σ belongs to $\Sigma ME(cp)$, every function in $B(\Sigma)$ is in $\Sigma MA(cp)$, and the natural set function on Σ is the "average" measure obtained through the integration over priors theorem. In the Ambiguity case, the inclusions $\Sigma MA(cp) \subset B(\Sigma)$ and $\Sigma ME(cp) \subset \Sigma$ are always strict for any system of probabilities cp by Theorem 8. The final part of Proposition 18 is interesting because, as anticipated, it gives us a condition guaranteeing that the set of subjectively measurable acts and events are independent of the choice of the conditional probabilities. A nonatomic measure on the quotient $\mathcal{C}/\mathcal{I}$ is ergodic if the measure of any saturated set (= union of equivalence classes) is either 0 or 1. Existence of such a measure is a property of the partition $\mathcal{I}$ and, as anticipated, it is satisfied in all the examples of Section 6.

Proof. (1) Let $cp = \{\mu_\iota\}_{\iota \in \mathcal{I}}$ be a collection of probability measures with μ_ι supported by ι (viewed as a subset of $\mathcal{C}$). For $\psi \in B(\mathcal{B})$, let $\tilde{\psi} : \mathcal{C}/\mathcal{I} \longrightarrow \mathbb{R}$ be defined by $\tilde{\psi}(\iota) = \int_{\mathcal{C}} \psi d\mu_\iota$. As noted in the proof of Theorem 8, the set

$$\Theta = \left\{ \psi \in B(\mathcal{B}) \mid (a)\ \tilde{\psi} \text{ is measurable; } (b) \int_{\mathcal{C}} \psi d\mu = \int_{\mathcal{C}/\mathcal{I}} \int_{\iota} \psi \mid_\iota d\mu_\iota d\mu' \right\}$$

is a linear subspace of $B(\mathcal{B})$. Since a function $f \in B(\Sigma)$ is subjectively measurable if and only if $\kappa(f) \in \Theta$, it follows that the class $\Sigma MA(cp)$ is the set $\kappa^{-1}(\kappa(B(\Sigma)) \cap \Theta)$. From the linearity of κ, it immediately follows that this is a linear subspace of $B(\Sigma)$. The second part of the statement follows at once from the linearity and the positivity of the integral as well as the linearity of κ.

(2) From (1), it follows immediately that $\Sigma ME(cp)$ is closed under finite disjoint unions. Moreover, since constant functions always belong to $(\kappa(B(\Sigma)) \cap \Theta)$, the class $\Sigma ME(cp)$ always contains $\varnothing$ and S.

(3) $\chi_E \in \Sigma ME(cp)$ implies $\int\limits_{\mathcal{C}} \kappa(\chi_E) d\mu = \int\limits_{\mathcal{C}/\mathcal{I}} \int\limits_{\iota} \kappa(\chi_E) \mid_\iota d\mu_\iota d\mu'$. Hence, $E \longmapsto \int\limits_{\mathcal{C}} \kappa(\chi_E) d\mu$ is an additive set function on $\Sigma ME(cp)$.

(4)-(5)-(6) An act (event) $f \in B(\Sigma)$ $(E \in \Sigma)$ is unambiguously measurable iff $\displaystyle\int_s f dP = \int_s f dP'$ $(P(E) = P'(E))$ for all P and P' in $\mathcal{C}$, which means that the mapping $\kappa(f)$ $(\kappa(\chi_E))$ is a constant mapping on $\mathcal{C}$. By (2), constant functions are always subjectively measurable, which proves the inclusions in (4) and (5). Finally, if the condition in part (6) is satisfied, the only subjectively measurable functions (modulo sets of measure 0) are constant and equality holds. ■

We conclude this section by determining the sets $\Sigma MA(cp)$ and $\Sigma ME(cp)$ in the aforementioned examples. We also notice, by means of an example, that the determination of the set $\Sigma MA(cp)$ permits us to derive a new object which, intuitively speaking, represents the understanding (or the subjective view) that the decision maker has of the (objective) set of states.

Example 19 (Ex. Sec. 6.1 cntd) *By [17], condition (6) in Proposition 18 is satisfied. Hence, an act is subjectively measurable iff it is unambiguously measurable. It is easy to check that for* $\mathcal{C} = ca_1^+(\Sigma)$, ΣMA *contains constant functions only and, consequently,* $\Sigma ME = \{\varnothing, S\}$.

Example 20 (Ex. Sec. 6.3 ctnd) *By [15, Theorem 1.2], condition (6) in Proposition 18 is satisfied and an act is subjectively measurable iff it is unambiguously measurable. In this case, the set* $\Sigma ME = UME$ *is given by*

$$\Sigma ME = UME = \{E \in \Sigma \mid \textit{either } E \subseteq R \textit{ or } E \subseteq R^c\}$$

Hence,

$$\Sigma MA = UMA = \overline{\overline{lin}} \{\chi_E \mid E \in \Sigma ME\}$$

where $\overline{\overline{\cdot}}$ *denotes the closed linear span of the set.*

Example 21 (Ex. Sec. 6.4 ctnd) *If we assume that all the cells of the partition are nonmeasurable sets, as it would be the case if we use the* σ*-algebra at the end of Sec. 6.4, then the sets of subjectively measurable acts and events are exactly as in the previous item.*

Example 22 *Consider a variation of the previous example where the state space* S *consists of only* 3 *points, labeled* R, B *and* Y. *The analysis of the example is essentially the same as in Sec. 6.4. Then, by the previous item, the set* ΣMA *is spanned by the indicator function of the point* R *and by the indicator function of the complement of* R *in* S. *Thus,* ΣMA *is the Euclidean space* $\mathbb{R}^2$ *while the set of all acts is* $\mathbb{R}^3$. *Effectively, on the basis of his information, the decision maker can only handle a two-point set of states while the actual one is a three-point set of states.*

9.1 Unambiguous events in the sense of [10] and [21]

There is an apparent similarity between our unambiguously measurable events and the unambiguous events of Ghirardato, Maccheroni, Marinacci [10] and of Nehring [21]: in both cases, the unambiguous events are those that are assigned the same measures by all the probabilities in a certain set. The same observation applies to the comparison between our unambiguously measurable acts and the *crisp* acts of [10]. In principle, however, these classes are different because the sets of probability measures intervening in their definition need not be the same. Obviously, a relation is to be expected; we will look into it in Section 11.

10 Foundations of non-EU theories

In our inquiry into the role played by Information in problems of decision making under uncertainty, we have been led to consider Bayesian decision makers who want to integrate over a set of priors. When the information is ambiguous, this operation is not possible and, consequently, there are acts that cannot be evaluated. Only a subset of the acts can be evaluated by using an expected utility functional. Thus, the problem of determining the behavior of these decision makers becomes that of extending this functional to the set of all acts *in a way that respects the decision maker's* *information*. We study this problem in the present section. Its solution will give us the class of non-EU theories that emerge as a response to Ambiguity.

We should like to stress, once again, that this extension must respect the decision maker's information. This means that no additional information, aside from that contained in the information structure originally available, can be used in the extension from the set of subjectively measurable acts to the set of all acts. In other words, an extension obtained by requiring that a certain property, say Π, be satisfied is legit (and meaningful) only if the decision maker can formulate Π irrespective of the information available to him. We will focus on extensions satisfying *monotonicity*: that is, we require that for any two acts $f, g \in B(\Sigma)$ if $f \geq g$ then the decision maker must prefer f to g. We believe that the necessity to restrict to monotone extensions is transparent, and requires no further explanation. Later, we will also introduce another property (translation invariance), which is also independent of the information. We believe that, in our setting, the introduction of this property is also mandatory, and we will argue in favor of our point of view. But, since some disagreement is possible, we take care of formulating our results also in absence of this property.

The present section is divided into three subsections. The first contains some basic lemmata that lead to the proof of the main theorem of this section. This is proved in the second subsection. It states that behavior emerging in the face of Ambiguity belongs to the class of Invariant Biseparable preferences (or, in fact, to a wider class, if translation invariance is not imposed). The third subsection introduces some symmetry considerations, which are mainly motivated by the examples of the previous sections. These considerations lead to singling out a special model within the class of IB preferences: the α-MEU model.

10.1 Some lemmata

Let K be a linear subspace of $B(\Sigma)$ which contains the constant functions, and let $\mathcal{E} : K \longrightarrow \mathbb{R}$ be a positive linear functional. In terms of interpretation, K is to be thought of as the subset of subjectively measurable acts and $\mathcal{E}$ as the expected utility functional defined on those. Let us define two new functionals $\underline{\mathrm{I}} : B(\Sigma) \longrightarrow \mathbb{R}$ and $\bar{I} : B(\Sigma) \longrightarrow \mathbb{R}$ as follows

$$\begin{aligned}\underline{\mathrm{I}}(f) &= \sup_{g\in K}\{\mathcal{E}(g) \mid g \le f\} \\ \bar{I}(f) &= \inf_{g\in K}\{\mathcal{E}(g) \mid g \ge f\}\end{aligned}$$

Lemma 23 *$\underline{I}$ and $\bar{I}$ are monotone, translation invariant extensions of $\mathcal{E}$. Moreover, $\underline{I} \le \bar{I}$. In addition, $\underline{I}$ is super-additive and $\bar{I}$ is sub-additive.*

Proof. (a) $\underline{\mathrm{I}}$ and $\bar{I}$ are extensions of I:

$\forall f \in B(\Sigma)$, the sets $\{g \le f \mid g \in K\}$ and $\{g \ge f \mid g \in K\}$ are both nonempty because f is bounded and K contains the constant functions. Thus, both $\underline{\mathrm{I}}$ and $\bar{I}$ are well-defined and they are extensions of $\mathcal{E}$ since $f \in K$ implies $\bar{I}(f) = \mathcal{E}(f) = \underline{\mathrm{I}}(f)$.

(b) $\underline{\mathrm{I}} \le \bar{I}$: For any $\psi, \varphi \in K$ such that $\psi \ge f \ge \varphi$, we have that $\mathcal{E}(\psi) \ge \mathcal{E}(\varphi)$ because $\mathcal{E}$ is positive. Hence,

$$\bar{I}(f) = \inf_{\{\psi \ge f \mid \psi \in K\}} \mathcal{E}(\psi) \ge \sup_{\{\varphi \le f \mid \varphi \in K\}} \mathcal{E}(\psi) = \underline{\mathrm{I}}(f)$$

(c) $\underline{\mathrm{I}}$ and $\bar{I}$ are monotone:

$$\begin{aligned} f \ge h \quad &\Longrightarrow \quad \{\psi \ge f \mid \psi \in K\} \subset \{\psi \ge h \mid \psi \in K\} \\ &\Longrightarrow \quad \bar{I}(f) \ge \bar{I}(h)\end{aligned}$$

similarly,

$$\begin{aligned} f \ge h \quad &\Longrightarrow \quad \{\psi \le f \mid \psi \in K\} \supset \{\psi \le h \mid \psi \in K\} \\ &\Longrightarrow \quad \underline{\mathrm{I}}(f) \ge \underline{\mathrm{I}}(h)\end{aligned}$$

(d) $\underline{\mathrm{I}}$ and $\bar{I}$ are translation invariant:

Let $\alpha \geq 0$ and $\beta \in \mathbb{R}$, and let $\mathbf{1}$ be the function that is identically equal to 1 on S. Then,

$$\begin{aligned}
\underline{\mathrm{I}}(\alpha f + \beta \mathbf{1}) &= \sup_{g \in K} \{\mathcal{E}(g) \mid g \leq \alpha f + \beta \mathbf{1}\} \\
&= \sup_{\psi \in K} \{\mathcal{E}(\alpha\psi + \beta\mathbf{1}) \mid \alpha\psi + \beta\mathbf{1} \leq \alpha f + \beta\mathbf{1}\} && \text{b/c K is a linear space} \\
& && \text{which contains the constants} \\
&= \sup_{\psi \in K} \{\alpha\mathcal{E}(\psi) + \beta \mid \psi \leq f\} && \text{by linearity of } \mathcal{E} \\
&= \alpha \sup_{\psi \in K} \{\mathcal{E}(\psi) \mid \psi \leq f\} + \beta \\
&= \alpha \underline{\mathrm{I}}(f) + \beta
\end{aligned}$$

which shows that $\underline{\mathrm{I}}$ is translation invariant. The proof for $\bar{I}$ is similar.

(e) $\bar{I}$ is sub-additive and $\underline{\mathrm{I}}$ is super-additive:

Let $f, h \in B(\Sigma)$. By definition,

$$\bar{I}(f) = \inf_{g \in K} \{\mathcal{E}(g) \mid g \geq f\}$$

Thus, $\forall \varepsilon > 0\ \exists g_1, g_2 \in K$ such that (1) $g_1 \geq f$, $g_2 \geq h$; and (2) $\bar{I}(f) > \mathcal{E}(g_1) - \varepsilon$ and $\bar{I}(h) > \mathcal{E}(g_2) - \varepsilon$. Since $g_1 + g_2 \geq f + h$, we have (by definition of $\bar{I}$) that

$$\bar{I}(f + h) \leq \mathcal{E}(g_1 + g_2) < \bar{I}(f) + \bar{I}(h) + 2\varepsilon$$

By letting $\varepsilon \to 0$, we conclude that

$$\bar{I}(f + h) \leq \bar{I}(f) + \bar{I}(h)$$

that is the sub-additivity of $\bar{I}$.

Similarly, $\forall \varepsilon > 0\ \exists g_1, g_2 \in K$ such that (1) $g_1 \leq f$, $g_2 \leq h$; and (2) $\underline{\mathrm{I}}(f) < \mathcal{E}(g_1) + \varepsilon$ and $\underline{\mathrm{I}}(h) < \mathcal{E}(g_2) + \varepsilon$. Hence,

$$\underline{\mathrm{I}}(f + h) \geq \mathcal{E}(g_1 + g_2) > \underline{\mathrm{I}}(f) + \underline{\mathrm{I}}(h) - 2\varepsilon$$

from which we conclude that

$$\underline{\mathrm{I}}(f + h) \geq \underline{\mathrm{I}}(f) + \underline{\mathrm{I}}(h)$$

that is the super additivity of $\underline{I}$. $\blacksquare$

In the next Lemma, we are going to see that any monotone extension of $\mathcal{E} : K \longrightarrow \mathbb{R}$ is sandwiched between $\underline{I}$ and $\bar{I}$.

Lemma 24 *If $J : B(\Sigma) \longrightarrow \mathbb{R}$ is a monotone extension of $\mathcal{E} : K \longrightarrow \mathbb{R}$, then $\forall f \in B(\Sigma)$*

$$\underline{I}(f) \leq J(f) \leq \bar{I}(f)$$

Proof. By the way of contradiction, assume that there exists an $f \in B(\Sigma)$ such that $J(f) > \bar{I}(f)$. Clearly, $f \in B(\Sigma)\backslash K$. Since $f \in B(\Sigma)$ and K contains the constant functions, the set

$$MK(f) = \{g \geq f \mid g \in K\} \neq \varnothing$$

and, by $J = \mathcal{E}$ *on* K and by the fact that J is monotone, we have that for every $g \in MK(f)$

$$\mathcal{E}(g) = J(g) \geq J(f) > \bar{I}(f)$$

which implies that $J(f)$ is a lower bound for $\mathcal{E}(g)$ on $MK(f)$, thus contradicting the definition of $\bar{I}(f)$ (as $\bar{I}(f) = \inf\limits_{g \in MK(f)} \mathcal{E}(g)$). We conclude that $\bar{I}(f) \geq J(f)$, $\forall f \in B(\Sigma)$.

In a similar fashion, suppose that $\exists f \in B(\Sigma)$ such that $J(f) <\underline{I}(f)$. Since the constant mappings are in K

$$mK(f) = \{g \leq f \mid g \in K\} \neq \varnothing$$

and, by the monotonicity of J and $J = \mathcal{E}$ on K, we have that $\forall g \in mK(f)$

$$\underline{I}(f) > J(f) \geq J(g) = \mathcal{E}(g)$$

which implies that $J(f)$ is an upper bound for $\mathcal{E}(g)$ on $mK(f)$, thus contradicting the definition of $\underline{I}(f)$ (as $\underline{I}(f) = \sup\limits_{g \in mK(f)} \mathcal{E}(g)$). We conclude that $\underline{I}(f) \leq J(f)$, $\forall f \in B(\Sigma)$. $\blacksquare$

The next Lemma gives us a representation of the functionals $\underline{I}$ and $\bar{I}$. It is a simple consequence of the Hahn-Banach theorem. Its proof is standard and is included here only for completeness.

Lemma 25 (representation of $\underline{I}$ and $\bar{I}$) *Assume that $E(\mathbf{1}) = 1$. There exists a unique convex,* *weak*-compact set* $\mathcal{L}(I)$ *of linear functionals such that*

$$\underline{I}(f) = \min_{L \in \mathcal{L}(I)} L(f) \qquad \textit{and} \qquad \bar{I}(f) = \max_{L \in \mathcal{L}(I)} L(f)$$

Proof. Let B' denote the set of all linear functionals on $B(\Sigma)$. Let

$$\mathcal{L}(I) = \left\{ L \in B' \mid \ L \text{ is an extension to } B(\Sigma) \text{ of } \mathcal{E},\, \bar{I} \geq L \geq \underline{\mathrm{I}} \text{ and } L(\mathbf{1}) = \mathcal{E}(\mathbf{1}) \right\}$$

Notice that all elements of $\mathcal{L}(I)$ are positive due to the monotonicity of $\underline{\mathrm{I}}$ and that they are continuous because they are bounded owing to the condition $L(\mathbf{1}) = \mathcal{E}(\mathbf{1})$. Moreover, owing to Lemma 24, the elements of $\mathcal{L}(I)$ are all the positive extensions of I. By Hahn-Banach, $\mathcal{L}(I)$ is nonempty. For $f \in B(\Sigma)$, let $K_f = lin\{K, f\}$ be the linear subspace spanned by K and f. The functional $\underline{\mathrm{L}}_f : K_f \longrightarrow \mathbb{R}$ defined by

$$\underline{\mathrm{L}}_f(\alpha f + \beta k) = \alpha \underline{\mathrm{I}}(f) + \beta \underline{\mathrm{I}}(k)$$

is linear on K_f, $\underline{\mathrm{L}}_f \geq \underline{\mathrm{I}}$ on K_f (by the super-additivity of $\underline{\mathrm{I}}$) and has the property that $\underline{\mathrm{L}}_f(f) = \underline{\mathrm{I}}(f)$. Again by Hahn-Banach, $\underline{\mathrm{L}}_f$ has an extension which belongs to $\mathcal{L}(I)$, which proves the formula for $\underline{\mathrm{I}}$. Similarly, the functional $\bar{L}_f : K_f \longrightarrow \mathbb{R}$ defined by

$$\bar{L}_f(\alpha f + \beta k) = \alpha \bar{I}(f) + \beta \bar{I}(k)$$

is linear on K_f, $\bar{L}_f \leq \bar{I}$ on K_f (by the sub-additivity of $\bar{I}$) and has the property that $\bar{L}_f(f) = \bar{I}(f)$. Again by Hahn-Banach, $\bar{L}_f$ has an extension which belongs to $\mathcal{L}(I)$, which proves the formula for $\bar{I}$. Uniqueness of $\mathcal{L}(I)$ follows from a standard separation argument as, for instance, in [13]. Convexity and weak*-compactness of $\mathcal{L}(I)$ are obvious. ■

10.2 ... all priors compatible with his limited information

When facing Ambiguity the decision maker is able to determine, on the basis of his information, an expected utility functional on the space of subjectively measurable acts. This determination encompasses all the information available to him. Once this is done, he must face the problem of evaluating those acts that are not subjectively measurable, that is he must somehow extend his evaluation to the set of all acts. As we observed, the most natural way of doing so is to demand that the extension should satisfy the property of monotonicity (which is automatically satisfied on the set of subjectively measurable acts). In this section, we begin by studying the class of all monotone extensions of the decision maker's functional on the subjectively measurable acts. From these, we will then single out those that respect the decision maker's information, that is those that use the knowledge of the EU functional on the space of subjectively measurable acts and that knowledge only. We will, then, use these to characterize the type of behavior that emerges in situations of Ambiguity.

Let $\mathcal{E}$ denote the expected utility functional on the set of subjectively measurable acts (Proposition 18). By Lemma 24, any monotone extension of $\mathcal{E}$ to $B(\Sigma)$ has the form

$$I(f) = \alpha(f)\underline{I}(f) + (1 - \alpha(f))\bar{I}(f)$$

where $\forall f \in B(\Sigma)$, $\alpha(f) \in [0, 1]$. By combining Lemma 25 with the Riesz representation theorem, we can re-write this

$$I(f) = \alpha(f) \min_{P \in \mathcal{C}} \int f dP + (1 - \alpha(f)) \max_{P \in \mathcal{C}} \int f dP \tag{2}$$

where $\mathcal{C}$ is a (uniquely determined) convex, weak*-compact subset of $ba_1^+(\Sigma)$. Thus, (2) gives us the class of all monotone extensions of $\mathcal{E}$. We must now isolate those extensions that do not use more information than the one available to the decision maker, which is the one embedded into the functional $\mathcal{E}$. Clearly, the functionals $\underline{I}$ and $\bar{I}$ are two such extension as they use only the set of Hahn-Banach extensions of $\mathcal{E}$. In general, this is not true for all functionals of the type (2). In fact, the condition translates into a restriction on the coefficient $\alpha(\cdot)$ in (2).

Definition 26 *An extension $I : B(\Sigma) \longrightarrow \mathbb{R}$ of the functional $\mathcal{E}$ is measurable with respect to the decision maker information iff*

$$\int f dP = \int g dP \qquad \textit{for all } P \in \mathcal{C} \qquad \Longrightarrow \qquad \alpha(f) = \alpha(g)$$

That is, if two acts have the same evaluation according to each and every Hahn-Banach extension, then they should be evaluated in the same way. Transparently, this condition guarantees that the decision maker uses only the knowledge of the (set of) Hahn-Banach extensions of $\mathcal{E}$ to define the functional I.

As we have already observed, both functionals $\underline{\mathrm{I}}$ and $\bar{I}$ are measurable with respect to the decision maker's information. These functionals also satisfy an extra property that is not necessarily shared by a general functional of the form (2): translation invariance, that is for all $a \geq 0$, $b \in \mathbb{R}$ it holds that $\underline{\mathrm{I}}(af + b\mathbf{1}) = a\underline{\mathrm{I}}(f) + b$ and $\bar{I}(af + b\mathbf{1}) = a\bar{I}(f) + b$. In the setting we have been focusing on, translation invariance is an important property. In fact, we have been assuming the existence of an affine utility function on the outcome space X, and this utility is unique only up to a positive affine transformation. In this context, the translation invariance of the preference functional I is the same as the equivariance of I with respect to transformations of the utility on the outcome space, that is the statement that the representation of the decision maker's preference does not depend on the choice of the utility (within the class of admissible utilities). Because of this, we believe that translation invariance is a property that should be imposed on all the admissible extensions of $\mathcal{E}$. Be that as it may, when imposed, translation invariance also translates into a restriction on the coefficient $\alpha(\cdot)$ in (2). The theorem below summarizes our findings (the proof of (4) is omitted as it follows from simple algebra).

Theorem 27 *A decision maker facing Ambiguity is characterized by a preference functional $I : B(\Sigma) \longrightarrow \mathbb{R}$ with the following properties:*

(1) I is a positive linear functional on the linear subspace of subjectively measurable acts, which is determined by (1) in Proposition 18.

(2) I is of the form

$$I(f) = \alpha(f) \min_{P \in \mathcal{C}} \int f dP + (1 - \alpha(f)) \max_{P \in \mathcal{C}} \int f dP$$

where $\mathcal{C}$ is a (unique) convex, weak-compact subset of $ba_1^+(\Sigma)$ and $\alpha(f) \in [0,1]$.*

(3) The coefficient $\alpha(\cdot)$ in (2) satisfies

$$\int f dP = \int g dP \qquad \text{for all } P \in \mathcal{C} \qquad \Longrightarrow \qquad \alpha(f) = \alpha(g)$$

(4) I is translation invariant iff for all $a \geq 0$, $b \in \mathbb{R}$ and $f \in B(\Sigma)$, the coefficient $\alpha(\cdot)$ in (2)

satisfies

$$\alpha(af + b\mathbf{1}) = \alpha(f)$$

With Theorem 27, we have characterized the class of preference functionals that emerge in response to Ambiguity. The next theorem expresses the same characterization in terms of properties of a preference relation on the set of acts.

Theorem 28 *Consider a decision maker whose preference functional satisfies properties (1) to (4) in Theorem 27. Then his preference relation on the set of all acts is an Invariant Bi-separable preference.*

Proof. This is an entirely standard argument. Axiom A5 of Section 2 is satisfied (by construction) by virtue of our identification of the acts with $B(\Sigma)$. Since the preference functional is $\mathbb{R}$-valued, A1 is satisfied. Property (4) implies that A2 is satisfied, and the monotonicity of the functional implies that so is A4. By virtue of a well-known elementary argument, monotonicity and translation invariance of the functional imply its sup-norm continuity which, in turn, implies A3. ■

Remark 29 *It is worth noticing that the combination of Theorem 27 and Theorem 28 casts an entirely new light on property (3) in Theorem 27. This property originally appeared in [10] as a feature of Invariant Bi-separable preferences. Now, Theorem 27 and Theorem 28 tell us something entirely new: this property is precisely what indicates that the decision maker extends his preference functional on the basis of limited information.*

As a consequence of the previous theorem and of Theorem 1, we also have

Corollary 30 *The preference functional of a decision maker facing Ambiguity is of the form*

$$I(f) = \int_{\mathcal{C}} \kappa(f) d\nu = \int_{\mathcal{C}} \int_{S} f dP d\nu(P)$$

where ν is a capacity on the Borel subsets of $\mathcal{C}$.

This corollary clarifies the intuition that we gave in Remark 4 of Section 3 and provides the formal ground for that suggestion: *Integrating over priors with bad information gives rise to a non-additive integral* or, if the reader prefers, *Lebesgue integration over a non-measurable partition is represented by a Choquet integral.*

10.3 Symmetry considerations

By Theorem 28, the preferences that emerge in situations of Ambiguity are Invariant Bi-separable. Moreover, Theorem 27 makes it clear that all of its subclasses obtain by imposing restrictions on the coefficient $\alpha(f)$, that is on the function $\alpha(\cdot) : B(\Sigma) \longrightarrow \mathbb{R}$ in (2). In this way, one derives, for instance, the models of maxmin-EU, maxmax-EU, Choquet EU, etc. In this section, we are going to focus on a particular restriction on the function $\alpha(\cdot)$, one that is motivated by the examples that we have encountered in this paper.

We begin with the three-point version of Ellsberg's experiment discussed at the end of Section 9. When we introduced Ellsberg's experiment in Section 6.2, we noticed that the behavior which is typically observed mirrors the symmetry in the information: to an information that is symmetric in B and Y, there corresponds a table of preferences that is symmetric in B and Y. We can now express this properly.

Example 31 (Ex. 22 Sec. 9 ctnd) *Consider Example 22 of Section 9, and let $\succsim$ be a preference relation represented by a functional satisfying conditions (1) to (3) in Theorem 27. The preference $\succsim$ is symmetric with respect to B and Y iff $\alpha(B) = \alpha(Y)$. Consequently, since the specification of the coefficient α is immaterial in the case of unambiguously measurable acts, the preference functional representing a symmetric preference is of the form*

$$I(f) = \alpha \min_{P \in \mathcal{C}} \int f dP + (1 - \alpha) \max_{P \in \mathcal{C}} \int f dP$$

Thus, the model with α constant emerges in the three-point example as the only model that treats B and Y symmetrically.

Example 32 (Ex. Sec. 6.3 and Ex. Sec. 6.4 ctnd) *Similar considerations can be made for the other two versions of Ellsberg' experiment studied in 6.3 and Sec. 6.4, and then again in Sec. 9. In these examples, however, the condition $\alpha(B) = \alpha(Y)$ is not sufficient to obtain the α-MEU model (i.e., the model with α constant). This is so because the condition $\alpha(B) = \alpha(Y)$ does not imply anything with regard to the (strict) subsets of B and Y and on the coefficient that should be assigned to the bets on those subsets.*

Example 33 (Ex. Sec. 6.1 ctnd) *In the measure equivalence example of Sec. 6.1, there are nearly no extra conditions that can be imposed in a natural way on the extension from the set of subjectively measurable (= unambiguous) acts. In fact, for $\mathcal{C} = ca_1^+(\Sigma)$, not even condition (3) bites, and every monotone extension is automatically measurable with respect to the decision maker's information. It is still meaningful, however, to talk about symmetric extensions as those extensions that treat all acts in the same way, that is $\alpha(f) = \alpha(g)$ for all $f, g \in B(\Sigma)$. In this sense, α-MEU is the only (fully) symmetric extension in the measure equivalence example.*

The corollary below summarizes the discussion of this section on the α-MEU model (notice that measurability of α-MEU with respect to the decision maker's information is automatic).

Corollary 34 *Suppose that the information structure available to the decision maker is such that condition (6) in Proposition 18 is satisfied. Then, α-MEU is the only fully symmetric measurable extension from the set of subjectively measurable acts.*

11 Perceived ambiguity and the set of predictives in non-EU theories

The search for indicators of the ambiguity perceived by a decision maker has been a recurrent theme in the literature on Ambiguous Events. For the class of preferences determined above, it is well-known that the set of priors appearing in the representations of those non-Eu functionals cannot be taken, at least at face value, as one such indicator. This is so because different combinations of sets of priors and functions $\alpha(\cdot)$ may give rise to the same functional. A well-studied example is provided by the α-MEU model encountered above: for $0 \leq \alpha \leq 1$, one can start with a set of priors $\mathcal{C}$, define the α-MEU functional by using this $\mathcal{C}$, and then show that there exists another representation of the same preference with a set of priors $\mathcal{C}' \neq \mathcal{C}$ and a (typically) non-constant functions $\alpha(\cdot)$ (see, for instance, [7]). Truth to say, there is nothing peculiar to α-MEU in this. It even happens in the context of Expected Utility: take a set of priors $\mathcal{C}$ and a measure μ on it, and assume that the information is the best possible like in Section 4; then, the pair $(\mathcal{C}, \mu)$ defines an EU functional but the exact same functional is defined by the pair $(\{P^*\}, \delta(P^*))$, where P^* is the average measure as in Section 4 and δ is the Dirac measure. It is also clear that there are, generally speaking, many pairs $(\mathcal{C}', \mu')$ which define the same linear functional as that defined by P^*. Classically, the (uniquely determined) measure P^* is the decision maker's "predictive": given his information – a pair of the form $(\mathcal{C}', \mu')$ and a measurable partition – a Bayesian decision maker picks the best (in a Bayesian sense) distribution to *predict* the next draw from set of states S. In principle, P^* need not even be in $\mathcal{C}'$ (but, necessarily, it is in its convex hull). The situation is analogous in the non-EU case: given his information (in particular, a set of possible probabilistic views of the world) the decision maker determines a set of predictives. The only difference is that this set is not a singleton, which is what "reveals" the presence of Ambiguity. Two questions need to be addressed, at this point: What is the set of predictives in the non-Eu case? and, What is an indicator of the Ambiguity perceived by the decision maker? In particular, one might want to ask if the two necessarily coincide. We are going to answer these questions in the remainder of this section and, in doing so, we will be able to refine on an observation that we made elsewhere ([2, Section 3.1]).

At this point of the exposition, it is quite obvious that an indicator of the Ambiguity perceived by the decision maker is the subspace ΣMA of subjectively measurable acts: the larger this subspace the less the Ambiguity which is perceived.[6] As it is desirable of any Ambiguity index, this indicator is independent of the decision maker's attitude toward Ambiguity. It is so because the subspace ΣMA is determined before the extension of the preference functional, and the attitude toward ambiguity appears only in that process of extension. Another indicator, equivalent to the one just proposed, but with the additional appeal of representing the Ambiguity by means of a set of probability measures, can be achieved as follows. The subspace ΣMA and the linear functional defined on it are uniquely associated to the set of Hahn-Banach extensions of that functional which, in turn, can be represented (via the Riesz theorem) by a set of probability measures. Let us denote this set by $\mathcal{C}_{\Sigma MA}$. Because the association is unique, this set provides us with the same information as ΣMA: the smaller ΣMA, the larger $\mathcal{C}_{\Sigma MA}$ and the larger the perceived Ambiguity. We can use these observations to make comparisons across decision makers with regard to the Ambiguity that they perceive as well as their aversion toward that Ambiguity. Let dm_i denote a generic decision maker and let ΣMA_i be his subspace of subjectively measurable acts.

[6] We dropped the reference to the system cp as that is immaterial to the discussion of this section.

Definition 35 *We say that dm_1 perceives more ambiguity than dm_2 iff $\Sigma MA_1 \subseteq \Sigma MA_2$ iff $\mathcal{C}_{\Sigma MA_1} \supseteq \mathcal{C}_{\Sigma MA_2}$.*

Also, by denoting by I_i the preference functional of dm_i and by $\alpha_i(\cdot)$ the function that appears in that functional as obtained in Section 10.2, we have

Definition 36 *We say that dm_1 is more ambiguity averse than dm_2 iff $\alpha_1(f) \geq \alpha_2(f)$ for every $f \in B(\Sigma)$.*

Given the meaning of $\mathcal{C}_{\Sigma MA_i}$, Definition 36 follows naturally because it states that dm_1 has a consistently more pessimistic attitude than dm_2. Be it clear, however, that *this comparison refers to the representations of I_i which use $C_{\Sigma MA_i}$ and to those representations only*, and does not extend to other representations of the decision makers' preference functionals. This qualification is necessary because the sets of probability measures appearing in those representations cannot be interpreted, at least not in an obvious way, as indicators of the perceived ambiguity. It is worth stressing that the comparison in Definition 36 does not require, unlike [10], the sets $\mathcal{C}_{\Sigma MA_1}$ and $\mathcal{C}_{\Sigma MA_2}$ to be in any pre-specified relation; that is, Definition 36 allows for comparisons of decision makers who might have very different perceptions of ambiguity.

On a logical ground, Definition 35 and Definition 36 seem to us rather uncontroversial. Their potential weakness is that the set ΣMA_i (or $\mathcal{C}_{\Sigma MA_i}$), which is their crucial ingredient, is not immediately derivable from the decision maker's choices. As we know, ΣMA_i is a linear subspace with the property that the decision maker's preference functional is linear on it. Thus, in an attempt to uncover ΣMA_i from the decision maker's choices, one would have to look for something like "the largest subspace where the preference functional is linear". An attempt of this sort, however, would inevitably run into two problems. The first, somewhat minor, is that the notion of "largest subspace where the preference functional is linear" is ill-defined. The second, more substantial, is that by following that strategy one might end up determining something other than ΣMA_i. The reason is that some "extra-linearity" may appear when extending the preference functional from ΣMA_i, which would result in the existence of subspaces where the preference is linear which are strictly larger than ΣMA_i. Thus, by treating one of these subspaces as ΣMA_i, one would effectively underestimate the ambiguity perceived by the decision maker (equivalently, one would determine a set of priors strictly smaller than $\mathcal{C}_{\Sigma MA_i}$). In passing, we should like to observe that the tendency of underestimating ambiguity is inherent to all those approaches that derive Ambiguity indicators on the basis of the geometric properties of the preference functional (hence, of the decision maker's choices). We will come back to this point momentarily. For now, we are going to elaborate more on the two issues just mentioned. This will shed light on the problem of finding the set of predictives and of unveiling its relation with the set $\mathcal{C}_{\Sigma MA}$. We begin with a simple example showing us that extra-linearities can pop up easily in the extension process.

Example 37 *Let $\mathcal{C}$ be an arbitrary, non-singleton, set of probability measures. Define the set UMA, and let $\mathcal{E}$ be the associated linear functional on it. Since $\mathcal{C}$ is not a singleton, $B(\Sigma)\backslash UMA$ is non-empty. Let I be the α-MEU extension of $\mathcal{E}$ to $B(\Sigma)$. Then, it is easy to see that if $\alpha = 1/2$ the functional I is linear on $lin\{f, UMA\}$ for every $f \in B(\Sigma)\backslash UMA$.*

The example shows that we can exhibit a decision maker for whom $\Sigma MA_i = UMA$, who has a non-linear (α-MEU) preference functional on the set of all acts and whose preference functional is linear on many subspaces which all contain ΣMA_i strictly. Elaboration on the same example also shows that we can exhibit a decision maker who satisfies all the above properties and whose preference functional is, in addition, never linear on $lin\{f, g, UMA\}$ for $f, g \in B(\Sigma)\backslash UMA$ and $f \neq g$. Thus, in particular, the notion of "largest subspace on which I is linear" is meaningless. This motivates the following definitions.

Definition 38 *Let I be a preference functional satisfying the conditions of Theorem 27. We say that*

(a) A subspace $\mathcal{S} \subset B(\Sigma)$ is an AI subspace for I if I is linear on $\mathcal{S}$.

(b) An AI subspace $\mathcal{S}$ is said to be saturated if for any $g \in B(\Sigma)\backslash\mathcal{S}$, $lin\{g, \mathcal{S}\}$ is not an AI subspace.

Let $\mathcal{SAI}$ denote the class of all saturated AI subspaces, and define

$$\mathcal{K}^* = \bigcap_{\mathcal{S}_\gamma \in \mathcal{SAI}} \mathcal{S}_\gamma$$

That is, $\mathcal{K}^*$ is the intersection of all saturated AI subspaces. Since $\mathcal{K}^*$ is itself a linear subspace and I is linear on $\mathcal{K}^*$, we can determine the set of all Hahn-Banach extensions of the restriction of I to

$\mathcal{K}^*$ and, hence, associate $\mathcal{K}^*$ to a set of probability measures $\mathcal{C}^*$. Proposition 39 below shows that $\mathcal{C}^*$ is the set $\mathcal{C}_{GMM}$ that Ghirardato el al. derive in [10] and that they interpret as an indicator of the ambiguity perceived by the decision maker. The proposition also shows that $\mathcal{K}^*$ is equal to the set of crisp acts of [10], and gives the relation of these two sets with $\mathcal{C}_{\Sigma MA}$ and ΣMA, respectively.

Proposition 39 *Let I be a preference functional satisfying the condition of Theorem 27. We have*

(a) The set $\mathcal{C}^*$ is equal to the set $\mathcal{C}_{GMM}$ in [10, Prop. 4]

(b) $\mathcal{K}^*$ is equal to the set of crisp acts in [10, Prop. 10]

(c) $\mathcal{K}^* \supseteq \Sigma MA \supseteq UMA$

(d) $\mathcal{C}^* \subseteq \mathcal{C}_{\Sigma MA}$

Proof. (a) Consider the set of all Hahn-Banach extensions of the restriction of I to $\mathcal{K}^*$. By Lemmata 24 and 25 and [2, Theorems 1 and 2], I admits a representation of the form

$$I(f) = \int_{\mathcal{C}^*} \kappa(f) d\nu$$

where ν is a capacity on the Borel subsets of $\mathcal{C}^*$. Define now a preference relation $\succsim^*$ on $B(\Sigma)$ by

$$f \succsim^* g \qquad iff \qquad \int f dP \geq \int g dP \qquad \text{for all } P \in \mathcal{C}^* \qquad iff \qquad \kappa(f) \geq \kappa(g)$$

By the monotonicity of the Choquet integral, we have that $\kappa(f) \geq \kappa(g)$ implies $I(f) \geq I(g)$, which shows that $f \succsim^* g$ implies $f \succsim g$, that is $\succsim^*$ is a sub-relation of the decision maker's preference relation $\succsim$. Trivially, the relation $\succsim^*$ satisfies the Independence Axiom; hence, $\succsim^*$ is also a sub-relation of the unambiguous preference relation of [10] since the latter is the maximal restriction of $\succsim$ which satisfies the Independence Axiom (see [10, Prop. 3 part (7)]). This shows that $\mathcal{C}^* \supseteq \mathcal{C}_{GMM}$. We want to show that equality holds. To begin, notice that the set of all Hahn-Banach extensions of I coincide on $\mathcal{K}^*$, that $I(\psi) = \int \psi dP$ for all $\psi \in \mathcal{K}^*$, and that this value is independent of the choice of P in $\mathcal{C}^*$. Thus, $\mathcal{K}^*$ is a subspace of (in fact, equal to) the linear subspace of $B(\Sigma)$ defined by

$$\left\{ l \in B(\Sigma) \mid \int l dP = \int l dP' \text{ for all } P, P' \in \mathcal{C}^* \right\}$$

Suppose now that the inclusion $\mathcal{C}^* \supseteq \mathcal{C}_{GMM}$ is strict, that is there exists a $P \in \mathcal{C}^* \backslash \mathcal{C}_{GMM}$. Let

$$\mathcal{K}' = \left\{ l \in B(\Sigma) \mid \int l dP = \int l dP' \text{ for all } P, P' \in \mathcal{C}_{GMM} \right\}$$

Then $\mathcal{K}' \supset \mathcal{K}^*$ strictly. But, [10, Prop. 10 part (iv)] implies that $\mathcal{K}'$ belongs to any saturated AI subspace, hence to their intersection. Since this contradicts the definition of $\mathcal{K}^*$, we conclude that $\mathcal{C}^* = \mathcal{C}_{GMM}$.

(b) By [10, Prop 10 part (iii)], $\mathcal{K}'$ is the set of crisp acts, and $\mathcal{C}^* = \mathcal{C}_{GMM}$ implies $\mathcal{K}' = \mathcal{K}^*$.

(c) Since the decision maker's preference relation is, by construction linear on ΣMA, ΣMA is contained in all saturated AI subspaces, hence in $\mathcal{K}^*$.

(d) Follows at once from (c). ■

Ghirardato et al. use the set $\mathcal{C}^*$ as an indicator of the ambiguity perceived by the decision maker. Since $\mathcal{C}^*$ is possibly smaller than $\mathcal{C}_{\Sigma MA}$, their indicator tends to underestimate this ambiguity: Our reasoning above shows that (a) we can start off with a decision maker whose space of subjectively measurable acts is strictly included in $\mathcal{K}^*$; (b) determine $\mathcal{C}_{\Sigma MA}$; (c) write down his preference functional as $I_{\mathcal{C}_{\Sigma MA}}(f) = \alpha(f) \min_{\mathcal{C}_{\Sigma MA}} \kappa(f) + (1 - \alpha(f)) \max_{\mathcal{C}_{\Sigma MA}} \kappa(f)$; and, finally (d) choose the function $\alpha(\cdot)$ so as to guarantee that the intersection of all saturated AI subspaces of $I_{\mathcal{C}_{\Sigma MA}}$ is precisely $\mathcal{K}^*$. Notice, in particular, that the choice of the function $\alpha(\cdot)$ enters the determination of $\mathcal{K}^*$, which shows that the information contained in $\mathcal{K}^*$ (equivalently, in $\mathcal{C}^*$) inevitably mixes the perception of ambiguity with the attitude toward it. A moment of thought shows that this problem is inherent to all those approaches that determine, like in the case of Ghirardato et al.,[7] indicators of Ambiguity by using only geometric properties of the preference functional.

The set $\mathcal{C}^*$ of Ghirardato et al., however, still plays a very important role, this time exactly because it is derived only on the basis of the geometric properties of the preference functional: the set $\mathcal{C}^*$ represents that Ambiguity that cannot be reduced by the decision maker when he extends the preference functional from ΣMA. As such, the set $\mathcal{C}^*$ (or rather the set of its extreme points) represents the set of predictives: those probability measures that the decision maker actually uses to evaluate the acts. In general, there is nothing strange about the fact that this set might be smaller than the one representing the ambiguity. As we saw above, extra-linearities (effectively

[7] Ghirardato et al. determine the set $\mathcal{C}^*$ by using the decision maker's unambiguous preference relation. Loosely speaking, this is [10, Prop. 10 part (7)] the part of the functional I which satisfies linearity.

"reductions" of Ambiguity) may emerge during the process of extension from ΣMA, and take the form of restrictions on the function $\alpha(\cdot)$ generated in that extension. These restrictions may be motivated by various considerations like, for instance, the symmetry considerations that we encountered in Ellsberg's experiments.

Example 40 *Restrictions of the form $\alpha(f) + \alpha(-f) = 1$ have the feature that the extended functional is always linear on $lin\{f, UMA\}$ and might express conditions on how to evaluate acts when on different sides of the market (buy or sell).*

To summarize, in our opinion the correct indicator of the Ambiguity perceived by the decision maker is the set $\mathcal{C}_{\Sigma MA}$ (or, equivalently, the linear subspace ΣMA). The set of predictives is the set $\mathcal{C}^*$ of Ghirardato, Maccheroni and Marinacci [10] and represents that *ambiguity that cannot be reduced* when evaluating acts. $\mathcal{C}^*$ is always contained in $\mathcal{C}_{\Sigma MA}$, and to use $\mathcal{C}^*$ as an indicator of the Ambiguity perceived by the decision maker may lead to underestimate that Ambiguity. We conclude by finding an "upper bound" for the set $\mathcal{C}_{\Sigma MA}$ that can be derived from the decision maker's choices. The basic idea is to relate the representation of Invariant Bi-separable preferences given by Ghirardato et al. to other possible representations, and try to detect the extra-linearities discussed above. The key observation is provided by part (c) of Proposition 39 above, which tells us that, as extra-linearities appear, the crisps act may grow from UMA to $\mathcal{K}^*$. For $\mathcal{C}$ an arbitrary set of probability measures, let us denote by $I_{\mathcal{C}}$ the functionals of the form

$$I_{\mathcal{C}}(f) = \alpha_{\mathcal{C}}(f) \min_{\mathcal{C}} \kappa(f) + (1 - \alpha_{\mathcal{C}}(f)) \max_{\mathcal{C}} \kappa(f) \tag{3}$$

where $f \in B(\Sigma)$, $\kappa(f) : \mathcal{C} \longrightarrow \mathbb{R}$ is defined by $\kappa(f)(P) = \displaystyle\int f dP$ and $\alpha_{\mathcal{C}}(f) \in [0,1]$. Also, let $\underset{\mathcal{C}}{var} \kappa(f) = \max_{\mathcal{C}} \kappa(f) - \min_{\mathcal{C}} \kappa(f)$. Now, let $\succsim$ be an Invariant Bi-separable preference. By [10], $\succsim$ admits a representation by a functional I of the type

$$I_{\mathcal{C}^*}(f) = \alpha_{\mathcal{C}^*}(f) \min_{\mathcal{C}^*} \kappa(f) + (1 - \alpha_{\mathcal{C}^*}(f)) \max_{\mathcal{C}^*} \kappa(f)$$

where $\mathcal{C}^*$ is the set of predictives discussed above. For $\mathcal{K}^*$ the corresponding set of crisp acts, let us say that an act $\zeta \in \mathcal{K}^*$ satisfies Property $\mathcal{L}$ if there exists a representation of $\succsim$ of the type (3) such that either (a) $\underset{\mathcal{C}}{var}\kappa(\zeta) = 0$; or, (b) $\alpha_{\mathcal{C}}(\zeta) = \frac{\underset{\mathcal{C}}{\max}\kappa(\zeta) - \int \zeta dP}{\underset{\mathcal{C}}{var}\kappa(\zeta)}$ for $P \in \mathcal{C}^*$.

Proposition 41 *Let $\hat{C}$ be largest subset of $ba_1^+(\Sigma)$ such that every act in $\mathcal{K}^*$ has Property $\mathcal{L}$. Then, $\mathcal{C}^* \subseteq \mathcal{C}_{\Sigma MA} \subseteq \hat{C}$.*

Notice that when $\mathcal{C} = \mathcal{C}^*$, every $\zeta \in \mathcal{K}^*$ satisfies Property $\mathcal{L}$ with $\underset{\mathcal{C}}{var}\kappa(\zeta) = 0$. When we move to representations of the same preference which use a $\mathcal{C}$ which strictly contains $\mathcal{C}^*$, for some $\zeta \in \mathcal{K}^*$ the property $\underset{\mathcal{C}}{var}\kappa(\zeta) = 0$ has to disappear. Yet, the new functional still has to stay linear on $\mathcal{K}^*$, which means that the function $\alpha_{\mathcal{C}}(\cdot)$ is to be chosen so to preserve this linearity. This is exactly what part (b) of Property $\mathcal{L}$ demands. Notice that an arbitrary set $\mathcal{C}$ may not have the property that every act in $\mathcal{K}^*$ has Property $\mathcal{L}$. This is so because the necessity of satisfying part (b) of Property $\mathcal{L}$ may be in conflict with other properties that the function $\alpha(\cdot)$ has to satisfy, for instance properties (3) and (4) in Theorem 27 (that is, if $\mathcal{C}$ is such that every act in $\mathcal{K}^*$ has Property $\mathcal{L}$, then $I_{\mathcal{C}}$ may not be a representation of $\succsim$). If there exists a set $\mathcal{C}$ strictly larger than $\mathcal{C}^*$ with the property that every act in $\mathcal{K}^*$ has Property $\mathcal{L}$, then we can think of $\mathcal{C}$ as representing the Ambiguity and of $\mathcal{C}^*$ as its set of predictives. From what we have seen above, it follows that $\mathcal{C}^* \subseteq \mathcal{C}_{\Sigma MA} \subseteq \hat{C}$.

Proposition 41 provides us with a method for estimating $\mathcal{C}_{\Sigma MA}$ as well as for identifying those situations in which $\mathcal{C}^* = \mathcal{C}_{\Sigma MA}$. Clearly, the effectiveness of Proposition 41 is greatly enhanced when some extra information, in the form of a restriction on the function $\alpha(\cdot)$, is available. For instance, it is clear that by restricting to the class of α-MEU preferences, or by simply knowing that the preference has an α-MEU representation, considerably reduces the ability of satisfying part (b) of Property $\mathcal{L}$ and, hence, makes it more likely that $\mathcal{C}^* = \mathcal{C}_{\Sigma MA}$.

APPENDICES

A Conditional measures

Let $(\mathcal{C}, \mathcal{B}, \mu)$ be a measure space, let $\mathcal{I}$ be a partition of $\mathcal{C}$ (modulo μ-measure 0 events) and let $\mathcal{C}/\mathcal{I}$ denote the quotient space. Let $\pi : \mathcal{C} \longrightarrow \mathcal{C}/\mathcal{I}$ be the canonical projection. The canonical σ-field on $\mathcal{C}/\mathcal{I}$ is the finest σ-field that makes the canonical projection measurable. The measure structure induced by π on $\mathcal{C}/\mathcal{I}$ consists of the canonical σ-field and the image measure (pushforward) of μ under π. We recall the following definition.

Definition 42 *A canonical system of conditional measures associated to the partition $\mathcal{I}$ is a family of measures $\{\mu_\iota, \iota \in \mathcal{I}\}$, with the following properties*

(i) for any $A \in \mathcal{B}$, the set $A \cap \iota$ is measurable in ι for almost all $\iota \in \mathcal{C}/\mathcal{I}$ and the function $\mu_\iota(A \cap \cdot) : \mathcal{C}/\mathcal{I} \to \mathbb{R}$ is measurable; and

(ii) for any $A \in \mathcal{B}$,

$$\mu(A) = \int_{\mathcal{C}/\mathcal{I}} \mu_\iota(A \cap \iota) d\mu'$$

where μ' is the pushforward of μ under the canonical projection $\pi : \mathcal{C} \longrightarrow \mathcal{C}/\mathcal{I}$.

B Standard Spaces

A *Polish space*, (X, τ), is a separable, completely metrizable topological space. Given the topology τ on X, the Borel σ–field is the one generated by the closed sets. A *Standard Borel space* is a Polish space stripped down to its Borel structure.

Given two measurable spaces, $(X_1, \mathcal{B}_1)$ and $(X_2, \mathcal{B}_2)$, a mapping $X_1 \longrightarrow X_2$ is called a Borel isomorphism if it is a bijection and is bimeasurable.

Borel isomorphism theorem (see [27, Theorem 3.3.13]) Any two uncountable standard Borel spaces are Borel isomorphic.

A Standard Borel space along with a finite nonatomic measure is a called a *Standard Lebesgue space*. A measurable set in a Standard Lebesgue space is a set which differs from a Borel set by a set of measure zero.

Given two measure spaces, $(X_1, \mathcal{B}_1, m_1)$ and $(X_2, \mathcal{B}_2, m_2)$, a measurable mapping $T : X_1 \longrightarrow X_2$ is a *measure preserving transformation* if for all $E \in \mathcal{B}_2$ we have

$$m_1(T^{-1}(E)) = m_2(E)$$

If T is bijective and its inverse T^{-1} is also measure-preserving, then T is an *invertible measure-preserving transformation.* Two measure spaces, $(X_1, \mathcal{B}_1, m_1)$ and $(X_2, \mathcal{B}_2, m_2)$, are isomorphic if there exists an invertible measure preserving transformation $T : X_1 \longrightarrow X_2$.

Isomorphism of Lebesgue Spaces (see [29, Theorem 2.1]) Any two Standard Lebesgue spaces are isomorphic.

C A Polish setting

In combination, the two assumptions below guarantee that the set of measures $\mathcal{C}$ in Theorem 1 is a Polish space.

Standard State Space The measurable space (S, Σ) is a standard Borel space.

Let $\succsim$ be a preference relation satisfying Axioms 1 to 5. Let $\trianglerighteq$ denote the unambiguous preference relation ([10], Sec. B.3) associated to $\succsim$.

Axiom of Monotone Continuity (see [10]) For all $x, y, z \in X$ such that $y \rhd z$, and all sequences of events $\{A_n\}_{n \geq 1} \subseteq \Sigma$ with $A_n \downarrow \emptyset$, there exists $\bar{n} \in \mathbb{N}$ such that $y \rhd x A_{\bar{n}} z$.

The Axiom of Monotone Continuity is equivalent to the property that all the measures in Theorem 1 are countably additive [10, Sec. B.3]. Let $\mathcal{P}(\Sigma)$ denote the space of regular Borel measures on Σ.

Theorem 43 *Let (S, Σ) be a standard Borel space. A preference relation $\succsim$ on $\mathcal{A}$ satisfies Axioms 1 to 5 and the Axiom of Monotone Continuity iff Theorem 1 holds and the set $\mathcal{C} \subset \mathcal{P}(\Sigma)$. In particular, $\mathcal{C}$ is a Polish space.*

In the course of the proof, we will denote by $\sigma(ba(\Sigma), Y)$ the weak topology on $ba(\Sigma)$ induced by a set of mappings Y.

Proof. Assume that Theorem 1 holds with $\mathcal{C} \subset \mathcal{P}(\Sigma)$. The function κ defined in Theorem 1 is sup-norm to sup-norm continuous as a consequence of the inequality

$$\begin{aligned} \|\kappa(f) - \kappa(g)\|_\infty &= \sup_{P \in \mathcal{C}} \left| \int f dP - \int g dP \right| \leq \sup_{P \in \mathcal{C}} \int |f - g| \, dP \\ &\leq \sup_{P \in \mathcal{C}} \int \sup_{s \in S} |f - g| \, dP = \|f - g\|_\infty \end{aligned}$$

Since the Choquet integral is sup-norm continuous, it follows that the preference defined by the functional in Theorem 1 satisfies Axioms 1 and 3. By the monotonicity and translation invariance of both the Lebesgue integral and the Choquet integral, Axioms 2 and 4 are also satisfied. Finally, the Axiom of Monotone Continuity is satisfied by [10, Sec. B.3]. Now, the converse. Since $\succsim$ satisfies Axioms 1 to 5, Theorem 1 holds as stated in Section 2. We want to show that, when the Axiom of Monotone Continuity is satisfied, Theorem 1 continues to hold with $\mathcal{C} \subset \mathcal{P}(\Sigma)$ and that this space is Polish. Our strategy will be as follows. We are going to change the topology on the set $ba(\Sigma)$ (hence, on $\mathcal{C}$). We will show that, when the the Axiom of Monotone Continuity is satisfied, this change in the topology leaves $\mathcal{C}$ compact. As the new topology is coarser than the old one, the functions $\kappa(f) : \mathcal{C} \longrightarrow \mathbb{R}$ are no-longer necessarily continuous. We are going to show, however, that even with the new topology $\kappa(f)$ is always a measurable function for the Borel class defined by the new topology, $\forall f \in B(\Sigma)$. This will guarantee that the functional defined in Theorem 1 is still well-defined and, being defined pointwise in the exact same way, it still represents the same preference.

From Theorem 1, we know that $\mathcal{C}$ is a weak*-compact subset of $(ba(\Sigma), \sigma(ba(\Sigma), B(\Sigma)))$. By the Axiom of Monotone Continuity, all the probabilities in $\mathcal{C}$ are countably additive. By the assumption that (S, Σ) is standard Borel, it follows that all the probabilities in $\mathcal{C}$ are regular. If we replace the topology $\sigma(ba(\Sigma), B(\Sigma))$ with the topology $\sigma(ba(\Sigma), C_b(S))$ – $C_b(S)$ the set of continuous bounded functions on S – then $\mathcal{C}$ remains compact because the new topology is weaker than the original one. In particular, $\mathcal{C}$ is closed. Finally, (S, Σ) standard implies that the space $\mathcal{P}(\Sigma)$ is Polish in the topology $\sigma(\mathcal{P}(\Sigma), C_b(S))$, and we conclude that $\mathcal{C}$ is Polish as well. Next, define κ as in Theorem 1. Since all the measures in $\mathcal{C}$ are bounded and countably additive, the Monotone Convergence

Theorem implies that κ is normal, that is

$$f_n \nearrow f \Longrightarrow \kappa(f_n) \nearrow \kappa(f), \qquad n \in \mathbb{N}$$

Let $E \in \Sigma$, and let χ_E denote the indicator function of the set E. Then, $\kappa(\chi_E)$ is obviously bounded and it is well-known that $\kappa(\chi_E)$ is measurable for the Borel σ-algebra generated by $\sigma(\mathcal{P}(\Sigma), C_b(S))$ [1, Lemma 14.16]; that is, $\kappa(\chi_E) \in B(\mathcal{B})$ for all $E \in \Sigma$. If $h \in B(\Sigma)$ is a simple function, then h can be written as a finite linear combination of indicator functions, and linearity of κ along with the previous observation imply that $\kappa(h) \in B(\mathcal{B})$. Finally, if $f \in B(\Sigma)$ is any function, then there exists a sequence of simple functions $\{f_n\} \subset B(\Sigma)$ such that $f_n \nearrow f$, and normality of κ implies that $\kappa(f) \in B(\mathcal{B})$. We conclude that when $\mathcal{C}$ is equipped with the Polish topology $\sigma(\mathcal{P}(\Sigma), C_b(S))$, the linear mapping κ takes $B(\Sigma)$ into a subset of $B(\mathcal{B})$. ■

References

[1] Aliprantis C.D. and K.C. Border (2006), Infinite Dimensional Analysis, 3rd ed. Springer-Verlag.

[2] Amarante M. (2009), Foundations of Neo-Bayesian Statistics, *Journal of Economic Theory* **144**, 2146-73.

[3] Anscombe F.J. and R.J. Aumann (1963), A definition of subjective probability, *Annals of Mathematical Statistics* **34**, 199-205.

[4] Arrow K. (1970), Essays in the theory of risk-bearing, North-Holland.

[5] Billingsley P. (1995), Probability and Measure, Wiley.

[6] Cornfeld I.P., S.V. Fomin and Y.G. Sinai (1982), Ergodic Theory, Springer-Verlag.

[7] Eichberger J., S. Grant, D. Kelsey and G. Koshevoy (2011), The α-MEU Model: A Comment, *Journal of Economic Theory* **48**, 1684-1698.

[8] Epstein L. and M. Schneider (2003), Recursive multiple priors, *Journal of Economic Theory* **113**, 1-31.

[9] Gardenfors P. and N-E Sahlin (1982), Unreliable Probabilities, Risk-taking, and Decision Making, *Synthese* **53**, 361-86.

[10] Ghirardato P., F. Maccheroni and M. Marinacci (2004), Differentiating ambiguity and ambiguity attitude, *Journal of Economic Theory* **118**, 133-173.

[11] Ghirardato P., F. Maccheroni, M. Marinacci and M. Siniscalchi (2003), Subjective Foundations for Objective Randomization: A New Spin on Roulette Wheels, *Econometrica* **71**, 1897-1908.

[12] I. Gilboa and M. Marinacci (2011), Ambiguity and the Bayesian Paradigm, mimeo, http://www.tau.ac.il/~igilboa/pdf/Gilboa_Marinacci_Ambiguity_and_Bayesian_Paradigm.pdf

[13] Gilboa I. and D. Schmeidler (1989), Maxmin expected utility with a non-unique prior, *Journal of Mathematical Economics* **18**, 141-53.

[14] Hanany E. and P. Klibanoff (2009), Updating Ambiguity Averse Preferences, *The B.E. Journal of Theoretical Economics* **9**, 37.

[15] Hjorth G. (2000), Classification and orbit equivalence relations, American Mathematical Society.

[16] Kast R., A. Lapied and P. Toquebeuf (2012), Updating Choquet capacities: a general framewrok, *Economics Bulletin* **32**, 1495-1503.

[17] Kechris, A. and N.E. Sofronidis (2001), A strong generic ergodicity property of unitary and self-adjoint operators, *Ergodic Theory and Dynamical Systems* **21**, 1459-79.

[18] Maccheroni F. and M. Marinacci (2005), A Strong Law of Large Numbers for Capacities, *The Annals of Probability* **33**, 1171-78.

[19] Marinacci M. (1997), A simple proof of a basic result for multiple priors, unpublished.

[20] Marinacci M. (2002), Learning from ambiguous urns, *Statistical Papers* **43**, 143-151.

[21] Nehring K. (2001), Ambiguity in the context of probabilistic beliefs, UC Davis, mimeo.

[22] Phelps R. R. (1966), Lectures on Choquet theorem, van Nostrand.

[23] Rokhlin, V.A. (1949), On the Fundamental Ideas of Measure Theory, Translations of the AMS, 71.

[24] Savage L.J. (1972), The Foundations of Statistics, Dover.

[25] Schmeidler D. (1989), Subjective probability and expected utility without additivity, *Econometrica* **57**, 571-87.

[26] Siniscalchi M. (2011), Dynamic choice under ambiguity, *Theoretical Economics* **6**, 379-421.

[27] Srivastava, S.M. (1998), A Course on Borel Sets, Springer-Verlag GTM.

[28] P. Walley (1991), Statistical reasoning with imprecise probabilities, Chapman and Hall.

[29] Walters P., F. W. Gehring and K. A. Ribet (2000), An Introduction to Ergodic Theory, Springer.

Récents cahiers de recherche du CIREQ
Recent Working Papers of CIREQ

Si vous désirez obtenir des exemplaires des cahiers, vous pouvez les télécharger à partir de notre site Web http://www.cireqmontreal.com/cahiers-de-recherche

If you wish to obtain copies of the working papers, you can download them directly from our website, http://www.cireqmontreal.com/cahiers-de-recherche

05-2013 Okasha, S., J.A. Weymark, W. Bossert, "Inclusive Fitness Maximization : An Axiomatic Approach", mai 2013, 17 pages

06-2013 Ehlers, L., B. Klaus, "House Allocation via Deferred-Acceptance", juillet 2013, 13 pages

07-2013 McCausland, W.J., A.A.J. Marley, "Bayesian Inference and Model Comparison for Random Choice Structures", juillet 2013, 26 pages

08-2013 Cardia, E., P. Gomme, "The Household Revolution : Childcare, Housework, and Female Labor Force Participation", juillet 2013, 44 pages

09-2013 Dutta, R., S. Horan, "Inferring Rationales from Choice : Identification for Rational Shortlist Methods ", mars 2013, 51 pages

10-2013 Dutta, R., R. Ishii, "Coordinating by Not Committing : Efficiency as the Unique Outcome", août 2013, 38 pages

11-2013 Amarante, M., "A Representation of Risk Measures", octobre 2013, 8 pages

12-2013 Amarante, M., "A Characterization of Exact Non-atomic Market Games", octobre 2013, 8 pages

13-2013 Dufour, J.-M., L. Khalaf, M. Voia, "Finite-Sample Resampling-Based Combined Hypothesis Tests, with Applications to Serial Correlation and Predictability", octobre 2013, 23 pages

14-2013 Zhang, H.J., J.-M. Dufour, J.W. Galbraith, "Exchange Rates and Commodity Prices : Measuring Causality at Multiple Horizons", septembre 2013, 41 pages

15-2013 Gaudet, G., P. Lasserre, "The Taxation of Nonrenewable Natural Resources", octobre 2013, 41 pages

16-2013 Benchekroun, H., G. Gaudet, "On the Effects of Mergers on Equilibrium Outcomes in a Common Property Renewable Asset Oligopoly", octobre 2013, 33 pages

17-2013 Andersson, T., L. Ehlers, "An Algorithm for Identifying Agent-k-Linked Allocations in Economies with Indivisibilities", novembre 2013, 7 pages

01-2014 Alvarez-Cuadrado,F., N.V. Long et M. Poschke, "Capital-Labor Substitution, Structural Change and the Labor Income Share", janvier 2014, 55 pages

02-2014 Long, N.V., "The Green Paradox under Imperfect Substitutability between Clean and Dirty Fuels", février 2014, 23 pages

03-2014 Doko Tchatoka, F., J.-M. Dufour, "Identification-Robust Inference for Endogeneity Parameters in Linear Structural Models", janvier 2014, 29 pages