Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 150 (2014) 642–667

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Parametric representation of preferences ☆

Nabil I. Al-Najjar *, Luciano De Castro

Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston, IL 60208, United States

Received 12 October 2011; final version received 13 February 2013; accepted 23 July 2013

Available online 27 December 2013

Abstract

A preference is invariant with respect to a set of transformations if the ranking of acts is unaffected by reshuffling the states under these transformations. For example, transformations may correspond to the set of finite permutations, or the shift in a dynamic choice model. Our main result is that any invariant preference must be parametric: there is a unique sufficient set of parameters such that the preference ranks acts according to their expected utility given the parameters. Parameters are characterized in terms of objective frequencies, and can thus be interpreted as objective probabilities. By contrast, uncertainty about parameters is subjective. The preferences for which the above results hold are only required to be reflexive, transitive, monotone, continuous, and mixture linear.

JEL classification: A1

Keywords: Decision making; Uncertainty; Parameters

1. Introduction

This paper develops a general model for representing preferences in terms of parameters. In our representation the decision maker decomposes the uncertainty he faces into: (1) objective

☆ The first version of this paper, circulated in 2010, applied to decision theory results from our earlier work, De Castro and Al-Najjar (2009). We thank Paolo Ghirardato, Ben Polak and Marciano Siniscalchi for extensive discussions and thoughtful comments on the 2009 version of the project. We also thank Simone Galperti for his research assistance.

* Corresponding author.

E-mail addresses: al-najjar@northwestern.edu (N.I. Al-Najjar), decastro.luciano@gmail.com (L. De Castro).

URLs: http://www.kellogg.northwestern.edu/faculty/alnajjar/htm/index.html (N.I. Al-Najjar), https://netfiles.uiuc.edu/luciano/www/ (L. De Castro).

0022-0531/$ – see front matter
http://dx.doi.org/10.1016/j.jet.2013.12.006

parameter risk that can be characterized in terms of empirical frequencies, and (2) subjective uncertainty about parameters. In the stylized example of repeated coin tosses, whether a coin turns up *Heads* or *Tails* in any single toss is idiosyncratic, being the outcome of a multitude of complex factors. Roughly, parameters are the lens through which a decision maker decomposes the data into patterns and noise.

We consider a preference over acts on a state space Ω. The state space in our formal model is abstract and need not have an intertemporal or product structure. For concreteness, assume throughout this Introduction that Ω has the product structure $S \times S \times \cdots$, where each coordinate S represents the outcome of some experiment. We say that a preference has a *parametric representation* if there are distributions $\{P^{\theta}\}_{\theta \in \Theta}$ indexed by a set of parameters Θ and a decomposition map $\vartheta : \Omega \to \Theta$ such that for any pair of acts f, g[1]:

$$f(\cdot) \succcurlyeq g(\cdot) \quad \Longleftrightarrow \quad \int_{\Omega} f \, dP^{\vartheta(\cdot)} \succcurlyeq \int_{\Omega} g \, dP^{\vartheta(\cdot)}. \tag{1}$$

The distribution $P^{\vartheta(\omega)}$ captures the statistical patterns the decision maker associates with a sequence of outcomes ω. When (1) holds we say that the parametrization (Θ, ϑ) is *sufficient* for the preference: the decision maker's ranking of acts contingent on parameters fully captures his non-contingent ranking. The connection to the notion of sufficiency in statistics is obvious and discussed further below.

Our main theorem identifies conditions under which a preference has a *parametric representation* with respect to a uniquely defined set of parameters. The key condition we use is that the preference is invariant with respect to transformations of the state space. Perhaps the best known example of such transformations is the group of finite permutations, where one requires the preference to be invariant with respect to reshuffling of the coordinates. Permutations give rise to the i.i.d. parameters and, with additional conditions, to de Finetti's [8] celebrated representation theorem. In this paper we consider general countable semi-groups of transformations which cover exchangeability, but also partial exchangeability, stationary distributions, Markovian structures, among others.

The sufficiency of a parametrization defines an operator:

$$f \stackrel{\Psi}{\longmapsto} \int_{\Omega} f \, dP^{\vartheta(\cdot)}$$

that maps the state-based acts $\mathcal{F}$ to their corresponding elements in the set of *parameter-based acts* $\mathbb{F}$. A binary relation on $\mathbb{F}$ is called an *aggregator* and reflects how the decision maker subjectively aggregates the parameters in making decisions. If the aggregator $\succcurlyeq$ satisfies our basic conditions of reflexivity, transitivity, monotonicity, and continuity, then there is a unique preference $\succcurlyeq$ on $\mathcal{F}$ such that for every $f, g \in \mathcal{F}$

$$f \succcurlyeq g \quad \Longleftrightarrow \quad \Psi(f) \succcurlyeq \Psi(g).$$

The preference $\succcurlyeq$ is necessarily invariant and satisfies our basic conditions.

This provides a general template to incorporate subjective parameters into most known decision models. First, start with a semi-group of transformations and let Θ be the corresponding subjective set of parameters. For example, if you start with the permutations, then Θ is the set of

1 The notation $(\cdot)$ emphasizes that we are dealing with acts that take ω as argument.

i.i.d. parameters. Second, propose an aggregator of parameter uncertainty, perhaps corresponding to some compelling set of axioms (e.g., Bayesian belief over parameters, Bewley-style incomplete preferences, ..., etc.). Third, derive an invariant preference $\succcurlyeq$ on the state-based acts $\mathcal{F}$.

After introducing a general methodology for aggregating of parameter uncertainty in Section 4, we turn in Section 5 to the special case of aggregators that take the form:

$$\mathcal{V}(F) = \int_{\Theta} \phi\big(F(\theta)\big)\,d\mu(\theta) = \int_{\Theta} \phi\bigg(\int_{\Omega} u(f)\,dP^{\theta}\bigg)\,d\mu(\theta), \tag{2}$$

for $F = \Psi(f)$, a von Neumann–Morgenstern utility function u, and a function $\phi : \mathbb{R} \to \mathbb{R}$. We call the preferences corresponding to such aggregators *second-order subjective expected utility preferences*. These are preferences that (to our knowledge) were first introduced by Neilson [23,24] and used by, among others, Nau [21,22], Ergin and Gul [13], Chew and Sagi [5], Strzalecki [25], Grant et al. [16].

The models of Neilson [23] and Strzalecki [25] have an interesting interpretation in our setting. They consider functionals of the form[2]

$$\int_{\Omega} \phi\Big(\sum u(c)p_{f(\omega)}(c)\Big)\,dP(\omega),$$

where f is an Anscombe–Aumann act and $p_{f(\omega)}(c)$ is the probability of consequence c under the lottery $f(\omega)$. Writing $u(f) \equiv \sum u(c)p_{f(\omega)}(c)$, we can write the above in our notation as:

$$\int_{\Omega} \phi\bigg(\int_{\Omega} u(f)\,d\delta_{\omega}\bigg)\,dP,$$

where δ_{ω} is the measure that puts mass 1 on the state ω. This can be interpreted in our setting as follows: if the preference is invariant with respect to the trivial identity transformation, then the parameters are simply the Dirac measures δ_{ω} that put unit mass on a state ω, and the space of parameters is in fact Ω itself. This coincides with the second-order subjective expected utility representation (2) with the trivial identity transformation.

The aggregator (2) may therefore be viewed as a generalization of Neilson's representation to coarser parametrizations (e.g., where parameters are i.i.d. distributions). Invariance with respect to non-trivial transformations means that the decision maker pools many states into risky events $\vartheta^{-1}(\theta)$, while in Neilson [23] and Strzalecki [25] the risky events are singletons. We interpret Neilson's [23] and Strzalecki's [25] decision makers as ones who do not not do such pooling, so each state is its own parameter. Section 5.2 discusses this in more details.

Our model also helps clarify Klibanoff et al.'s [18] functional form which is similar to (2), but where the outer integral is over mixtures of parameters $\Delta(\Theta)$, rather than parameters Θ, and where their behavioral data includes choices over "second-order acts." As we discuss briefly in Section 5.3, and in more detail in our companion paper Al-Najjar and Castro [1], the behavioral content of their model is different from what we have here and in the other papers cited above as it requires the decision maker to express rankings over unobservable objects. Klibanoff et al. [18] assume that the decision maker can make bets that pay depending on which probability distribution on Ω obtains. The outcome of such "second-order acts" is unobservable, even in principle and in idealized experiments where infinite amount of data is available. For example, given two

[2] A main contribution of Strzalecki [25] is to characterize when ϕ gives rise to multiplier preferences.

i.i.d. parameters θ, θ', a second-order act would have to specify what the decision maker gets at a distribution $\frac{1}{2}\theta + \frac{1}{2}\theta'$, and which in turn may be different from what he would get at, say, $\frac{1}{4}\theta + \frac{3}{4}\theta'$. Even with infinite amount of data, all that one observes is $\vartheta(\omega)$ which, in this example, is either θ or θ', so payments contingent on whether a distribution $\alpha\theta + (1-\alpha)\theta'$ 'occurred' has no behavioral meaning in term of even hypothetical revealed preference experiments. Our framework does not appeal to unobservable second-order acts. Parameter-based acts are just ordinary acts (i.e., functions of ω) that happen to be measurable with respect to events of the form $\vartheta^{-1}(\theta) \subset \Omega$.

We close with two additional connections to the literature. First, parameters are obviously central in statistical theory and its applications. Although Bayesian and classical statistics differ in their approach to inference, both use data to learn the value of an unknown underlying parameter. In the statistics literature, parameters are usually formalized as extreme points of convex sets of distributions; see, for example, Dynkin [11], Dawid [6], Lauritzen [20] among others. From the perspective of economic and game theoretic modeling, the treatment of parameters in statistics is not completely satisfactory: parameters are either objective, a datum handed down as part of the description of the statistical model or, in Bayesian statistics, they are subjective but require a commitment to a Bayesian model of inference. In this paper parameters are part of the decision maker's subjective model of the world (reflected in his invariance assumptions) but require no commitment to a specific approach to resolving parameter uncertainty.

The second connection concerns the interpretation of subjectivity vs. objectivity in decision making. A common distinction is between objective parameters, interpreted as risk, and the subjective aggregation over parameters, interpreted as 'model uncertainty' (see for instance, the literature on robustness in macroeconomics). While intuitively compelling, formalizing this distinction is quite challenging.[3] In our framework, parameters are derived from the invariance of preferences to transformations, and they are therefore subjective. But parameters are also linked via a subjective ergodic theorem to the empirical frequencies, which are objective.

2. Preliminaries

2.1. States, acts and preferences

Given a Polish space X, i.e., a complete separable metrizable space with the Borel σ-algebra $\mathcal{X}$, let $\Delta(X)$ be the set of countably additive probability measures on $(X, \mathcal{X})$. Note that $\Delta(X)$ itself is a Polish space with its standard (weak*) topology. We will consider only measurable functions between Polish spaces, unless explicitly stated otherwise.

Our primitive is a binary relation $\succcurlyeq$ on acts defined on a compact Polish state space Ω with the Borel σ-algebra Σ. Assume that the space of consequences X is (a convex subset of) an Euclidean space. For instance, we can have $X = \Delta(C)$ where C is finite. Under the usual convex combination operation, the set X is a mixture space in the sense of Herstein and Milnor [17]. We describe mixtures of elements of X abstractly because they can be interpreted as either lotteries over C or as frequencies. Our framework and main results will shed some light on how the two might be connected.

[3] In a recent paper, Gilboa et al. [15] argue for a separation between objective and subjective parts of a preference. Roughly, the objective part in their model is an incomplete preference that satisfies Bewley's axioms. Their approach is quite different from ours, where parameters are ergodic distributions that can be characterized in terms of objective empirical frequencies.

An *act* is any measurable function:

$$f : \Omega \to X.$$

An act that takes the constant value x is, with some abuse of notation, denoted x. Let $\mathcal{F}$ be the set of all $\succcurlyeq$-bounded acts; that is, for each $f \in \mathcal{F}$, there exists $x, y \in X$ such that $x \succcurlyeq f(\omega) \succcurlyeq y$ for all $\omega \in \Omega$. The decision maker's choice behavior is represented by a preference relation $\succcurlyeq$ on $\mathcal{F}$. We assume that $\succcurlyeq$ satisfies the following conditions.

Assumption 1 *(Order properties).* $\succcurlyeq$ is reflexive and transitive on $\mathcal{F}$ and complete on X.

Next we introduce the usual monotonicity assumption:

Assumption 2 *(Monotonicity).* If $f(\omega) \succcurlyeq g(\omega)$ for all $\omega \in \Omega$, then $f \succcurlyeq g$.

Write $f^n \to f$ if f^n converges to f pointwise. The following pointwise continuity condition is equivalent to countable additivity of the prior under subjective expected utility.

Assumption 3 *(Continuity).* Given a pair of acts $f, g \in \mathcal{F}$, if there are sequences $\{f^n\}$, $\{g^n\}$ and acts $a, b \in \mathcal{F}$ such that: (i) $f^n \to f$ and $g^n \to g$; (ii) $a(\omega) \succcurlyeq f^n(\omega), g^n(\omega), f(\omega), g(\omega) \succcurlyeq b(\omega), \forall\omega \in \Omega, n \in \mathcal{N}$; and (iii) $f^n \succcurlyeq g^n$ for all $n \in \mathcal{N}$, then $f \succcurlyeq g$.[4]

Next we assume Herstein and Milnor's [17] linearity in mixtures:

Assumption 4 *(Mixture linearity).* Let $x, y, z \in X$. Then $x \sim y \Rightarrow \frac{1}{2}x + \frac{1}{2}z \sim \frac{1}{2}y + \frac{1}{2}z$.

By Herstein and Milnor's [17] theorem (see details in Appendix A), there is an affine function $u : X \to R$, unique up to positive affine transformations, such that for any pair of constant acts f, g, that take values $\ell_f, \ell_g \in X$ respectively,

$$f \succcurlyeq g \quad \Longleftrightarrow \quad \int\limits_{c \in C} u(c)\, d\ell_f \geqslant \int\limits_{c \in C} u(c)\, d\ell_g.$$

We will sometimes use the convenient notation $u(\ell), \ell \in X$ to denote the expected utility $\int_{c \in C} u(c)\, d\ell$. Finally, we assume that the preference is not trivial.

Assumption 5 *(Non-triviality).* There are $x, y \in X$ such that $x \succ y$.

2.2. Transformations and ergodicity

Next we introduce standard mathematical notions of transformations and parametrizations. A *transformation* is any measurable function $\tau : \Omega \to \Omega$. Thus, starting with a state ω, τ generates a sequence of states $\omega, \tau\omega, \tau^2\omega, \ldots$. We will also deal with (countable) semi-groups of

[4] Our continuity assumption is similar to Ghirardato et al.'s [14] B3. They require that, if $f^n \to f$ and $g^n \to g$ pointwise and $f^n \succcurlyeq g^n$ for each n, then $f \succcurlyeq g$. Note that they do not require the sequences to be bounded by a function b.

transformations Γ.[5] It is useful to recall the following standard definitions (below, γ will denote a generic element of the semi-group Γ):

- An event E is γ-invariant if $E = \gamma^{-1}(E)$; E is Γ-invariant if it is γ-invariant for every $\gamma \in \Gamma$.
- $\{\mathbf{1}, \tau, \tau^2, \ldots\}$ is the semi-group of transformations generated by τ, where $\mathbf{1}$ is the identity transformation.
- $\mathcal{E}_\gamma \subset \Sigma$ is the σ-algebra generated by the γ-invariant events; $\mathcal{E}_\Gamma$ is the σ-algebra $\bigcap_{\gamma \in \Gamma} \mathcal{E}_\gamma$.
- A probability measure P is γ-invariant if $P(\gamma^{-1}(E)) = P(E)$ for every $E \in \Sigma$; P is Γ-invariant if it is γ-invariant for every $\gamma \in \Gamma$.
- P is γ-ergodic if either $P(E) = 0$ or $P(E) = 1$ for every γ-invariant event E; P is Γ-ergodic if it is γ-ergodic for every $\gamma \in \Gamma$.

It is well-known that the set of Γ-invariant measures is convex and its extreme points are the Γ-ergodic measures.

2.3. *Parametrizations*

We will be interested in representing preferences in terms of parameters. Fix a countable semi-group of transformations Γ, and write the set of Γ-ergodic measures as $\{P^\theta\}_{\theta \in \Theta}$, with Θ denoting an index set of parameters. Viewed as a set of probability measures, Θ inherits the relativized topology and σ-algebra of $\Delta(\Omega)$.[6] A standard definition of decomposition map with respect to a semi-group is given by Varadarajan [26].[7,8]

Definition 1 *(Decomposition maps and parametrizations).* Fix a σ-algebra $\mathcal{E} \subset \Sigma$, a set of probability measures $\mathcal{P}$, and a subset $\{P^\theta\}_{\theta \in \Theta} \subset \mathcal{P}$ with index set Θ. A function $\vartheta : \Omega \to \Theta$ is a *decomposition map* (with respect to $\mathcal{E}$, $\mathcal{P}$, $\{P^\theta\}_{\theta \in \Theta}$) if

(i) ϑ is measurable;
(ii) $P^\theta(\vartheta^{-1}(\theta)) = 1$ for all $\theta \in \Theta$; and
(iii) for every $A \in \Sigma$, $P^{\vartheta(\omega)}(A)$ is a version of the conditional probability of A given $\mathcal{E}$ for every $P \in \mathcal{P}$.[9]

Refer to (Θ, ϑ) as a *parametrization* and Θ as the set of *parameters*.

If Γ is a semi-group of transformations then we refer to (Θ, ϑ) as the Γ*-parametrization* if $\mathcal{E} = \mathcal{E}_\Gamma$, $\mathcal{P}$ is the set of Γ-invariant probability measures, $\{P^\theta\}_{\theta \in \Theta}$ is the set of Γ-ergodic

5 A semi-group is a set Γ together with an operation "$\cdot$" satisfying closure—$\forall \gamma, \zeta \in \Gamma$, $\gamma \cdot \zeta \in \Gamma$—and associativity—$\forall \gamma, \zeta, \varrho \in \Gamma$, $(\gamma \cdot \zeta) \cdot \varrho = \gamma \cdot (\zeta \cdot \varrho)$. In our case, the operation considered is function composition.

6 We will always assume that there is at least one Γ-ergodic measure.

7 Varadarajan [26] focuses on groups instead of semi-groups, but we were able to extend his main constructions to our case. It is necessary to work with semi-groups to consider the shift transformation applied to sequences that are one-sided infinite. The shift generates a group when applied to sequences that are doubly infinite.

8 The definition is standard and essentially that of a sufficient statistic. See, for example, Billingsley [2, p. 450], Varadarajan [26] and Dynkin [11].

9 The key point is that the conditional distribution $P^{\vartheta(\omega)}$ does not depend on P.

measures, and ϑ is Γ-invariant (i.e., $\vartheta(\gamma(\omega)) = \vartheta(\omega), \forall\gamma \in \Gamma$). If Γ is the semi-group generated by τ, then we abuse terminology and refer to (Θ, ϑ) as the τ-parametrization.

The ergodic decomposition theorem (see, e.g., Varadarajan [26]) shows, under general conditions, that a decomposition map exists.[10] Note that such decomposition is a purely mathematical object that may bear little or no connection to choice behavior. The next section develops such connection.

3. Invariance and sufficient statistics

The central concept in this paper is invariance to transformations of the state space. Invariance is a central, foundational concept in statistical inference and, as we show later, in connecting the notions of risk and uncertainty.

3.1. Invariance

Intuitively, a transformation τ is a rearrangement of the state space, and invariance refers to the property that the preference remains the same after the states have been thusly rearranged. At a minimum, invariance with respect to a single transformation τ should require that for any act f,

$$f \sim f \circ \tau.$$

For a concrete example, suppose that Ω has a *product structure*, i.e., $\Omega = S \times S \times \cdots$ with each coordinate S interpreted as modeling the random outcome of an experiment of interest (a coin toss, an econometric model, and so on). Write a generic state ω in terms of the infinite sequence of coordinate values $(s^1, s^2, \ldots)$. Consider the permutation transformation:

$$\left(s^1, s^2, \ldots\right) \xrightarrow{\pi} \left(s^2, s^1, \ldots\right).$$

Invariance with respect to this permutation formalizes the intuition that the decision maker views the first and second experiments as similar. More generally, invariance with respect to the set of finite permutations indicates that the decision maker is indifferent to relabellings of the coordinates, and leads to the concept of exchangeability. Another example is the *shift transformation*:

$$\left(s^1, s^2, \ldots\right) \xrightarrow{T} \left(s^2, s^3, \ldots\right).$$

Invariance with respect to this transformation corresponds to a decision maker with stationary preferences.

We will be interested in invariance with respect to *sets* of transformations. At a minimum, starting with a transformation τ we would like to consider its iterates $\tau^2, \tau^3, \ldots$. If we are to incorporate a set of transformations Γ in our model, it seems natural to require that Γ be closed

[10] It is also essentially unique, in the sense of Lemma 4.4 in Varadarajan [26]. We comment further on this below. We cannot use directly Varadarajan's result because we work with semi-groups instead of groups and, more importantly, we do not assume—as we did in a previous version of this paper—that a set which is μ-null for all invariant measures μ is also $\succcurlyeq$-null. Without this property (previously called Bayesian consensus), it is not clear what is the meaning of Varadarajan's map for the preference. In Appendix B.2, the Bayesian consensus is *proved* as a consequence of the existence of ϑ, which is directly established.

under composition: given two transformations $\gamma_1, \gamma_2 \in \Gamma$ their composition $\gamma_1 \circ \gamma_2$ should also belong to Γ. For example, we need the process of shifting by two coordinates $T^2 \equiv T \circ T$ to also be a legitimate transformation (i.e., belongs to Γ). This amounts to saying that Γ is a semi-group of transformations.[11] We note finally that we do not require transformations γ to have an inverse (which is why we work with semi-groups rather than groups). For example, the shift T is not invertible, but our results apply to the semi-group obtained by T and its iterates $\{T, T^2, \ldots\}$.

Definition 2 *(Invariance).* Let Γ be a countable semi-group of transformations. The preference $\succcurlyeq$ is Γ*-invariant* if for all acts $f \in \mathcal{F}$, integer n, and $\gamma_1, \ldots, \gamma_n \in \Gamma$,

$$f \sim \frac{f \circ \gamma_1 + \cdots + f \circ \gamma_n}{n}. \tag{3}$$

If Γ is the semi-group generated by τ, then we abuse terminology and call $\succcurlyeq$ τ-invariant.

Note that without the linear structure on the space of consequences (for instance, if consequences were just a finite set C), the averages in (3) would not make sense. The linear structure ensures that these conditions are behaviorally meaningful. Thus, letting u be an affine utility function on consequences, we have for every ω:

$$u\left(\frac{f \circ \gamma_1 + \cdots + f \circ \gamma_n}{n}(\omega)\right) = \frac{u \circ f \circ \gamma_1(\omega) + \cdots + u \circ f \circ \gamma_n(\omega)}{n}.$$

Note that this last condition incorporates the decision maker's risk attitude, expressed in u, while (3) is free from such reference. We think of invariance as part the decision maker's understanding of similarity in the problem he faces, and thus should not be confounded with his attitude towards risk.

The invariance condition is interesting only when we consider 'coarse' parametrizations with respect to which the preference is invariant. To make this formal, note first that if $\Gamma \subset \Gamma'$ then Γ'-invariance implies Γ-invariance. Note further than every preference is invariant with respect to the trivial semi-group $\{\mathbf{1}\}$ that consists of the identity transformation $\mathbf{1}$, defined by $\mathbf{1}(\omega) = \omega$.[12] Invariance has more of a bite when we consider rich sets of transformations with intuitive structures. See Section 3.3 for discussion and examples.

3.2. Subjective ergodic theory and sufficient statistics

In the remainder of this section, we restrict attention to semi-groups generated by a single transformation τ. In Section 3.3 we show that they can be the basis for a general theory to model invariance relative to general classes of transformations, e.g., the group of finite permutations that give rise to exchangeability.

Next we introduce the concept of sufficient parametrizations:

[11] The other axiom of semi-groups, associativity, is automatically satisfied for the composition of functions $\circ$.

[12] Because in this case, (3) reduces to: $f \sim \frac{f \circ \mathbf{1} + \cdots + f \circ \mathbf{1}}{n} = f$, which is guaranteed by reflexivity.

Definition 3 *(Sufficiency).* A parametrization (Θ, ϑ) is *sufficient* for a preference $\succcurlyeq$ if ϑ is the essentially unique function satisfying[13]:

$$\forall f, g \in \mathcal{F}, \quad f \succcurlyeq g \quad \Longleftrightarrow \quad \int_\Omega f\, dP^{\vartheta(\cdot)} \succcurlyeq \int_\Omega g\, dP^{\vartheta(\cdot)}. \tag{4}$$

A parametrization (Θ, ϑ) is sufficient for $\succcurlyeq$ if in ranking f and g, it is enough for the decision maker to examine the acts $\int_\Omega f\, dP^{\vartheta(\cdot)}$ and $\int_\Omega g\, dP^{\vartheta(\cdot)}$ that aggregate, slice by slice, the acts f and g using the parameters. In words, the integrals with respect to the parameters (the RHS of (4)) are sufficient summary of how $\succcurlyeq$ ranks all acts. The notion of parametric preference has bite only when there is a non-trivial parametrization.

The above definition of sufficiency for preferences is closely related to the standard concept of sufficiency in mathematical statistics. Recall that a measurable function $\kappa : \Omega \to A$, where A is an abstract measurable space, is a *sufficient statistic* for a family of probability distributions $\mathcal{P}$ if the conditional distributions $P(\cdot \mid \kappa)$ do not depend on $P \in \mathcal{P}$. Roughly, κ is sufficient if it captures all the relevant information contained in a state ω: given knowledge that $\kappa(\omega) = \bar{\kappa}$, no further information about ω is useful in drawing an inference about P. By analogy, ϑ is a sufficient statistic for the family of all τ-invariant preferences (that satisfy our other conditions).

Every transformation τ gives rise to *empirical limits* of an act:

$$f^\star(\omega) \equiv \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f\left(\tau^j \omega\right),$$

a concept that connects subjective probability and frequencies. The following theorem adapts the findings in De Castro and Al-Najjar [7] to a setting with general consequences:

Theorem 1. *Given a transformation τ, there is a (τ-) parametrization (Θ, ϑ) such that for every τ-invariant preference $\succcurlyeq$ satisfying assumptions* 1–5:

1. (Θ, ϑ) *is sufficient for* $\succcurlyeq$.
2. *For every act f, $f \sim f^\star$ and for all ω outside a $\succcurlyeq$-null set Ω',*[14] *$f^\star$ exists and*

$$f^\star(\omega) = \int_\Omega f\, dP^{\vartheta(\omega)}. \tag{5}$$

3.3. Parametric preferences

The central concept of this paper is that of a parametric preference:

Definition 4 *(Parametric preference).* We say that $\succcurlyeq$ is a *parametric preference* (*with parameters* Θ) if it has a sufficient parametrization (Θ, ϑ).

[13] That is, if ϑ' is another function satisfying (4), the set $\{\omega \in \Omega\colon\ \vartheta(\omega) \neq \vartheta'(\omega)\}$ is $\succcurlyeq$-null.

[14] We use the standard definition of null events: $E \subset \Omega$ is $\succcurlyeq$*-null* if for all acts f, g, h:

$$\begin{bmatrix} f(\omega), & \text{if } \omega \in E \\ h(\omega), & \text{if } \omega \notin E \end{bmatrix} \sim \begin{bmatrix} g(\omega), & \text{if } \omega \in E \\ h(\omega), & \text{if } \omega \notin E \end{bmatrix}.$$

In general, we are interested in more parsimonious (more restrictive) parametrization than what is delivered by a single transformation τ. For example, if Ω has product structure and the transformation τ is just the shift T, then Θ is the set of stationary ergodic measures, which includes the set of i.i.d. distributions, but also all stationary non-i.i.d. distributions (e.g., all k-stage Markov processes). If $\succcurlyeq$ is in addition invariant to all finite permutations (exchangeable, as we define below), then $\succcurlyeq$ continues to be T-invariant and (Θ, ϑ) remains a parametrization for $\succcurlyeq$. But this parametrization does not take into account the additional restriction that $\succcurlyeq$ is also permutation-invariant.

In situations where there is a specific semi-group Γ corresponds to invariance properties the agent perceives in his environment (e.g., the exchangeable or Markovian cases discussed below), we use the following result which is adapted from De Castro and Al-Najjar [7] to a context with utility:

Theorem 2. *Let (Θ, ϑ) be a sufficient parametrization of $\succcurlyeq$, Γ any countable semi-group, and Θ_Γ the corresponding set of Γ-ergodic probabilities. Assume that $u(X) = \mathcal{R}$, $\Theta_\Gamma \subset \Theta$, and that ϑ is Γ-invariant. Then the following are equivalent*:

1. *$f \sim f \circ \gamma$ for all $\gamma \in \Gamma$ and $f \in \mathcal{F}$.*
2. *There exists a map $\tilde{\vartheta} : \Omega \to \Theta_\Gamma$ such that $(\Theta_\Gamma, \tilde{\vartheta})$ is a sufficient parametrization of $\succcurlyeq$.*
3. *$\vartheta^{-1}(\Theta \setminus \Theta_\Gamma)$ is $\succcurlyeq$-null.*

Condition 3 in the theorem captures the intuition of a *parametric restriction*: parameters outside Θ_Γ, although part of the parametrization implied by τ-invariance, are irrelevant for the preference.

3.3.1. Exchangeability

The classic example of invariance and its implications is de Finetti's [8] notion of exchangeability and his representation theorem. To introduce exchangeability, assume that the state space has the product structure $\Omega = S \times S \times \cdots$. Let Π be the group of all finite permutations, with a generic permutation denoted π. The following result is an immediate corollary of Theorem 2.

Corollary 3. *Assume that $\succcurlyeq$ is T-invariant and that (Θ, ϑ) is a sufficient T-parametrization. Let $\Theta_\Pi \subset \Theta$ be the set of i.i.d. measures. Then the following are equivalent*:

1. *$f \sim f \circ \pi$ for every $\pi \in \Pi$ and $f \in \mathcal{F}$.*
2. *There exists a map $\tilde{\vartheta} : \Omega \to \Theta_\Pi$ such that $(\Theta_\Pi, \tilde{\vartheta})$ is a sufficient parametrization of $\succcurlyeq$.*

Note that, as in Theorems 2, we do not require full-invariance with respect to the semi-group in the sense of (3).

The exchangeable case has attracted the most interest in statistics and in decision theory. There is also a large literature that studies weaker notions of exchangeability, usually referred to as "partial exchangeability."[15] See Section 5.3 for review of related literature in decision theory.

[15] See Diaconis and Freedman [10] for a general treatment.

3.3.2. Markov parameters

In many applications, it is natural to consider parameters with a non-trivial correlation structure. This simplest such case is that of Markov processes, which includes i.i.d. distributions as special case. Diaconis and Freedman [9] characterize the group of transformations $\mathcal{M}$ that gives rise to Markov parameters. Using their characterization, we can use Theorem 2 to obtain an analogue of Corollary 3 for the Markovian case.

3.3.3. The trivial parametrization

The 'finest' parametrization (i.e., the one with the most parameters) is the one where each state ω is its own parameter, the Dirac measure δ_ω that puts unit mass on that state. This parametrization represents the polar case where the decision maker makes no connections between states. This intuition is confirmed by the next proposition showing that the trivial parametrization corresponds to the semi-group $\{\mathbf{1}\}$, where $\mathbf{1}$ is the identity transformation: $\mathbf{1}(\omega) = \omega$. For the next proposition, assume an abstract Ω that does not necessarily have a product structure.

Proposition 4. *The (essentially unique) parametrization (Θ, ϑ) corresponding to the trivial semi-group $\{\mathbf{1}\}$ is one where:*

- *the parameters are the Dirac measures δ_ω, $\omega \in \Omega$;*
- *ϑ is the identity.*

Let I_A denote the indicator function of an event A, so $I_A^\star(\omega)$ is the empirical frequency under τ of the event A at state ω.

Proof of Proposition 4. For the trivial group $\mathbf{1}$, the limit $f^\star$ trivially exists for all $\omega \in \Omega$. In particular, $I_A^\star(\omega) = I_A(\omega) = \delta_\omega(A)$, for all $A \in \Sigma$. Therefore, we can define the decomposition map $\vartheta(\omega) = \delta_\omega$ and $\Theta = \Omega$ will be the parameter space. All the properties of the decomposition map are easily seen to be satisfied. □

Every preference is trivially $\mathbf{1}$-invariant, imposing no restrictions on the preference. Interpreting invariance as a belief in a similarity relationship between states, under the trivial transformation every state is similar only to itself. In the case of coin tosses, under the trivial parametrization, no finite amount of data will enable the decision maker to learn the true parameter. Compare this with exchangeable parametrizations where it is easy to devise (classical or Bayesian) procedures that 'learn' the true i.i.d. parameter.

3.3.4. Discussion

Next we turn to some issues of interpretation:

- *Parameters and frequencies*: The probability measure $P^{\vartheta(\omega)}$ can be constructed by observing the (deterministic) sequence of $\tau(\omega), \tau^2(\omega), \ldots$, for all ω in a set Ω' that is the complement of a $\succcurlyeq$-null set. In Appendix A, Theorem A.9 shows that information about the frequencies starting with a typical initial state ω is sufficient to derive the parameter $\vartheta(\omega)$. The distribution $P^{\vartheta(\omega)}$ is therefore nothing but a compact way to represent the empirical frequencies at ω.
- *Parameters and objectivity*: Parameters P^θ may be interpreted as objective in that they can be constructed from the long-run frequencies. On the other hand, how the decision maker

treats parameter uncertainty is, in this sense, subjective in that it cannot be linked to objective frequencies.

- *Taste over consequences vs. invariance judgment*: The preference $\succcurlyeq$ distills not just the decision maker's judgment of similarity but also, among other things, his ranking of consequences and his risk attitude. The decision maker's similarity judgment is conceptually distinct from such taste issues. For example, when facing a sequence of statistical experiments, the decision maker judgment whether the experiments are, say, exchangeable is an assertion of a statistical connection between experiments that ought to be unrelated to his taste over consequences. The order of quantifiers in Theorem 1 is important: the same parametrization (Θ, ϑ) works simultaneously for all Γ-invariant preferences, regardless of the decision maker's utility function over consequences or his risk attitude.

4. Aggregating parameter uncertainty

We introduce the concepts of parameter-based acts and aggregators to provide a general methodology for modeling attitudes towards parameter uncertainty. This section aims at a general treatment of aggregators, regardless of whether they have known or interesting axiomatic characterizations. In Section 5 we use our approach to characterize second-order expected utility aggregators based on primitive preference properties.

4.1. Parameter-based acts

A *parameter-based act* is any measurable function

$$F : \Theta \to X.$$

Contrast this with (ordinary) *state-based acts* which are defined on Ω. Let $\mathbb{F}$ denote the set of parameter-based acts. As a notational convention, we denote state-based acts with lower case letters f, g, h and parameter-based acts by the upper case letters F, G, H.

It is usually more convenient to introduce assumptions regarding how the decision maker treats parameter-uncertainty directly on acts defined in terms of parameters. To avoid ambiguity, we refer to a binary relation $\succcurlyeq$ on $\mathbb{F}$ as an *aggregator* because it describes how the decision maker aggregates uncertainty about parameters. Since an aggregator is just a preference on an auxiliary state space Θ, the properties of reflexivity, transitivity, monotonicity, and continuity can be defined similarly to the corresponding properties of preferences.

This section's objective is to provide a template for how to start with properties of the aggregator $\succcurlyeq$ and translate them into properties of a primitive preference on $\mathcal{F}$. Our main tool is the operator $\Psi : \mathcal{F} \to \mathbb{F}$:

$$\Psi(f)(\theta) = \int_{\Omega} f \, dP^{\theta},$$

which relates state-based and parameter-based acts. The following proposition establishes its usefulness in linking aggregators to preferences: starting with an abstract decomposition map and an aggregator, we can construct a preference on the underlying state space.

Proposition 5. *Let (Θ, ϑ) be a decomposition map. Then for any aggregator $\succcurlyeq$ on $\mathbb{F}$ satisfying assumptions 1–5, there is a uniquely defined preference $\succcurlyeq$ on $\mathcal{F}$ satisfying the same assumptions such that (Θ, ϑ) is a sufficient parametrization for $\succcurlyeq$, that is*:

$$f \succcurlyeq g \quad \Longleftrightarrow \quad \Psi(f) \succcurlyeq \Psi(g). \tag{6}$$

Conversely, given a preference $\succcurlyeq$ satisfying assumptions 1–5 and its sufficient parametrization (Θ, ϑ), there is an aggregator $\succcurlyeq$ on $\mathbb{F}$ satisfying the same assumptions and (6).

Next, consider a situation where we only know that there is a functional $\mathcal{V} : \mathbb{F} \to \mathcal{R}$ which represents the aggregator $\succcurlyeq$:

$$F \succcurlyeq G \quad \Longleftrightarrow \quad \mathcal{V}(F) \geqslant \mathcal{V}(G).$$

Proposition 6. *Let (Θ, ϑ), $\succcurlyeq$ and $\succcurlyeq$ satisfy all the conditions of Proposition 5. Then, there is a functional $\mathcal{V}$ representing $\succcurlyeq$ if and only if there are functions $V : \tilde{\mathbb{F}} \to \mathcal{R}$ and $u : X \to \mathcal{R}$, where $\tilde{\mathbb{F}} \subset \mathcal{R}^{\Theta}$ and u is affine, such that*

$$f \succcurlyeq g \quad \Longleftrightarrow \quad V\left(\theta \mapsto \int_{\Omega} u(f)\, dP^{\theta}\right) \geqslant V\left(\theta \mapsto \int_{\Omega} u(g)\, dP^{\theta}\right). \tag{7}$$

Proof. By Proposition 5 and the assumption, $f \succcurlyeq g \Leftrightarrow \Psi(f) \succcurlyeq \Psi(g) \Leftrightarrow \mathcal{V}(\Psi(f)) \succcurlyeq \mathcal{V}(\Psi(g))$. By monotonicity, if $F(\theta) \sim G(\theta)$ for all $\theta \in \Theta$, then $F \sim G$. Since $\succcurlyeq$ and $\succcurlyeq$ are complete in X and satisfy the Herstein–Milnor axiom, there is an affine function $u : X \to \mathcal{R}$ representing the preference (and the aggregator) over X. Therefore, we can write the functional $\mathcal{V}$ as $V \circ u$, where $V : \tilde{\mathbb{F}} \to \mathcal{R}$ and $\tilde{\mathbb{F}} = u(\mathbb{F}) \subset \mathcal{R}^{\Theta}$. Since u is affine, $u(\Psi(f)(\theta)) = \int_{\Omega} u(f)\, dP^{\theta}$, which establishes (7). The converse is trivial. □

4.2. Uncertainty averse aggregators

For examples of aggregators, consider the class of uncertainty averse preferences characterized by Cerreia, Maccheroni, Marinacci, and Montrucchio [3] (henceforth CMMM). This class is very broad and covers most models of uncertainty aversion in the literature, applied to the set $\tilde{\mathbb{F}} \subset \mathbb{F}$ of parameter valued acts assuming finitely many values. Here the aggregator characterized is

$$F \succcurlyeq G \quad \Longleftrightarrow \quad \min_{\mu \in \Delta^{\sigma}(\bar{\mu})} \Phi\left(\int_{\Theta} u(F)\, d\mu, \mu\right) \geqslant \min_{\mu \in \Delta^{\sigma}(\bar{\mu})} \Phi\left(\int_{\Theta} u(G)\, d\mu, \mu\right) \tag{8}$$

where: $\bar{\mu} \in \Delta(\Theta)$, $\Delta^{\sigma}(\bar{\mu})$ is the set of countably additive probability measures which are absolutely continuous with respect to $\bar{\mu}$, $u : X \to \mathcal{R}$ is an affine function with $u(X) = \mathcal{R}$, and $\Phi : \mathcal{R} \times \Delta(\Theta) \to (-\infty, \infty]$ is a function satisfying certain technical conditions; see CMMM for details.

If (Θ, ϑ), $\succcurlyeq$ and $\succcurlyeq$ satisfy all the conditions of Proposition 5 and $\succcurlyeq$ satisfies the axioms A.1–A.8 of CMMM on $\tilde{\mathbb{F}}$, one can conclude by CMMM's Theorem 7 that there exist an affine $u : X \to \mathcal{R}$, with $u(X) = \mathcal{R}$, a function[16] $\Phi : \mathcal{R} \times \Delta(\Theta) \to (-\infty, \infty]$ and $\bar{\mu} \in \Delta(\Theta)$ such that, for all $f, g \in \tilde{\mathcal{F}}$,[17]

$$f \succcurlyeq g \iff \min_{\mu \in \Delta^\sigma(\bar{\mu})} \Phi\left(\int_\Theta \left(\int_\Omega u(f)\, dP^\theta\right) d\mu, \mu\right) \geqslant \min_{\mu \in \Delta^\sigma(\bar{\mu})} \Phi\left(\int_\Theta \left(\int_\Omega u(g)\, dP^\theta\right) d\mu, \mu\right).$$

The converse is also true.

One could substitute CMMM's axioms by some other set of axioms Ax provided that an aggregator $\succcurlyeq$ satisfies Ax if and only if there exist functions $A : \tilde{\mathbb{F}} \to \mathcal{R}$ and affine $u : X \to \mathcal{R}$, $\tilde{\mathbb{F}} = u(\mathbb{F}) \subset \mathcal{R}^\Theta$, such that

$$F \succcurlyeq G \iff A\big(u\big(F(\cdot)\big)\big) \geqslant A\big(u\big(G(\cdot)\big)\big).$$

Then an analogous representation would hold with the obvious adaptations.

5. Second-order expected utility

In this section we examine in greater details second-order expected utility aggregators, which generalize the model introduced by Neilson [23,24].[18] The theorem of this section, in contrast to the results of the last section, is stated in terms of primitive assumptions on preferences, rather than abstract aggregators.

5.1. Expected utility aggregators

Given a transformation τ, recall that $\mathcal{E}_\tau$ is the set of events that are τ-invariant, and that it is a sub-σ-algebra of Σ. Define $\mathcal{F}_\tau$ to be the subset of acts that are measurable with respect to $\mathcal{E}_\tau$.

Definition 5. Given a transformation τ, a preference $\succcurlyeq$ has an *expected utility representation on* $\mathcal{F}_\tau$ if

- There exists a function $\varphi : X \to \mathcal{R}$ and a countably additive probability measure ν on $(\Omega, \mathcal{E}_\tau)$ such that for any $f, g \in \mathcal{F}_\tau$

$$f \succcurlyeq g \iff \int_\Omega \varphi(f)\, d\nu \geqslant \int_\Omega \varphi(g)\, d\nu \tag{9}$$

  and:
- The function φ is unique up to positive affine transformations, and the measure ν is unique.

Behavioral axioms characterizing this property are standard: they reduce to applying the Savage model a decision problem where events are restricted to $\mathcal{E}_\tau$.

[16] Φ satisfy some technical conditions. See [3] for details.

[17] $\tilde{\mathcal{F}}$ denotes the set of finitely valued acts.

[18] As noted in the Introduction, related models include, among others, Nau [21,22], Klibanoff et al. [18,19], Ergin and Gul [13], Chew and Sagi [5], Strzalecki [25], and Grant et al. [16].

Theorem 7. *Suppose that (Θ, ϑ) is sufficient for $\succcurlyeq$. The following statements are equivalent:*

1. *$\succcurlyeq$ has a subjective expected utility representation on $\mathcal{F}_\tau$.*
2. *There is a probability measure μ on Θ, and a function $\phi : \mathcal{R} \to \mathcal{R}$ such that, for any pair of acts f, g:*

$$f \succcurlyeq g \iff \int_\Theta \phi\left(\int_\Omega u(f)\, dP^\theta\right) d\mu \geqslant \int_\Theta \phi\left(\int_\Omega u(g)\, dP^\theta\right) d\mu. \tag{10}$$

If a preference can be represented as in (10), then μ is unique, and the restriction of the function ϕ to $u(X)$ is unique up to positive affine transformations.

Moreover, if (Θ, ϑ) is a Γ-parametrization, then $\succcurlyeq$ is Γ-invariant.

This theorem is closely related to Theorem 6 of Cerreia et al. [4]. We offer different foundations, however. In our model, parameters are ergodic distributions that emerge out of invariance and similarity, and with a well-established connections with objective empirical frequencies.

Proof of Theorem 7. Assume (1) in the statement of the theorem and let (φ, μ) be as in the definition so (9) holds. Since φ and u both represent the same preference on X, there must be a monotone increasing function $\phi : u(X) \to \mathcal{R}$ such that $\varphi(c) = \phi(u(c))$ for every consequence c. Since ν is defined on $(\Omega, \mathcal{E}_\tau)$, it is associated to the measure $\mu = \nu \circ \vartheta^{-1}$ on Θ.

Fix f, g and let F, G be the corresponding parameter-based acts, that is, $F(\theta) = \int_\Omega f\, dP^\theta$ and similarly for G. Then:

$$f \succcurlyeq g \iff \int_\Omega f\, dP^{\vartheta(\cdot)} \succcurlyeq \int_\Omega g\, dP^{\vartheta(\cdot)} \tag{11}$$

$$\iff \int_\Omega \varphi\left(\int_\Omega f\, dP^{\vartheta(\cdot)}\right) d\nu \geqslant \int_\Omega \varphi\left(\int_\Omega g\, dP^{\vartheta(\cdot)}\right) d\nu \tag{12}$$

$$\iff \int_\Theta \varphi\left(\int_\Omega f\, dP^\theta\right) d\mu \geqslant \int_\Theta \varphi\left(\int_\Omega g\, dP^\theta\right) d\mu \tag{13}$$

$$\iff \int_\Theta \phi \circ u\left(\int_\Omega f\, dP^\theta\right) d\mu \geqslant \int_\Theta \phi \circ u\left(\int_\Omega g\, dP^\theta\right) d\mu \tag{14}$$

$$\iff \int_\Theta \phi\left(\int_\Omega u(f)\, dP^\theta\right) d\mu \geqslant \int_\Theta \phi\left(\int_\Omega u(g)\, dP^\theta\right) d\mu. \tag{15}$$

In the above: (11) follows from the definition of a sufficient parametrization; (12) follows from condition (9) and the fact that the acts in (11) are in $\mathcal{F}_\tau$; (13) follows from the definition of μ; (14) follows from $\varphi(c) = \phi(u(c))$ for every c; and finally (15) follows from the fact that u is linear. Note that ϕ can be moved along with u inside the integral only if it is linear.

Conversely, if we assume that (2) in the theorem holds, we can redo the above equivalences and obtain (12), which implies that $\succcurlyeq$ has a subjective expected utility representation on $\mathcal{F}_\tau$.

Finally, we assume (10) and show that $\succcurlyeq$ is Γ-invariant. Fix $\gamma_1, \ldots, \gamma_n \in \Gamma$ and act f. Then:

$$\begin{aligned}\int \varphi\left(\int \frac{1}{n}\sum_{j=0}^{n-1} f \circ \gamma_j \, dP^\theta\right) d\mu &= \int \varphi\left(\frac{1}{n}\sum_{j=0}^{n-1}\int f \circ \gamma_j \, dP^\theta\right) d\mu \\ &= \int \varphi\left(\frac{1}{n}\sum_{j=0}^{n-1}\int f \, dP^\theta\right) d\mu \\ &= \int \varphi\left(\int f \, dP^\theta\right) d\mu.\end{aligned}$$

This concludes the proof. □

The reason Theorem 7 is not covered by Propositions 6 is that in these propositions, the aggregator is already represented by an affine u. Here, φ need not be affine; in fact, if were, then the model above collapses to a standard expected utility preference.

Parametrization partitions the state space into events $\{\vartheta^{-1}(\theta)\}_\Theta$ within which variability is treated as objective risk, in the sense that the decision maker applies the same risk attitude given by u that he applies to objective lotteries.

For further intuition, consider the state space $\Omega = \{H, T\}^\infty$ and two decision makers with preferences $\succcurlyeq$ and $\succcurlyeq'$ with the same utility functions u so they display identical attitudes towards objective risk. Suppose that $\succcurlyeq$ is invariant only with respect to the trivial semi-group $\{\mathbf{1}\}$, while $\succcurlyeq'$ is exchangeable. Then the sets of parameters are $\Theta = \Omega$ and $\Theta' = [0, 1]$ respectively. Define the second-order probabilities μ and μ' to be the uniform distributions on Θ and Θ' respectively. Finally, assume that ϕ is strictly concave and identical for both preferences. Consider the act f that pays 1 dollar if the first toss is H and 0 otherwise, and let θ be the probability of H. Then the parameter-based act corresponding to f yields utility $\delta_\omega \mapsto u(f(\omega))$ for the first decision maker and $\theta \mapsto \theta u(1) + (1-\theta)u(0)$ for the second. The overall value of the act under the representation (10) is, respectively, $0.5\phi(u(1)) + 0.5\phi(u(0))$ and $\int \phi[\theta u(1) + (1-\theta)u(0)]\, d\mu$. The decision maker with finer parametrization perceives less risk and more uncertainty than the decision maker with the coarser parametrization.

5.2. Separating objective and subjective uncertainties

The double integral representation in Theorem 7 includes as a special case expected utility models as well as some of the ambiguity aversion preferences studied in the literature. In this subsection we consider the models of Neilson [23,24] and the related work by Strzalecki [25].[19] Neilson considers the representation:

$$V_N(f) = \int \phi\big[u\big(f(\omega)\big)\big]\, d\mu(\omega). \tag{16}$$

To relate this to the framework of Theorem 7 and Proposition 4, we first note that every preference is invariant with respect to the trivial semi-group $\{\mathbf{1}\}$. A representation via the functional V_N obtains if no additional structure on the decision maker's perception of his environment is imposed:

[19] Strzalecki [25] introduces additional structure that ensures that ϕ has the specific functional form corresponding to multiplier preferences. The form of the function ϕ is not a focus of the present paper.

Corollary 8. *For every preference relation* $\succcurlyeq$ *the following are equivalent*:

1. $\succcurlyeq$ *satisfies assumptions* 1–4*, is invariant with respect to the trivial semi-group* $\{\mathbf{1}\}$*, and has an expected utility representation on parameter-based acts.*
2. *There is a probability measure* μ *on* Θ*, and a function* $\phi : \mathcal{R} \to \mathcal{R}$ *such that* $\succcurlyeq$ *can be represented by the functional* V_N.

The uniqueness properties of μ*,* ϕ *hold as in Theorem* 7.

Introducing invariance with respect to a non-trivial semi-group Γ captures the idea that the decision maker treats as objective risk not just the objective lotteries on consequences, but also all uncertainty conditional on knowledge of the value of the parameter. If (Θ, ϑ) denotes the parametrization corresponding to Γ, then our model accommodates the parameters as an additional source of objective uncertainty, yielding the functional form derived in Theorem 7.

5.3. Literature review

Epstein and Seo [12] consider invariance with respect to the group of permutations Π, as well as weaker notions that give rise to parameters that are sets of probabilities. See our earlier working paper, De Castro and Al-Najjar [7], for a detailed discussion of their work. In a paper subsequent to our work, Klibanoff et al. [19] showed equivalent forms of the invariance condition and used them to characterize various ambiguity aversion models.

Cerreia et al. [4] provide a framework which incorporates the statistical concept of sufficiency into decision theoretic models. They study a decision maker with information represented by a set of probability measures $\mathbb{P}$ and show how this can give rise to behavioral characterizations in terms of the set $\mathbb{S}(\mathbb{P})$ of strong extreme points of $\mathbb{P}$. Our framework is different and potentially complementary. We take as primitive the invariance of a preference with respect to transformations, capturing the decision maker's perception of similarity between experiments. We show that these deterministic transformations give rise to probability distributions that are sufficient for the preference, in the sense of (4). These distributions are characterized in terms of frequencies and are related to commonly used parameters in statistics. In Cerreia et al. [4] information is given in the form of probabilities $\mathbb{P}$ which need not have a connection to frequencies or parameters.

Klibanoff et al. [18] provide a model with similar representation which, in our notation, has the form:

$$\int_{\Delta(\Omega)} \phi\left(\int_{\Omega} u\big(f(\omega)\big)\, dP(\omega) \right) d\nu(P). \tag{17}$$

They interpret ν as the decision maker's subjective uncertainty about the 'true' objective process P, and the support of ν as the set of 'true' processes or parameters the decision maker views as possible.

In terms of foundations, Klibanoff et al. [18] postulate two preferences: One preference $\succcurlyeq$ over the set of state-based acts $\mathcal{F}$, and second preference $\breve{\succcurlyeq}$ over "second-order acts," which is the set $\mathbf{F}$ of all functions of the form:

$$\mathbf{f} : \Delta(\Omega) \to X.$$

It is not possible to formally compare this approach with the framework of this paper because neither second-order acts nor integration over $\Delta(\Omega)$ in (17) have a behavioral meaning in our model. Here we briefly highlight the main issues. A detailed discussion appears in our companion note, Al-Najjar and Castro [1]. The second-order expected utility model (10) differs from (17) along two dimensions:

- *Foundations*: The behavioral foundations underlying (17) require decision makers to express a preference $\breve{\succcurlyeq}$ over second-order acts.
- *Functional form*: In (17) the outer integral is over mixtures of parameters so decision makers have beliefs about randomizations over parameters.

To illustrate these issues, consider a repeated coin toss setting with only two possible i.i.d. parameters $\theta_0 \neq \theta_1$. A second-order act restricted to this domain is a function $\mathbf{f} : [0, 1] \to \mathcal{R}$ where $\mathbf{f}(\alpha)$ is the decision maker's payoff when the 'true' distribution is the mixture $P_\alpha \equiv \alpha P^{\theta_0} + (1-\alpha)P^{\theta_1}$.

For concreteness, interpret $\mathbf{f}$ as a contractual promise to pay 0 if $\alpha \leqslant 0.5$ and 100 otherwise. Then $\mathbf{f}(0)$ and $\mathbf{f}(1)$ make sense since the payoff of the decision maker can be determined based on the objective long-run frequency of *Heads*. The second-order act $\mathbf{f}$, on the other hand, requires the decision maker to contemplated what his payoff will be if θ_0 is selected with probability $1/3$, say. It is difficult to think of what meaning to attribute to such contract. For example, a contract that pays 100 if a Democrat wins the next U.S. presidential election and 0 otherwise is meaningful because the payment is contingent on events that can be objectively verified. On the other hand, a contract that pays 100 if a Democrat wins the election with probability $1/3$ or less, and pays 0 otherwise, treats the probability of a Democrat winning the White House as it were an objective entity that can be measured, rather than a subjective state of mind of the decision maker.

In our model, a parameter θ is a label associated with the event $\vartheta^{-1}(\theta) \subset \Omega$. Bets on parameters are bets on such events, and parameter-based acts are act in the usual sense, determining for each state ω an unambiguous consequence $F(\vartheta(\omega))$. Bets on the probability with which a parameter occurred have no similar meaning even in idealized thought experiments with infinite data. All that is observed in the limit is the state ω, from which one can infer which parameter occurred, but not the probability with which it did occur. Preferences that incorporates second-order acts are inconsistent not just with the second-order expected utility model (10) but with all parametric preferences covered in this paper, regardless of functional form.[20]

Appendix A. Preliminary results

A.1. $\succcurlyeq$-null sets

Lemma A.1. *Let E_n be $\succcurlyeq$-null for all $n \in \mathcal{N}$. Then, $E = \bigcup_{n\in\mathcal{N}} E_n$ is $\succcurlyeq$-null.*

[20] A referee raised the issue that parameters in our setting are also unobservable, in the sense that they require an infinite amount of data to verify. For example, in a sequence of coin tosses, verifying whether the event "the limiting average of *Heads* is less than $\frac{1}{2}$" occurred is not possible with finite data. There is nothing special about our model in this regard. Consider, for instance, Savage's setting with state space [0, 1]. To determine which consequence obtains under some act, one may have to verify which state actually occurred, even though this requires checking the infinite decimal expansion of a real number, which may be infeasible in practice. In decision theoretic frameworks, constraints like these are modeled by limiting the feasible sets available to the decision maker, not as part of the abstract framework itself.

Proof. Let f, g, h be arbitrary acts. Define $A^N \equiv \bigcup_{n=1}^N E_n$; $f^N \equiv f1_{A^N} + x1_{E\setminus A^N} + h1_{E^c}$, and $g^N \equiv g1_{A^N} + x1_{E\setminus A^N} + h1_{E^c}$. Observe that $f^N = \sum_{n=1}^N f1_{E_n} + x1_{E\setminus A^N} + h1_{E^c}$ and $g^N = \sum_{n=1}^N g1_{E_n} + x1_{E\setminus A^N} + h1_{E^c}$. Using the nullness of E_n for each $n = 1, 2, \dots$, we have:

$$\begin{aligned} f^N &= f1_{E_1} + f1_{E_2} + \dots + f1_{E_N} + x1_{E\setminus A^N} + h1_{E^c} \\ &\sim g1_{E_1} + f1_{E_2} + f1_{E_3} + \dots + f1_{E_N} + x1_{E\setminus A^N} + h1_{E^c} \\ &\sim g1_{E_1} + g1_{E_2} + f1_{E_3} + \dots + f1_{E_N} + x1_{E\setminus A^N} + h1_{E^c} \\ &\quad \dots \\ &\sim g1_{E_1} + g1_{E_2} + \dots + g1_{E_N} + x1_{E\setminus A^N} + h1_{E^c} \\ &= g^N. \end{aligned}$$

It is easy to see that $f^N \to f1_E + h1_{E^c}$ and $g^N \to g1_E + h1_{E^c}$. Since $f, g, h \in \mathcal{F}$, there exist $\bar{x}_a, \underline{x}_a$ such that $\bar{x}_a \succcurlyeq a(\omega) \succcurlyeq \underline{x}_a, \forall \omega \in \Omega$, for $a = f, g, h$. Define $u(\omega) \equiv \max\{\bar{x}_f, \bar{x}_g, \bar{x}_h\}$ and $l(\omega) \equiv \min\{\underline{x}_f, \underline{x}_g, \underline{x}_h\}$. Note that these values are well-defined because $\succcurlyeq$ is complete on X. Therefore, continuity implies $f1_E + h1_{E^c} \sim g1_E + h1_{E^c}$. Since f, g, h are arbitrary, E is null. □

Lemma A.2. *If $A \subset E$, $A \in \Sigma$ and E is $\succcurlyeq$-null then A is $\succcurlyeq$-null.*

Proof. Let $f, g, h \in \mathcal{F}$. Define $f' \equiv f1_A + h1_{E\setminus A}$ and $g' \equiv g1_A + h1_{E\setminus A}$. Since E is null, we have:

$$\begin{bmatrix} f'(\omega), & \text{if } \omega \in E \\ h(\omega), & \text{if } \omega \notin E \end{bmatrix} \sim \begin{bmatrix} g'(\omega), & \text{if } \omega \in E \\ h(\omega), & \text{if } \omega \notin E \end{bmatrix}.$$

Note, however that the left and right side above are respectively:

$$\begin{bmatrix} f(\omega), & \text{if } \omega \in A \\ h(\omega), & \text{if } \omega \notin A \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} g(\omega), & \text{if } \omega \in A \\ h(\omega), & \text{if } \omega \notin A \end{bmatrix}.$$

Therefore, A is $\succcurlyeq$-null. □

Definition A.3. Let $\mathcal{N}$ denote the set of $\succcurlyeq$-null sets and let $\mathcal{H}$ be a sub-σ-field of Σ. Let $\overline{\mathcal{H}}$ denote the following class of sets:

$$\overline{\mathcal{H}} \equiv \{A \in \Sigma \colon \exists B \in \mathcal{H},\ A \Delta B \in \mathcal{N}\},$$

where $A\Delta B \equiv (A \cap B^c) \cup (A^c \cap B)$.

Lemma A.4. *$\overline{\mathcal{H}}$ is a σ-field containing $\mathcal{H}$. More precisely, $\overline{\mathcal{H}} = \mathcal{H} \vee \mathcal{N}$ is the smallest σ-field containing both $\mathcal{H}$ and $\mathcal{N}$.*

Proof. It is obvious that $\overline{\mathcal{H}} \supset \mathcal{H}$ and $\emptyset \in \overline{\mathcal{H}}$. If $A \in \overline{\mathcal{H}}$, let $B \in \mathcal{H}$ be such that $A\Delta B \in \mathcal{N}$. Since $B^c \in \mathcal{H}$ and $A^c \Delta B^c = A\Delta B$, then $A^c \in \overline{\mathcal{H}}$. Finally, assume that $\{A_n\}_{n\in\mathcal{N}} \subset \overline{\mathcal{H}}$. Then there exist $\{B_n\}_{n\in\mathcal{N}} \subset \mathcal{H}$, such that $E_n \equiv A_n \Delta B_n \in \mathcal{N}$. Let $A \equiv \bigcup_{n\in\mathcal{N}} A_n$ and $B = \bigcup_{n\in\mathcal{N}} B_n$. It is clear that $B \in \mathcal{H}$ and

$$\begin{aligned}
A \Delta B &= \left[\left(\bigcup_{n\in\mathcal{N}} A_n\right)\cap\left(\bigcup_{n\in\mathcal{N}} B_n\right)^c\right]\cup\left[\left(\bigcup_{n\in\mathcal{N}} A_n\right)^c\cap\left(\bigcup_{n\in\mathcal{N}} B_n\right)\right] \\
&= \left[\left(\bigcup_{n\in\mathcal{N}} A_n\right)\cap\left(\bigcap_{n\in\mathcal{N}} B_n^c\right)\right]\cup\left[\left(\bigcap_{n\in\mathcal{N}} A_n^c\right)\cap\left(\bigcup_{n\in\mathcal{N}} B_n\right)\right] \\
&\subset \left[\bigcup_{n\in\mathcal{N}} \left(A_n\cap B_n^c\right)\right]\cup\left[\bigcup_{n\in\mathcal{N}} \left(A_n^c\cap B_n\right)\right] \\
&= \bigcup_{n\in\mathcal{N}} E_n.
\end{aligned}$$

The set $\bigcup_{n\in\mathcal{N}} E_n$ is $\succcurlyeq$-null by Lemma A.1. Since $A\Delta B$ is Σ-measurable and is contained in the $\succcurlyeq$-null set $\bigcup_{n\in\mathcal{N}} E_n$, Lemma A.2 shows that $A\Delta B \in \mathcal{N}$. This establishes that $\overline{\mathcal{H}}$ is a σ-field.

Finally, it is clear that $\overline{\mathcal{H}} \supset \mathcal{H}\cup\mathcal{N}$. Since it is a σ-field, then $\overline{\mathcal{H}} \supset \mathcal{H}\vee\mathcal{N}$. On the other hand, if $A\in\overline{\mathcal{H}}$, there exists $B\in\mathcal{H}$ such that $A\Delta B\in\mathcal{N}$. Then, $E = A\setminus B \subset A\Delta B$ is Σ-measurable and therefore, $\succcurlyeq$-null. But $A = B\cup E\in\mathcal{H}\cup\mathcal{N}$ and therefore, $A\in\mathcal{H}\vee\mathcal{N}$. □

A.2. Reduction to real-valued functions

Under our assumptions, Herstein and Milnor's [17] theorem implies the existence of a linear function $u: X\to\mathcal{R}$, unique up to affine transformations, representing $\succcurlyeq$ on X. By linear, we mean that $u(\alpha x + (1-\alpha)y) = \alpha u(x) + (1-\alpha)u(y)$, for every $x, y\in X$ and $\alpha\in[0,1]$. Moreover, since u is linear, $I \equiv u(X)\subset\mathcal{R}$ is a convex subset of $\mathcal{R}$, that is, I is an interval. By non-triviality (there exists x, y such that $x\succ y$) and taking an affine transformation of u if needed, we can assume that $[0,1]\subset I$. Moreover, pointwise continuity implies that u is continuous.

Now, $u: X\to\mathcal{R}$ induces a preference order on the set $\mathcal{D}$ of the Σ-measurable functions $f:\Omega\to I$ as follows: for each $f, g\in\mathcal{D}$,

$$f \succcurlyeq^{\mathcal{D}} g \equiv u^{-1}(f) \succcurlyeq u^{-1}(g). \tag{18}$$

In fact, (18) is not completely formal, since u is not invertible in general and, therefore, $u^{-1}\circ f:\Omega\to X$ is actually a correspondence rather than an act in $\mathcal{F}$. However, by monotonicity all selections of this correspondence will be indifferent, so that $\succcurlyeq^{\mathcal{D}}$ is well-defined by (18). Observe that since u represents $\succcurlyeq$ when restricted to consequences, we have $x, y\in I$, $x\geqslant y \Leftrightarrow x\succcurlyeq^{\mathcal{D}} y$.

Conversely, given a preference $\succcurlyeq^{\mathcal{D}}$ on $\mathcal{D}$ and function $u: X\to I$ we can define a preference $\succcurlyeq'$ on $\mathcal{F}$ by the following: for any $f, g\in\mathcal{F}$,

$$f \succcurlyeq' g \equiv u(f) \succcurlyeq^{\mathcal{D}} u(g). \tag{19}$$

It is easy to see that if we start with $\succcurlyeq$ on $\mathcal{F}$, obtain $\succcurlyeq^{\mathcal{D}}$ on $\mathcal{D}$ by (18) and use this $\succcurlyeq^{\mathcal{D}}$ together with u in (19) to define a preference $\succcurlyeq'$, then $\succcurlyeq$ and $\succcurlyeq'$ coincide.

In sum, a preference $\succcurlyeq$ on X-valued functions $\mathcal{F}$ defines a preference $\succcurlyeq^{\mathcal{D}}$ on real-valued functions $\mathcal{D}$ and a preference $\succcurlyeq^{\mathcal{D}}$ on $\mathcal{D}$ *together with a function* $u: X\to\mathcal{R}$ defines a preference $\succcurlyeq$ on $\mathcal{F}$. The next proposition establishes a useful link between the two:

Lemma A.5. *Consider one of the following two cases*:

1. *It is given a preference* $\succcurlyeq$ *on* $\mathcal{F}$ *satisfying our assumptions and let* $\succcurlyeq^{\mathcal{D}}$ *be defined as in* (18).

2. *It is given a preference* $\succcurlyeq^{\mathcal{D}}$ *on* $\mathcal{D}$ *and a linear* $u : X \to \mathcal{R}$, *let* $\succcurlyeq$ *be defined by* (19).

In any case, $\succcurlyeq$ *is* Γ*-invariant if and only if* $\succcurlyeq^{\mathcal{D}}$ *is* Γ*-invariant.*

Proof. Fix an act $f : \Omega \to X$. Since $f \in \mathcal{F}$ is bounded, there exist $\underline{x}, \bar{x}$ such that $\bar{x} \succcurlyeq f(\omega) \succcurlyeq \underline{x}$, for all $\omega \in \Omega$. Herstein and Milnor [17] also show that for any z satisfying $\bar{x} \succcurlyeq z \succcurlyeq \underline{x}$, there exists a unique $\alpha \in [0,1]$ such that $z \sim \alpha\bar{x} + (1-\alpha)\underline{x}$. Therefore, $u(f(\Omega)) \subset u([\underline{x}, \bar{x}])$, where $[\underline{x}, \bar{x}] \equiv \{\alpha\bar{x} + (1-\alpha)\underline{x}\colon \alpha \in [0,1]\}$, and the function u is invertible when restricted to $[\underline{x}, \bar{x}]$; in this proof, u^{-1} will denote the inverse function of this restriction.

Since u is linear in X, then for every $\omega \in \Omega$,

$$u\left(\frac{f\circ\gamma_1 + \cdots + f\circ\gamma_n}{n}(\omega)\right) = \frac{u\circ f\circ\gamma_1(\omega) + \cdots + u\circ f\circ\gamma_n(\omega)}{n}. \tag{20}$$

We claim that u^{-1} is also linear. To see this, observe that:

$$\begin{aligned} &u\big(\alpha z + (1-\alpha)w\big) = \alpha u(z) + (1-\alpha)u(w) \\ &\quad\Rightarrow\quad u^{-1}\big[u\big(\alpha z + (1-\alpha)w\big)\big] = u^{-1}\big(\alpha u(z) + (1-\alpha)u(w)\big) \\ &\quad\Rightarrow\quad \alpha z + (1-\alpha)w = u^{-1}\big(\alpha u(z) + (1-\alpha)u(w)\big). \end{aligned}$$

If we put $u(z) = a$ and $u(w) = b$, so that $z = u^{-1}(a)$ and $w = u^{-1}(b)$, the last equation is just:

$$u^{-1}\big(\alpha a + (1-\alpha)b\big) = \alpha u^{-1}(a) + (1-\alpha)u^{-1}(b),$$

that is, u^{-1} is linear as we claimed.

Now, assume that $\succcurlyeq^{\mathcal{D}}$ is Γ-invariant, that is, for every $\gamma_1, \ldots, \gamma_n \in \Gamma$ and $\tilde{f} \in \mathcal{D}$, the following holds:

$$\tilde{f} \sim \frac{\tilde{f}\circ\gamma_1 + \cdots + \tilde{f}\circ\gamma_n}{n}. \tag{21}$$

Fix $f \in \mathcal{F}$. From (19),

$$f \sim \frac{f\circ\gamma_1 + \cdots + f\circ\gamma_n}{n} \quad\Longleftrightarrow\quad u(f) \sim^{\mathcal{D}} u\left(\frac{f\circ\gamma_1 + \cdots + f\circ\gamma_n}{n}\right).$$

Using (20) and (21), we obtain that $\succcurlyeq$ is Γ-invariant. The proof of the converse statement is analogous. □

The above results shows that it is enough to consider preferences over bounded real valued functions with values in $I = u(X)$. Since u is affine, $u(X)$ will be an interval $I \subset \mathcal{R}$ (which may be the whole $\mathcal{R}$). We can calibrate u so that the two outcomes $x, y \in X$ assumed to exist in Assumption 5, have values 0 and 1, respectively. In particular, this implies that the interval $[0,1] \subset I \subset \mathcal{R}$ and that for any $x, y \in [0,1]$, $x > y \Leftrightarrow x \succ^{\mathcal{D}} y$.

In next sections, we will consider only $\succcurlyeq^{\mathcal{D}}$ and, for convenience, we will drop the superscript $\mathcal{D}$, denoting it only by $\succcurlyeq$. The following result summarizes the properties of $\succcurlyeq^{\mathcal{D}}$ that we will need and which are implied by the assumptions on $\succcurlyeq$ given in the body of the paper.

Corollary A.6. $\succcurlyeq^{\mathcal{D}}$ *is defined for functions* $f : \Omega \to I \subset \mathcal{R}$ *and satisfies the following*:

1. (Preorder). $\succcurlyeq^{\mathcal{D}}$ *is reflexive and transitive.*
2. (Monotonicity). *If* $f(\omega) \geqslant g(\omega)$ *for all* $\omega \in \Omega$, *then* $f \succcurlyeq^{\mathcal{D}} g$.
3. (Pointwise continuity). *Suppose that for a given pair of acts* $f, g \in \mathcal{D}$ *there are sequences* $\{f^n\}$, $\{g^n\}$ *such that*: (i) $f^n \to f$ *and* $g^n \to g$; (ii) $|f^n(\omega)| \leqslant b(\omega)$ *and* $|g^n(\omega)| \leqslant b(\omega)$, *for all* ω *and some* $b \in \mathcal{D}$; *and* (iii) $f^n \succcurlyeq^{\mathcal{D}} g^n$ *for all* n. *Then* $f \succcurlyeq^{\mathcal{D}} g$.
4. (Non-triviality). *For any* $x, y \in I$, $x \geqslant y \Leftrightarrow x \succcurlyeq^{\mathcal{D}} y$ *and* $[0, 1] \subset I$.

In the next section, we will prove theorems about preferences $\succcurlyeq^{\mathcal{D}}$ defined on measurable functions $f : \Omega \to I \subset \mathcal{R}$. It is useful to observe that to a preference $\succcurlyeq^{\mathcal{D}}$ corresponds more than one $\succcurlyeq$ on $\mathcal{F}$, because we can take different utility functions. To clarify this, suppose that we begin with a preference $\succcurlyeq$ on $\mathcal{F}$ obtain $\succcurlyeq^{\mathcal{D}}$ as in (18) using the u related to $\succcurlyeq$, and now consider $\succcurlyeq^{\mathcal{D}}$ with another linear function $u' : X \to \mathcal{R}$. If we use u' and $\succcurlyeq^{\mathcal{D}}$ as in (19), we obtain $\succcurlyeq'$:

$$f \succcurlyeq' g \equiv u'(f) \succcurlyeq^{\mathcal{D}} u'(g).$$

Observe that if $u \neq u'$ then $\succcurlyeq$ and $\succcurlyeq'$ will be different as well. However, we have the following:

Lemma A.7. $\succcurlyeq$ *and* $\succcurlyeq'$ *have the same null sets. Moreover,* $\succcurlyeq$ *and* $\succcurlyeq^{\mathcal{D}}$ *have the same null sets.*

Proof. Since $\succcurlyeq^{\mathcal{D}}$ can be obtained from $\succcurlyeq'$ using u' (instead of u) in (18), it is enough to show that $\succcurlyeq$ and $\succcurlyeq^{\mathcal{D}}$ have the same null sets. Let $f, g, h \in \mathcal{F}$. Then, (18) and (19) imply that:

$$\begin{bmatrix} f(\omega), & \text{if } \omega \in A \\ h(\omega), & \text{if } \omega \notin A \end{bmatrix} \sim \begin{bmatrix} g(\omega), & \text{if } \omega \in A \\ h(\omega), & \text{if } \omega \notin A \end{bmatrix}$$
$$\Longleftrightarrow \quad \begin{bmatrix} u(f(\omega)), & \text{if } \omega \in A \\ u(h(\omega)), & \text{if } \omega \notin A \end{bmatrix} \sim^{\mathcal{D}} \begin{bmatrix} u(g(\omega)), & \text{if } \omega \in A \\ u(h(\omega)), & \text{if } \omega \notin A \end{bmatrix}. \qquad \square$$

A.3. Subjective ergodic theorems

Consider a preference $\succcurlyeq$ defined on the set $\mathcal{D}$ of all Σ-measurable functions $f : \Omega \to I \subset \mathcal{R}$, satisfying all the assumptions listed in Corollary A.6. In this section, we will simplify notation by writing $\succcurlyeq$ instead of $\succcurlyeq^{\mathcal{D}}$. No confusion should arise since we consider no other preference here. The next two theorems appear in De Castro and Al-Najjar [7]; the notation is modified to fit the current setup:

Theorem A.8 *(The subjective ergodic theorem). The following conditions are equivalent*:

1. $\succcurlyeq$ *is* τ*-invariant.*
2. *For every act* f, *the empirical limit* $f^\star$ *is well-defined off a* $\succcurlyeq$*-null event.*

In this case, $f^\star \sim f$,[21] *and* $f^\star$ *is* τ*-invariant, that is,* $f^\star(\tau\omega) = f^\star(\omega)$, *whenever the limit exists. If* $\succcurlyeq$ *is* τ*-ergodic, then* $f^\star$ *is constant except in a* $\succcurlyeq$*-null set.*

For stating the next theorem, we need some notation. Let $\Delta(\Omega)$ be the set of all probability measures in Ω, endowed with its usual weak*-topology. Let $\mathcal{P}^{er}_\tau \subset \Delta(\Omega)$ denote the set of all

[21] Extend $f^\star$ arbitrarily at ω's where the limit does not exist.

τ-ergodic probability measures. As usual, it is convenient to write this set of τ-ergodic measures will be indexed by a set of parameters Θ, that is, $\mathcal{P}_\tau^{er} = \{P^\theta\}_{\theta\in\Theta}$. Of course, this set of parameters can be itself identified with $\mathcal{P}_\tau^{er}$ and thus inherit its topological and measurable structure.

Theorem A.9. *For every τ-invariant preference $\succcurlyeq$, there exists a decomposition map $\vartheta : \Omega \to \Theta$ such that (Θ, ϑ) is sufficient for $\succcurlyeq$.*

Appendix B. Proofs

B.1. Proof of Theorem 1

Let $\mathbb{P}$ denote the set of preferences on $\mathcal{F}$ satisfying assumptions 1–5. As discussed in Section A.2, for each preference $\succcurlyeq \in \mathbb{P}$ there is a linear utility function $u : X \to \mathcal{R}$ that represents $\succcurlyeq$ on X. Analogously, let $\mathbb{P}^{\mathcal{D}}$ denote the set of preferences on real-valued functions $\mathcal{D}$ satisfying the properties stated in Corollary A.6. As discussed in Section A.2, for each $\succcurlyeq \in \mathbb{P}$ it corresponds a $\succcurlyeq^{\mathcal{D}} \in \mathbb{P}^{\mathcal{D}}$, defined by (18) and, conversely, to each $\succcurlyeq^{\mathcal{D}} \in \mathbb{P}^{\mathcal{D}}$ and linear u corresponds a $\succcurlyeq \in \mathbb{P}$, as defined by (19).

Fix $f \in \mathcal{F}$ and a linear function $v : X \to \mathcal{R}$, and define the sets:

$$A \equiv \left\{\omega \in \Omega \colon \exists \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f\left(\tau^j\omega\right)\right\}$$

and

$$\tilde{A}_v \equiv \left\{\omega \in \Omega \colon \exists \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} v\left[f\left(\tau^j\omega\right)\right]\right\}.$$

Notice that the above sets do not depend on any preference.

Since X is finite dimensional, we can find a countable set $\mathcal{U}$ of linear functions $u : X \to \mathcal{R}$ such that $x^n \to x$ if and only if $u(x^n) \to u(x)$ for every $u \in \mathcal{U}$ (a finite dimensional space has only one topology). Define $\tilde{A} \equiv \bigcap_{u\in\mathcal{U}} \tilde{A}_u$. Thus, $A = \tilde{A}$.

Theorem A.8 shows that $\tilde{A}_u^c = \Omega \setminus \tilde{A}_u$ is $\succcurlyeq^{\mathcal{D}}$-null for any $\succcurlyeq^{\mathcal{D}} \in \mathbb{P}^{\mathcal{D}}$. By Lemma A.1, $\tilde{A}^c$ is also $\succcurlyeq^{\mathcal{D}}$ for any $\succcurlyeq^{\mathcal{D}} \in \mathbb{P}^{\mathcal{D}}$. By Lemma A.7, $A^c = \tilde{A}^c$ is $\succcurlyeq$-null for any $\succcurlyeq \in \mathbb{P}$.

By Theorem A.9, there exists a decomposition map $\vartheta : \Omega \to \Theta$ such that (Θ, ϑ) is sufficient for $\succcurlyeq^{\mathcal{D}}$ for any $\succcurlyeq^{\mathcal{D}} \in \mathbb{P}^{\mathcal{D}}$. In particular, this means that $u(f) \sim^{\mathcal{D}} \int u(f)\, dP^{\vartheta(\cdot)} = u(\int f\, dP^{\vartheta(\cdot)})$, because u is linear. By (19), $f \sim \int f\, dP^{\vartheta(\cdot)}$ and (Θ, ϑ) is sufficient for $\succcurlyeq \in \mathbb{P}$. By construction, $u(f)^\star(\omega) = \int u(f)\, dP^{\vartheta(\omega)}$ in a set Ω' whose complement is $\succcurlyeq$-null and, repeating the above argument, we can have $f^\star(\omega) = \int_\Omega f\, dP^{\vartheta(\omega)}$ in this set. □

B.2. Proof of Theorem 2

Although the assumption $u(\Delta(C)) = \mathcal{R}$ rules out C finite, this is not a problem for this and related results.

(2) $\Rightarrow$ (1): Since there exists a map $\tilde{\vartheta} : \Omega \to \Theta_\Gamma$ such that $(\Theta_\Gamma, \tilde{\vartheta})$ is a sufficient parametrization of $\succcurlyeq$ and $P^{\vartheta(\cdot)} \circ \gamma^{-1} = P^{\vartheta(\cdot)}$, $\forall\gamma \in \Gamma$,

$$f \circ \gamma \sim \int (f \circ \gamma)\, dP^{\tilde{\vartheta}(\cdot)} = \int f\, d\left(P^{\tilde{\vartheta}(\cdot)} \circ \gamma^{-1}\right) = \int f\, dP^{\tilde{\vartheta}(\cdot)} \sim f,$$

which establishes (1).

(1) $\Rightarrow$ (2): It is enough to establish that $A \equiv \vartheta^{-1}(\Theta \setminus \Theta_\Gamma)$ is $\succcurlyeq$-null. Let $A_\gamma \equiv \{\omega \in A\colon\ P^{\vartheta(\omega)} \neq P^{\vartheta(\omega)} \circ \gamma^{-1}\}$. Then, $A = \bigcup_{\gamma \in \Gamma} A_\gamma$. Since Γ is countable, it is enough to prove that A_γ is $\succcurlyeq$-null for each γ.

Fix $\gamma \in \Gamma$ and denote $\Theta_\gamma \equiv \{\theta \in \Theta_\Gamma^c\colon\ P^\theta \neq P^\theta \circ \gamma^{-1}\}$. Observe that $A_\gamma = \bigcup_{\theta \in \Theta_\gamma} \vartheta^{-1}(\theta) = \vartheta^{-1}(\Theta_\gamma)$. For each $\theta \in \Theta_\gamma$, let $B^\theta \subset \vartheta^{-1}(\theta)$ be such that $\alpha^\theta \equiv P^\theta(B^\theta) \neq P^\theta(\gamma^{-1}(B^\theta)) \equiv \beta^\theta$.

We will first prove that the set $\tilde{A}_\gamma \equiv \bigcup_{\theta \in \Theta_\gamma} B^\theta$ is $\succcurlyeq$-null. For an absurd, assume that $\tilde{A}_\gamma$ is not $\succcurlyeq$-null, that is, there exist $f, g \in \mathcal{F}$ such that

$$f' \equiv \begin{bmatrix} f(\omega), & \text{if } \omega \in \tilde{A}_\gamma \\ g(\omega), & \text{if } \omega \notin \tilde{A}_\gamma \end{bmatrix} \quad \text{and} \quad g$$

are incomparable. Since (Θ, ϑ) is sufficient, we can assume that $f, g \in \mathcal{F}_\vartheta$. This means that $\vartheta(\omega) = \vartheta(\omega') \Rightarrow f(\omega) = f(\omega')$ and a similar condition hold for g. Since $\alpha^\theta \neq \beta^\theta$, for each $\theta \in \Theta_\gamma$, we can find x^θ and y^θ such that:

$$\begin{cases} \alpha^\theta u(x^\theta) + (1 - \alpha^\theta) u(y^\theta) = u(f(\omega)), \\ \beta^\theta u(x^\theta) + (1 - \beta^\theta) u(y^\theta) = u(g(\omega)), \end{cases}$$

for every $\omega \in \vartheta^{-1}(\theta)$.

Define $h \in \mathcal{F}$ as follows:

$$h(\omega) = \begin{cases} x^\theta, & \text{if } \omega \in B^\theta, \theta \in \Theta_\gamma, \\ y^\theta, & \text{if } \omega \in \vartheta^{-1}(\theta) \setminus B^\theta, \theta \in \Theta_\gamma, \\ g(\omega), & \text{otherwise.} \end{cases}$$

Therefore, if $\vartheta(\omega) = \theta \in \Theta_\gamma$,

$$\begin{aligned} \int u(h)\, dP^{\vartheta(\omega)} &= u(x^\theta) P^\theta(B^\theta) + u(y^\theta) P^\theta(\vartheta^{-1}(\theta) \setminus B^\theta) \\ &= \alpha^\theta u(x^\theta) + (1 - \alpha^\theta) u(y^\theta) \\ &= u(f(\omega)). \end{aligned}$$

If $\vartheta(\omega) = \vartheta(\gamma(\omega)) = \theta \in \Theta_\gamma$,[22]

$$\begin{aligned} \int u(h \circ \gamma)\, dP^{\vartheta(\omega)} &= u(x^\theta) P^\theta(\gamma^{-1}(B^\theta)) + u(y^\theta) P^\theta[\gamma^{-1}(\vartheta^{-1}(\theta) \setminus B^\theta)] \\ &= \beta^\theta u(x^\theta) + (1 - \beta^\theta) u(y^\theta) \\ &= u(g(\omega)). \end{aligned}$$

On the other hand, if $\vartheta(\omega) = \theta \notin \Theta_\gamma$, then $P^\theta = P^\theta \circ \gamma^{-1}$, which implies that:

$$\int h \circ \gamma\, dP^{\vartheta(\omega)} = \int h\, d(P^{\vartheta(\omega)} \circ \gamma^{-1}) = \int h\, dP^{\vartheta(\omega)}.$$

Also, in this case,

$$\int h\, dP^{\vartheta(\omega)} = \int g\, dP^{\vartheta(\omega)} = g(\omega) = f'(\omega),$$

because we chose $g \in \mathcal{F}_\vartheta$.

22 Recall that ϑ is Γ-invariant.

Then $h \sim \int h\,dP^{\vartheta(\cdot)} = f'$ and $h \circ \gamma \sim \int (h \circ \gamma)\,dP^{\vartheta(\cdot)} = g$, but yet f' and g are incomparable, which contradicts $h \sim h \circ \gamma$. The contradiction establishes that $\tilde{A}_\gamma$ is $\succcurlyeq$-null.

The above argument can now be applied to $A_\gamma \setminus \tilde{A}_\gamma = \bigcup_{\theta \in \Theta_\gamma} [\vartheta^{-1}(\theta) \setminus B^\theta]$ to conclude that $A_\gamma \setminus \tilde{A}_\gamma$ is also $\succcurlyeq$-null. Since A_γ is the union of two $\succcurlyeq$-null sets, it is $\succcurlyeq$-null. This concludes the proof. □

B.3. *Proof of Proposition 5*

Given a decomposition map ϑ, we identify the parameter-based acts with acts $\mathcal{F}_\Gamma \subset \mathcal{F}$ that are measurable with respect to $\mathcal{E}_\Gamma$ as follows.[23] Define $M^\vartheta : \mathcal{F} \to \mathcal{F}_\Gamma$ by:

$$M^\vartheta(f)(\omega) = \int_\Omega f\,dP^{\vartheta(\omega)}. \tag{22}$$

Since the decomposition map defines a universal conditional expectation, the map M^ϑ acts as an identity in $\mathcal{F}_\Gamma$. Notice that if we have $\vartheta(\omega) = \theta$, then $M^\vartheta(f)(\omega) = \Psi(f)(\theta)$. That is, we have $M^\vartheta(f)(\cdot) = \Psi(f)(\vartheta(\cdot))$ and, conversely, $\Psi(f)(\cdot) = M^\vartheta(f)(\vartheta^{-1}(\cdot))$. Therefore, when restricted to $\mathcal{F}_\Gamma$, the map $\Psi : \mathcal{F}_\Gamma \to \mathbb{F}$ can be seen as one-to-one (up to functions that differ on $\succcurlyeq$-null sets). Therefore, the inverse $\Psi^{-1} : \mathbb{F} \to \mathcal{F}_\Gamma$ is given by:

$$\Psi^{-1}(F)(\omega) = F\big(\vartheta(\omega)\big).$$

Since $M^\vartheta(f)(\cdot) = \Psi(f)(\vartheta(\cdot))$, $\Psi(f)(\cdot) = M^\vartheta(f)(\vartheta^{-1}(\cdot))$, $f \sim M^\vartheta(f)$ is equivalent to (6).

Given an aggregator $\succcurlyeq$ satisfying assumptions 1–5, define $\succcurlyeq$ on $\mathcal{F}_\Gamma$ by (6). Imposing that (Θ, ϑ) is sufficient for $\succcurlyeq$, this defines $\succcurlyeq$ uniquely. It is easy to see that assumptions 1–5 hold (continuity holds by the dominated convergence theorem and the above definition).

Conversely, given $\succcurlyeq$, define $\succcurlyeq$ by:

$$F \succcurlyeq G \equiv \Psi^{-1}(F) \succcurlyeq \Psi^{-1}(G).$$

Given that Ψ is one-to-one up to null sets, (6) also holds. Again, it is easy to see that $\succcurlyeq$ satisfies assumptions 1–5. □

References

[1] N.I. Al-Najjar, L.D. Castro, Observability and 'Second-Order Act', Northwestern University, 2010.
[2] P. Billingsley, Probability and Measure, 3rd ed., Wiley Ser. Prob. Math. Stat., Prob. Math. Stat., John Wiley & Sons Inc., A Wiley–Interscience Publication, New York, 1995.
[3] S. Cerreia, F. Maccheroni, M. Marinacci, L. Montrucchio, Uncertainty averse preferences, J. Econ. Theory 146 (2011) 1275–1330.
[4] S. Cerreia, F. Maccheroni, M. Marinacci, L. Montrucchio, Ambiguity and robust statistics, J. Econ. Theory 148 (2013) 974–1049.
[5] S. Chew, J. Sagi, Small worlds: modeling attitudes toward sources of uncertainty, J. Econ. Theory 139 (1) (2008) 1–24.
[6] A.P. Dawid, Intersubjective statistical models, in: G. Koch, F. Spizzichino (Eds.), Exchangeability in Probability and Statistics, North-Holland, Amsterdam, 1982, pp. 217–232.
[7] L. De Castro, N.I. Al-Najjar, A Subjective Foundation of Objective Probability, Northwestern University, 2009.

[23] This proof is written in terms of a Γ-parametrization, but it can be easily adapted for a parametrization without reference to semi-groups.

[8] B. de Finetti, La prévision: ses lois logiques, ses sources subjectives, Ann. Inst. Henri Poincaré 7 (1937) 1–68.
[9] P. Diaconis, D. Freedman, De Finetti's theorem for Markov chains, Ann. Probab. 8 (1) (1980) 115–130.
[10] P. Diaconis, D. Freedman, Partial exchangeability and sufficiency, in: Proceedings of the Indian Statistical Intitute Golden Jubilee International Conference on Statistics: Applications and New Directions, 1984, pp. 205–236.
[11] E.B. Dynkin, Sufficient statistics and extreme points, Ann. Probab. 6 (5) (1978) 705–730.
[12] L. Epstein, K. Seo, Symmetry of evidence without evidence of symmetry, Theoretical Econ. 5 (2010) 313–368.
[13] H. Ergin, F. Gul, A subjective theory of compound lotteries, J. Econ. Theory (2009) 899–929.
[14] P. Ghirardato, F. Maccheroni, M. Marinacci, M. Siniscalchi, A subjective spin on roulette wheels, Econometrica 71 (6) (2003) 1897–1908.
[15] I. Gilboa, F. Maccheroni, M. Marinacci, D. Schmeidler, Objective and subjective rationality in a multiple prior model, Econometrica 78 (2010) 755–770.
[16] S. Grant, B. Polak, T. Strzalecki, Second-order expected utility, discussion paper, Harvard University, 2009.
[17] I. Herstein, J. Milnor, An axiomatic approach to measurable utility, Econometrica 21 (2) (1953) 291–297.
[18] P. Klibanoff, M. Marinacci, S. Mukerji, A smooth model of decision making under ambiguity, Econometrica 73 (6) (2005) 1849–1892.
[19] P. Klibanoff, S. Mukerji, K. Seo, Relevance and Symmetry, Northwestern University, 2010.
[20] S. Lauritzen, Extreme point models in statistics, Scand. J. Statist. (1984) 65–91.
[21] R. Nau, Uncertainty aversion with second-order utilities and probabilities, in: 2nd International Symposium on Imprecise Probabilities and Their Applications, June 2001, Ithaca, New York, 2001.
[22] R. Nau, Uncertainty aversion with second-order utilities and probabilities, Manage. Sci. 52 (1) (2006) 136.
[23] W. Neilson, Ambiguity aversion: an axiomatic approach using second order probabilities, discussion paper, mimeo, 1993.
[24] W. Neilson, A simplified axiomatic approach to ambiguity aversion, J. Risk Uncertainty 41 (2010) 113–124.
[25] T. Strzalecki, Axiomatic foundations of multiplier preferences, Econometrica 79 (2011) 47–73.
[26] V. Varadarajan, Groups of automorphisms of Borel spaces, Trans. Amer. Math. Soc. 109 (2) (1963) 191–220.