Parametric Representation of Preferences[*][†]

Nabil I. Al-Najjar[‡]   and   Luciano De Castro[§]

First draft: February 2010;
This version: September 2010

Abstract

A preference is invariant with respect to a transformation τ if its ranking of acts is unaffected by a reshuffling of the states under τ. We show 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. This property holds for all non-trivial preferences, provided only that they are reflexive, transitive, monotone, continuous and mixture linear.

Contents

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) systematic 'patterns' captured by the parameters, and (2) idiosyncratic 'noise' around these patterns. In the stylized example of repeated coin tosses, for instance, 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 structure (as for example in applications of spatial statistics to networks). But 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]

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

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 parameterization (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

---

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

conditions, to de Finetti (1937)'s celebrated representation theorem. In this paper we consider general countable semi-groups of transformation which cover exchangeability, but also partial exchangeability, stationary distributions, Markovian structures, among others.

The sufficiency of a parametrization defines an operator:

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

that maps the state-based acts $\mathcal{F}$ to their corresponding elements in the set of *parameter-based acts* $\mathbb{F}$. We call a binary relation on $\mathbb{F}$ an *aggregator*, reflecting 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 $f, g \in \mathcal{F}$

$$f \succcurlyeq g \iff \Psi(f) \succcurlyeq \Psi(g).$$

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 (*e.g.,* start with permutations, so Θ is the set of 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 preferences, ... etc). Third, derive an invariant preference $\succcurlyeq$ on the state-based acts $\mathcal{F}$.

In Section 4 we provide classes of examples of aggregators and a 'template' for translating theories about aggregating parameter uncertainty into preferences on the primitive states. In Section 4.3 we develop in greater detail an instance of this template, focusing on aggregators of the form:

$$\mathcal{V}(F) = \int_\Theta \phi(F(\theta)) \, d\mu(\theta) = \int_\Theta \phi \left(\int_\Omega u(f) \, dP^\theta \right) d\mu(\theta), \qquad (2)$$

for $F = \Psi(f)$, a von Neumann-Morgenstern utility function u, and a function from reals to reals ϕ. We call the preferences corresponding to such

2

aggregators *second-order subjective expected utility preferences.* These are preferences that (to our knowledge) were first introduced by Neilson((1993), (2009)) and used by, among others, Nau (2001), Nau (2006), Ergin and Gul (2009), Chew and Sagi (2008), Strzalecki (2010), Grant, Polak, and Strzalecki (2009).

The models of Neilson (1993) and Strzalecki (2010) have an interesting interpretation in our setting. They consider functionals of the form[2]

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

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:

$$\int_\Omega \phi \left(\int_\Omega u(f)\, d\delta_\omega \right)\, dP,$$

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; see Section 3.3.3.

The aggregator (2) may therefore be viewed as a generalization of Neilson's representation to coarser parameterizations (*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 (1993) and Strzalecki (2010) the risky events are singletons. We interpret Neilson (1993)'s and Strzalecki (2010)'s decision makers as ones who do not not do such pooling, so each state is its own parameter. Section 4.3 discusses this in more details.

Our model also helps clarify Klibanoff, Marinacci, and Mukerji (2005)'s functional form which is similar to (2), but where the outer integral is over

---

[2]Strzalecki (2010) consider an important special case of this function, namely the one corresponding to multiplier preferences.

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 4.4 and in greater details in the companion note, Al-Najjar and De Castro (2010), 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, Marinacci, and Mukerji 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 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 $\alpha\theta + (1-\alpha)\theta'$ has 'occurred' do not correspond to typical 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 (1978), Dawid (1982), Lauritzen (1984) among others. From the perspective of economic and game theoretic modeling, the treatment of parameters in statistics is unhelpful: 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. ob-

jectivity 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. In Al-Najjar, De Castro, and Pai (2010) we use this connection with frequencies to test the decision maker's invariance assumption. Roughly, we show that the parametrization (Θ, ϑ) can be tested, but the subjective aggregator over parameters cannot. In the simplest setting of exchangeability, that the data is consistent with the i.i.d.ness assumption can be tested, but the subjective belief μ over parameters cannot.

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.

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 is $\Delta(C)$ where C is finite.[4] We will consider only measurable functions between Polish spaces, unless explicitly stated otherwise.

Under the usual convex combination operation, the set $\Delta(C)$ is a mixture space in the sense of Herstein and Milnor (1953). We describe mixtures of elements of C abstractly because they can be interpreted as either lotteries

---

[3]In a recent paper, Gilboa, Maccheroni, Marinacci, and Schmeidler (2008) argue for a separation between objective and subjective parts of a preference. Roughly the objective part is an incomplete preference that satisfies Bewley's axioms. Their approach is quite different from our approach which emphasizes the role of parameters.

[4]This assumption is convenient but not essential to most of our results.

over C or as frequencies. Our framework and main results will shed some light on how the two might be connected.

An *act* is any measurable function:

$$f : \Omega \to \Delta(C).$$

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 $x, y \in \Delta(C)$ 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 $\Delta(C)$.*

Next we introduce the usual monotonicity assumption (Savage (1954)'s P3):

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) $a(\omega) \succcurlyeq f^n(\omega), g^n(\omega), f(\omega), g(\omega) \succcurlyeq b(\omega), \forall \omega \in \Omega, n \in \mathbb{N}$; and (iii) $f^n \succcurlyeq g^n$ for all $n \in \mathbb{N}$, then $f \succcurlyeq g$.*[5]

Next we assume Herstein and Milnor (1953)'s linearity in mixtures:

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

---

[5]Our continuity assumption is similar to Ghirardato, Maccheroni, Marinacci, and Siniscalchi (2003)'s 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.

By Herstein and Milnor (1953)'s theorem (see details in the appendix), there is an affine function $u : \Delta(C) \to R$, unique up to positive affine transformations, such that for any pair of constant acts f, g, that take values $\ell_f, \ell_g \in \Delta(C)$ respectively,

$$f \succcurlyeq g \iff \int_{c \in C} u(c) \, d\, \ell_f \geq \int_{c \in C} u(c) \, d\, \ell_g.$$

We will sometimes use the convenient notation $u(\ell), \ell \in \Delta(C)$ 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* $x, y \in \Delta(C)$ *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 transformations Γ.[6] 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 generated by $\{\mathcal{E}_\gamma, \gamma \in \Gamma\}$.

- 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 $\gamma \in \Gamma$.

---

[6]A semigroup is a set Γ together with an operation "·" 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.

- 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 Parameterizations

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 $\Delta(\Omega)$ (see Definition A.21).[7] A standard definition of decomposition map with respect to a semi-group is given in Varadarajan (1963). Since some of our theorems refer to decompositions without explicit reference to a semi-group, we provide a more abstract definition below.[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 measures, and ϑ is Γ-invariant (i.e.,

---

[7]We will always assume that there is at least one Γ-ergodic measure.

[8]The definition is standard and essentially that of a sufficient statistic. See, for example, Billingsley (1995, p. 450), Varadarajan (1963) and Dynkin (1978).

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

$\vartheta(\gamma(\omega)) = \vartheta(\omega), \forall \gamma \in \Gamma$). *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 (1963)) 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:

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

---

[10]It is also essentially unique, in the sense of Lemma 4.4 in Varadarajan (1963). We comment further on this below. We cannot use directly Varadarajan's result because 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 the appendix, the Bayesian consensus is *proved* as a consequence of the existence of ϑ, which is directly established.

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 relabelings of the coordinates, and leads to the concept of exchangeability. Another example is the *shift transformation*:

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

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 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 the finite set C), the averages in (3)

---

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

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:

---

[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:*

1. *There is a decomposition map $\vartheta : \Omega \to \Theta$ such that for every pair of acts f and g*

$$f \succcurlyeq g \iff \int_\Omega f \, dP^{\vartheta(\cdot)} \succcurlyeq \int_\Omega g \, dP^{\vartheta(\cdot)}. \tag{4}$$

2. *ϑ is essentially unique: for every ϑ' satisfying (4), the set $\{\omega \in \Omega : \vartheta(\omega) \neq \vartheta'(\omega)\}$ is $\succcurlyeq$-null.*

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 $P(\cdot \,|\, \kappa)$ 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.

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:*

12

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',[13] $f^{\star}$ *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).

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.

The following theorem provides a tighter description of the set of parameters that takes into account additional restrictions:

Theorem 2 *Let* (Θ, ϑ) *be a sufficient parametrization of* $\succcurlyeq$*, and let* $\overline{\Theta} \subset \Theta$*. Then there exists a semi-group* Γ *of measurable transformations* $\gamma : \Omega \to \Omega$*, such that the following are equivalent:*

---

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

$$\left[\begin{array}{ll} f(\omega), & \text{if } \omega \in E \\ h(\omega), & \text{if } \omega \notin E \end{array} \right] \sim \left[\begin{array}{ll} g(\omega), & \text{if } \omega \in E \\ h(\omega), & \text{if } \omega \notin E \end{array} \right].$$

1. $f \sim f \circ \gamma$ *for all* $\gamma \in \Gamma$ *and* $f \in \mathcal{F}$;[14]

2. *There exists a map* $\overline{\vartheta} : \Omega \to \overline{\Theta}$ *such that* $(\overline{\Theta}, \overline{\vartheta})$ *is a sufficient parametrization of* $\succcurlyeq$;

3. $\vartheta^{-1}(\Theta \setminus \overline{\Theta})$ *is* $\succcurlyeq$-*null.*

Moreover, $\overline{\Theta}$ *is* Γ-*invariant.*

Condition 3 in the theorem captures the intuition of a *parametric restriction*: parameters outside $\overline{\Theta}$, although part of the parametrization implied by τ-invariance, are irrelevant for the preference. Theorem 2 makes the conceptual point that all parametric restrictions can be reduced to invariance conditions, not that the semi-group it constructs admits a simple description. There are situations where a specific semi-group of interest Γ naturally corresponds to a specific set of parameters (for instance, the set of permutations and i.i.d. probabilities). In such cases, Theorem 2 is not very helpful because the semi-group whose existence it asserts may bear little resemblance to Γ. The following theorem is more useful in this case:

Theorem 3 *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* $u(\Delta(C)) = \mathbb{R}$,[15] $\Theta_\Gamma \subset \Theta$, *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* $\succcurlyeq$.

Although our main focus is on decomposition maps that correspond to semi-groups, this is clearly not essential. Parametric preferences are well-defined and the equivalence of (2) and (3) in Theorem 2 holds even if we work

---

[14]Note that this does not require the full-invariance with respect to the semi-group in the sense of (3).

[15]Although the assumption $u(\Delta(C)) = \mathbb{R}$ rules out C finite, this is not a problem for this result. See comments in its proof.

directly with a decomposition map with respect to an abstract σ-algebras, since Definition 1 of decomposition maps allows for this case. The semi-group structure is used here only to characterize or restrict the set of parameters that are relevant for a preference.

3.3.1 Exchangeability

The classic example of invariance and its implications is de Finetti (1937)'s 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 3.

Corollary 1 *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. (Θ_Π, ϑ) is sufficient for $\succcurlyeq$.

Note that, as in Theorems 2 and 3, we do not require full-invariance with respect to the semi-group in the sense of (3). Epstein and Seo (2010) were the first to propose invariance with respect to the group of permutations Π: for every $f \in \mathcal{F}$, integer n, and $\pi_1, \ldots, \pi_n \in \Pi$

$$f \sim \frac{f \circ \pi_1 + \cdots + f \circ \pi_n}{n}. \tag{6}$$

They also studied weaker versions of this criterion of exchangeability and their implications under maxmin expected utility. See De Castro and Al-Najjar (2009) for a more detailed discussion of their work. In a paper subsequent to this one, Klibanoff, Mukerji, and Seo (2010) showed various interesting equivalent forms of this condition.

15

3.3.2 Markov Parameters

There is a large literature that studies weaker notions of exchangeability, usually referred to as "partial exchangeability." See Diaconis and Freedman (1984) for a general treatment.

An important case is when the set of parameters are Markov processes. This is a set of parameters that includes as special case the set of i.i.d. distributions. Diaconis and Freedman (1980) characterize the group of transformations $\mathcal{M}$ that gives rise to Markov parameters.

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 3.1 *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 3.1: 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 **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.[16] 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 $\succsim$-null set. In the Appendix, 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 risks because they are derived from the empirical frequencies, which are objective. Uncertainty, on the other hand, concerns how the decision maker aggregates over parameters. A Bayesian, for instance, would rank acts according to their expected utility with respect to a subjective prior over Θ, while a classical statistician may follow very different procedures.

- *Taste over Consequences vs. Invariance Judgment:* The preference $\succsim$ 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,

---

[16]See Jackson, Kalai, and Smorodinsky (1999) for a formal notion of learnability.

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 parameterization (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

Given our weak assumptions, the parametric preferences framework is consistent with a broad range of approaches for aggregating uncertainty about parameters. Here we provide a general methodology for doing so.

4.1 Parameter-Based Acts

A *parameter-based act* is any measurable function

$$F : \Theta \to \Delta(C).$$

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 prim-

itive 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 4.1 *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:*

$$f \succcurlyeq g \iff \Psi(f) \succcurlyeq \Psi(g). \tag{7}$$

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 (7).

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

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

Proposition 4.2 *Let (Θ, ϑ), $\succcurlyeq$ and $\succcurlyeq$ satisfy all the conditions of Proposition 4.1. Then, there is a functional $\mathcal{V}$ representing $\succcurlyeq$ if and only if there are functions $V : \tilde{\mathbb{F}} \to \mathbb{R}$ and $u : \Delta(C) \to \mathbb{R}$, where $\tilde{\mathbb{F}} \subset \mathbb{R}^\Theta$ and u is affine, such that*

$$f \succcurlyeq g \iff V\left(\theta \mapsto \int_\Omega u(f) \, dP^\theta\right) \geqslant V\left(\theta \mapsto \int_\Omega u(g) \, dP^\theta\right). \tag{8}$$

Proof: By Proposition 4.1 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 $\Delta(C)$ and satisfy the Herstein-Milnor axiom, there is an affine function $u : \Delta(C) \to \mathbb{R}$ representing the

19

preference (and the aggregator) over $\Delta(C)$. Therefore, we can write the functional $\mathcal{V}$ as $V \circ u$, where $V : \tilde{\mathbb{F}} \to \mathbb{R}$ and $\tilde{\mathbb{F}} = u(\mathbb{F}) \subset \mathbb{R}^\Theta$. Since u is affine, $u\left(\Psi(f)(\theta)\right) = \int_\Omega u(f)\, dP^\theta$, which establishes (8). The converse is trivial. ∎

4.1.1   Uncertainty Averse Aggregators

For concrete examples of aggregators, consider the class of uncertainty averse preferences characterized by Cerreia, Maccheroni, Marinacci, and Montrucchio (2008) (henceforth CMMM). This class is very broad and covers most models of uncertainty aversion in the literature. Here the aggregator characterized is

$$F \succcurlyeq G \iff \min_{\mu \in \Delta^\sigma(\bar{\mu})} \Phi\left(\int_\Theta u(F)\, d\mu, \mu\right) \geqslant \min_{\mu \in \Delta^\sigma(\bar{\mu})} \Phi\left(\int_\Theta u(G)\, d\mu, \mu\right) \quad (9)$$

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 $\bar{\mu}$, $u : \Delta(C) \to \mathbb{R}$ is an affine function with $u(\Delta(C)) = \mathbb{R}$, and $\Phi : \mathbb{R} \times \Delta(\Theta) \to (-\infty, \infty]$ is a function satisfying certain technical conditions; see CMMM for details.

Proposition 4.3 *Let (Θ, ϑ), $\succcurlyeq$ and $\succcurlyeq$ satisfy all the conditions of Proposition 4.1. The following statements are equivalent:*

(i) *$\succcurlyeq$ satisfies the axioms $A.1 - A.8$ of CMMM.*

(ii) *there exist an affine $u : \Delta(C) \to \mathbb{R}$, with $u(\Delta(C)) = \mathbb{R}$, a function[17] $\Phi : \mathbb{R} \times \Delta(\Theta) \to (-\infty, \infty]$ and $\bar{\mu} \in \Delta(\Theta)$ such that, for all $f, g \in \mathcal{F}$,*

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

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

Proof: By CMMM's Theorem 7, condition (i) holds if and only if (9) is true. Since u is affine, if $F = \Psi(f)$ then $u(F(\theta)) = \int_\Omega u(f)\, dP^\theta$. This concludes the proof. ∎

---

[17]Φ satisfy some technical conditions. See CMMM for details.

In this proposition, 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 $A : \tilde{\mathbb{F}} \to \mathbb{R}$ and affine $u : \Delta(C) \to \mathbb{R}$, $\tilde{\mathbb{F}} = u(\mathbb{F}) \subset \mathbb{R}^{\Theta}$, such that

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

Then the proposition would hold with the obvious adaptations.

4.2 Second-Order Expected Utility

We next consider in greater details a (parametric) second-order expected utility model. This may be viewed as generalizing the model introduced by Neilson((1993), (2009)). As noted in the introduction, this type of preferences gained considerable attention in the literature recently.[18]

As we shall see below, parameters are more integral to the second-order expected utility model than the other preference models discussed earlier. This makes it an especially attractive class to examine the implications of parametric preferences framework in details.

Definition 5 *We say that $\succcurlyeq$ has an* expected utility representation on parameter-based acts *if*

- *There exists a function $\varphi : \Delta(C) \to \mathbb{R}$ and a countably additive probability measure ν on Θ such that for any $F, G \in \mathbb{F}$*

$$F \succcurlyeq G \iff \int_{\Theta} \varphi(F) \, d\nu \geq \int_{\Theta} \varphi(G) \, d\nu \tag{10}$$

  and:

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

---

[18]Related models include, among others, Nau (2001), Klibanoff, Marinacci, and Mukerji (2005), Nau (2006), Ergin and Gul (2009), Chew and Sagi (2008), Strzalecki (2010), Grant, Polak, and Strzalecki (2009), and Klibanoff, Mukerji, and Seo (2010).

Conditions characterizing the existence of an expected utility representation amount to imposing the appropriate version of the Savage axioms on $\mathbb{F}$.

Theorem 4 *Let* (Θ, ϑ), $\succcurlyeq$ *and* $\succcurlyeq$ *satisfy all the conditions of Proposition 4.1. The following statements are equivalent:*

1. $\succcurlyeq$ *has a subjective expected utility representation;*

2. *There is a probability measure* μ *on* Θ, *and a function* $\phi : \mathbb{R} \to \mathbb{R}$ *such that, for any pair of acts* f, g:

$$f \succcurlyeq g \iff \int_\Theta \phi \left(\int_\Omega u(f) \, dP^\theta \right) d\mu \geq \int_\Theta \phi \left(\int_\Omega u(g) \, dP^\theta \right) d\mu. \quad (11)$$

If a preference can be represented as in (11), then μ *is unique, and the restriction of the function* ϕ *to* $u(\Delta(C))$ *is unique up to positive affine transformations.*

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

Proof: Assume (1) in the statement of the theorem and let (φ, μ) be as in the definition so (10) holds. Since φ and u both represent the same preference on $\Delta(C)$, there must be a monotone increasing function $\phi : u(\Delta(C)) \to \mathbb{R}$ such that $\varphi(c) = \phi(u(c))$ for every consequence c.

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:

$$f \succcurlyeq g \iff \int_\Omega f \, dP^{\vartheta(\cdot)} \succcurlyeq \int_\Omega g \, dP^{\vartheta(\cdot)} \quad (12)$$

$$\iff F \succcurlyeq G \quad (13)$$

$$\iff \int_\Theta \varphi \left(F \right) d\mu \geq \int_\Theta \varphi \left(G \right) d\mu \quad (14)$$

$$\iff \int_\Theta \phi \circ u \left(\int_\Omega f \, dP^\theta \right) d\mu \geq \int_\Theta \phi \circ u \left(\int_\Omega g \, dP^\theta \right) d\mu \quad (15)$$

$$\iff \int_\Theta \phi \left(\int_\Omega u(f) \, dP^\theta \right) d\mu \geq \int_\Theta \phi \left(\int_\Omega u(g) \, dP^\theta \right) d\mu \quad (16)$$

In the above: (12) follows from the definition of a sufficient parameterization; (13) follows from Proposition 4.1; (14) follows from condition (10) and the fact that the acts in (13) are parameter-based; (15) follows from $\varphi(c) = \phi(u(c))$ for every c; and finally (16) 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, assume that (2) in the theorem holds. Then, we can repeat the above arguments to get the equivalence of (16) and (14). Now, by Proposition 4.1, we get the equivalence (14) and (13), which means that $\succcurlyeq$ has a subjective expected utility representation.

Finally, we assume (11) and show that $\succcurlyeq$ is Γ-invariant. Fix $\gamma_1, \ldots, \gamma_n \in \Gamma$ and act f. Then:

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

$$
=\; \int \varphi \left(\frac{1}{n} \sum_{j=0}^{n-1} \int f \, dP^\theta \right) d\mu
$$

$$
=\; \int \varphi \left(\int f \, dP^\theta \right) d\mu.
$$

This concludes the proof. ∎

The reader may wonder why we need to offer a different proof, since Proposition 4.2 and especially Proposition 4.3 seem to contain, as a special case, expected utility aggregators. There is a subtle difference: in the two propositions, the aggregator is already represented by an *affine* u, while in Theorem 4 φ need not be linear. Note that if φ is linear, then the model above collapses to a standard expected utility preference.

One way to think of a parametrization is as a process in which the decision maker partitions the state space into events $\{\vartheta^{-1}(\theta)\}_\Theta$ within which variability in consequences is treated as objective risk, in the sense that he applies the same risk attitude given by u that he applies to objective lotteries. To provide 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

23

$\succcurlyeq$ 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 now the act f that pays 1 dollar if the first toss is H and 0 otherwise. Then for the first decision maker, the parameter-based act corresponding to f yields utility $\delta_\omega \mapsto u\big(f(\omega)\big)$ while for the second, it is $\theta \mapsto \theta u(1) + (1 - \theta)u(0)$, where we take θ to be the probability of H. The overall value of the act under the representation (11) 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 former is obtained by applying a mean preserving spread to the latter, so it is smaller. The decision maker with finer parameterization perceives less risk and more uncertainty than the one whose parametrization is coarser.

4.3 Separating Objective and Subjective Uncertainty

The double integral representation in Theorem 4 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((1993), (2009)) and the related work by Strzalecki (2010). Neilson considers the representation:

$$V_N(f) = \int \phi\left[u(f(\omega))\right] d\mu(\omega). \tag{17}$$

The next proposition is a simple consequence of Theorem 4 and Proposition 3.1 characterizing Neilson's representation in terms of the trivial semi-group $\{\mathbf{1}\}$:

Proposition 4.4 *For every preference relation $\succcurlyeq$ the following are equivalent:*

1. *The preference $\succcurlyeq$ 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;*

24

2. *There is a probability measure μ on Θ, and a function $\phi : \mathbb{R} \to \mathbb{R}$ such that $\succcurlyeq$ can be represented by the functional V_N.*

The uniqueness properties of μ, ϕ hold as in Theorem 4.

In an important recent paper Strzalecki (2010) introduces additional structure that ensures that ϕ has the specific functional form corresponding to multiplier preferences:

$$\phi_\alpha(r) = \begin{cases} -\exp\left(-\frac{r}{\alpha}\right) & \alpha < \infty, \\ r & \alpha = \infty. \end{cases}$$

Here α is interpreted as a measure of uncertainty aversion: when $\alpha = \infty$ the decision maker is expected utility maximizer, while for finite α he treats subjective uncertainty about the state ω differently from the way he treats objective lotteries over consequences.

Neilson((1993), (2009)) and Strzalecki (2010) implicitly assume that the decision maker has no additional (subjective) structure on the decision problem in the form of non-trivial invariance. In the context of a sequence of coin tosses, the decision maker may believe that there are common factors that connect consecutive tosses, justifying an assumption of exchangeability, for instance. In applications of multiplier preferences to dynamic macro and finance models, it is common to assume that the decision maker has a parametric representation of the uncertainty he faces—for instance, that the time series of interest are governed by some Markovian process P^θ of unknown parameter θ. In this case, it seems plausible that, conditional on knowledge that the parameter is θ, the decision maker treats P^θ as objective risk.

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 we can extend Neilson((1993), (2009)) and Strzalecki (2010) models to accommodate the parameters as an additional source of objective uncertainty, yielding the functional form: $\int \phi \left(\int u(f) \, dP^\theta \right) d\mu(\theta)$ derived in Theorem 4.

25

4.4 "Second-Order" Acts

Klibanoff, Marinacci, and Mukerji (2005) provide a model with similar representation which, in our notation, has the form:

$$\int_{\Delta(\Omega)} \phi \left(\int_{\Omega} u(f(\omega)) \, dP(\omega) \right) d\nu(P). \tag{18}$$

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, Marinacci, and Mukerji 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:

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

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 (18) have a behavioral meaning in our model. Here we briefly highlight the main issues. A detailed discussion is in Al-Najjar and De Castro (2010).

The second-order expected utility model (11) differs from (18) along two fundamental dimensions:

- *Foundations:* The behavioral foundations underlying (18) require decision makers to express a preference $\breve{\succcurlyeq}$ over second-order acts.

- *Functional Form:* In (18) the outer integral is over mixtures of parameters so decision makers have beliefs about randomizations over parameters.

The two issues above are closely related. We illustrate this with a simple example and refer the reader to Al-Najjar and De Castro (2010) for a more comprehensive argument. 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 $\mathbf{f} : [0,1] \to \mathbb{R}$. Here, $\mathbf{f}(0)$ and $\mathbf{f}(1)$ denote the consequences obtained if θ_0 or θ_1 occur, respectively, while $\mathbf{f}(\alpha)$ is the consequence when the 'true' distribution is the mixture $P_\alpha \equiv \alpha P^{\theta_0} + (1-\alpha) P^{\theta_1}$.

In any parametric preference, regardless of functional form, a bet on a parameter θ_i is just a compact way of expressing a bet on the event $\vartheta^{-1}(\theta_i) \subset \Omega$. Parameter-based acts are nothing more than ordinary acts, compactly expressed using parameters. For $0 < \alpha < 1$, statements like "P_α occurred" has no behavioral meaning. Even in highly idealized thought experiments, all that is observed is the state ω, from which one can infer whether the parameter is either θ_0 or θ_1, in our example. The mixture P_α, on the other hand, is not observed, even in the limit with infinite amount of data. As a result, a second-order act $\mathbf{f}$ cannot be executed and a preference $\breve{\succsim}$ over such acts is not meaningful in our model. The functional form (18) reflects this: the only way to identify the belief ν is by asking the decision maker to rank second-order acts.

Any preference that incorporates second-order acts is inconsistent not just with the second-order expected utility model (11) but all parametric preferences covered in this paper, regardless of functional form.

A   Preliminary results

A.1 $\succsim$-Null sets

Lemma A.1 *Let C_n be $\succsim$-null for all $n \in \mathbb{N}$. Then, $C = \cup_{n \in \mathbb{N}} C_n$ is $\succsim$-null.*

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

$$
\begin{aligned}
f^N &= f1_{C_1} + f1_{C_2} + \cdots + f1_{C_N} + x1_{C \backslash A^N} + h1_{C^c} \\
&\sim g1_{C_1} + f1_{C_2} + f1_{C_3} + \cdots + f1_{C_N} + x1_{C \backslash A^N} + h1_{C^c} \\
&\sim g1_{C_1} + g1_{C_2} + f1_{C_3} + \cdots + f1_{C_N} + x1_{C \backslash A^N} + h1_{C^c} \\
&\quad \cdots \\
&\sim g1_{C_1} + g1_{C_2} + \cdots + g1_{C_N} + x1_{C \backslash A^N} + h1_{C^c} \\
&= g^N.
\end{aligned}
$$

It is easy to see that $f^N \to f1_C + h1_{C^c}$ and $g^N \to g1_C + h1_{C^c}$. Since $f, g, h \in \mathcal{F}$, there exist $\overline{x}_a, \underline{x}_a$ such that $\overline{x}_a \succsim a(\omega) \succsim \underline{x}_a, \forall \omega \in \Omega$, for $a = f, g, h$. Define $u(\omega) \equiv \max\{\overline{x}_f, \overline{x}_g, \overline{x}_h\}$ and $l(\omega) \equiv \min\{\underline{x}_f, \underline{x}_g, \underline{x}_h\}$. Note that these values are well-defined because $\succsim$ is complete on $\Delta(C)$. Therefore, continuity implies $f1_C + h1_{C^c} \sim g1_C + h1_{C^c}$. Since f, g, h are arbitrary, C is null. ∎

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

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

$$
\left[\begin{array}{ll} f'(\omega), & \text{if } \omega \in E \\ h(\omega), & \text{if } \omega \notin E \end{array} \right] \sim \left[\begin{array}{ll} g'(\omega), & \text{if } \omega \in E \\ h(\omega), & \text{if } \omega \notin E \end{array} \right].
$$

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

$$
\left[\begin{array}{ll} f(\omega), & \text{if } \omega \in A \\ h(\omega), & \text{if } \omega \notin A \end{array} \right] \text{ and } \left[\begin{array}{ll} g(\omega), & \text{if } \omega \in A \\ h(\omega), & \text{if } \omega \notin A \end{array} \right].
$$

Therefore, A is $\succsim$-null. ∎

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

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

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

Lemma A.4 *$\overline{\mathcal{H}}$ is a σ-field cointaining $\mathcal{H}$. 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 $\{A_n\}_{n\in\mathbb{N}} \subset \overline{\mathcal{H}}$. Then there exist $\{B_n\}_{n\in\mathbb{N}} \subset \mathcal{H}$, such that $C_n \equiv A_n \Delta B_n \in \mathcal{N}$. Let $A \equiv \cup_{n\in\mathbb{N}} A_n$ and $B = \cup_{n\in\mathbb{N}} B_n$. It is clear that $B \in \mathcal{H}$ and

$$
\begin{aligned}
A\Delta B &= [(\cup_{n\in\mathbb{N}} A_n) \cap (\cup_{n\in\mathbb{N}} B_n)^c] \cup [(\cup_{n\in\mathbb{N}} A_n)^c \cap (\cup_{n\in\mathbb{N}} B_n)] \\
&= [(\cup_{n\in\mathbb{N}} A_n) \cap (\cap_{n\in\mathbb{N}} B_n^c)] \cup [(\cap_{n\in\mathbb{N}} A_n^c) \cap (\cup_{n\in\mathbb{N}} B_n)] \\
&\subset [\cup_{n\in\mathbb{N}} (A_n \cap B_n^c)] \cup [\cup_{n\in\mathbb{N}} (A_n^c \cap B_n)] \\
&= \cup_{n\in\mathbb{N}} C_n.
\end{aligned}
$$

The set $\cup_{n\in\mathbb{N}} C_n$ is $\succsim$-null by Lemma A.1. Since $A\Delta B$ is Σ-measurable and is contained in the $\succsim$-null set $\cup_{n\in\mathbb{N}} C_n$, 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, $C = A \setminus B \subset A\Delta B$ is Σ-measurable and therefore, $\succsim$-null. But $A = B \cup C \in \mathcal{H} \cup \mathcal{N}$ and therefore, $A \in \mathcal{H} \vee \mathcal{N}$. ∎

A.2 Reduction to real-valued functions

Under our assumptions, Herstein and Milnor (1953)'s theorem implies the existence of a linear function $u : \Delta(C) \to \mathbb{R}$, unique up to affine transformations, representing $\succsim$ on $\Delta(C)$. By linear, we mean that $u(\alpha x + (1-\alpha)y) = \alpha u(x) + (1-\alpha)u(y)$, for every $x, y \in \Delta(C)$ and $\alpha \in [0,1]$. Moreover, since u

is linear, $I \equiv u(\Delta(C)) \subset \mathbb{R}$ is a convex subset of $\mathbb{R}$, 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:

Lemma A.5 $u : \Delta(C) \to \mathbb{R}$ *is continuous.*[19]

Proof: Let $x^n, x \in \Delta(C), \forall n \in \mathbb{N}$, $x^n \to x$ and assume that there exists $\delta > 0$ such that for any $m \in \mathbb{N}$, there exists $n_m > m$ such that $u(x^{n_m}) \notin (u(x) - \delta, u(x) + \delta)$. Define $A = \{j \in \mathbb{N} : x^{n_j} \succcurlyeq x\}$ and $B = \{j \in \mathbb{N} : x \succcurlyeq x^{n_j}\}$. Since $\succcurlyeq$ is complete on $\Delta(C)$, $A \cup B = \mathbb{N}$ and at least one of the two sets is infinite. Without loss of generality, assume that A is infinite. Then, for any $j \in A$, $u(x^{n_j}) > u(x) + \delta$. Since $u(\Delta(C))$ is convex, there exists y such that $u(x) < u(y) < u(x) + \delta < u(x^{n_j})$. Consider the sequence $y^j = y, \forall j \in \mathbb{N}$. Then, $x^{n_j} \succcurlyeq y^j$, $x^{n_j} \to x$ and $y^j \to y$. Pointwise continuity implies that $x \succcurlyeq y$, which contradicts $u(y) > u(x)$. ∎

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

$$f \succcurlyeq^{\mathcal{D}} g \equiv u^{-1}(f) \succcurlyeq u^{-1}(g). \tag{19}$$

In fact, (19) is not completely formal, since u is not invertible in general and, therefore, $u^{-1} \circ f : \Omega \to \Delta(C)$ 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 (19). 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 $u : \Delta(C) \to I$ we can define a preference $\succcurlyeq'$ on $\mathcal{F}$ by the following: for any $f, g \in \mathcal{F}$,

$$f \succcurlyeq' g \equiv u(f) \succcurlyeq^{\mathcal{D}} u(g). \tag{20}$$

_____

[19]This lemma holds even if C is just a Polish space. If C is finite, $\Delta(C)$ is finite dimensional and continuity follows immediately from the linearity of u. Thus, this lemma is relevant only to infinite dimensional spaces.

It is easy to see that if we start with $\succcurlyeq$ on $\mathcal{F}$, obtain $\succcurlyeq^{\mathcal{D}}$ on $\mathcal{D}$ by (19) and use this $\succcurlyeq^{\mathcal{D}}$ together with u in (20) to define a preference $\succcurlyeq'$, then $\succcurlyeq$ and $\succcurlyeq'$ coincide.[20]

In sum, a preference $\succcurlyeq$ on $\Delta(C)$-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* $u : \Delta(C) \to \mathbb{R}$ defines a preference $\succcurlyeq$ on $\mathcal{F}$. The next proposition establishes a useful link between the two:

Lemma A.6 *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 (19).*

2. *It is given a preference $\succcurlyeq^{\mathcal{D}}$ on $\mathcal{D}$ and a linear $u : \Delta(C) \to \mathbb{R}$, let $\succcurlyeq$ be defined by (20).*

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

Proof: Fix an act $f : \Omega \to \Delta(C)$. Since $f \in \mathcal{F}$ is bounded, there exist $\underline{x}, \overline{x}$ such that $\overline{x} \succcurlyeq f(\omega) \succcurlyeq \underline{x}$, for all $\omega \in \Omega$. Herstein and Milnor (1953) also show that for any z satisfying $\overline{x} \succcurlyeq z \succcurlyeq \underline{x}$, there exists a unique $\alpha \in [0, 1]$ such that $z \sim \alpha \overline{x} + (1 - \alpha)\underline{x}$. Therefore, $u(f(\Omega)) \subset u([\underline{x}, \overline{x}])$, where $[\underline{x}, \overline{x}] \equiv \{\alpha \overline{x} + (1 - \alpha)\underline{x} : \alpha \in [0, 1]\}$, and the function u is invertible when restricted to $[\underline{x}, \overline{x}]$; in this proof, u^{-1} will denote the inverse function of this restriction.

Since u is linear in $\Delta(C)$, then for every $\omega \in \Omega$,

$$u\left(\frac{f \circ \gamma_1 + \cdots + f \circ \gamma_n}{n}(\omega)\right) = \frac{u \circ f \circ \gamma_1(\omega) + \cdots + u \circ f \circ \gamma_n(\omega)}{n}. \quad (21)$$

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

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

---

[20] Given $f, g \in \mathcal{F}$, (19) implies that $u(f) \succcurlyeq^{\mathcal{D}} u(g) \Leftrightarrow f \succcurlyeq g$, from the discussion following (19). However,

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:

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

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

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

$$\tilde{f} \sim \frac{\tilde{f} \circ \gamma_1 + \cdots + \tilde{f} \circ \gamma_n}{n}. \tag{22}$$

Fix $f \in \mathcal{F}$. From (20),

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

Using (21) and (22), 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 $I = u(\Delta(C))$. Since u is affine, $u(\Delta(C))$ will be an interval $I \subset \mathbb{R}$ (which may be the whole $\mathbb{R}$). We can calibrate u so that the two outcomes $x, y \in \Delta(C)$ assumed to exist in Assumption 5, have values 0 and 1, respectively. In particular, this implies that the interval $[0, 1] \subset I \subset \mathbb{R}$ 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.1 $\succcurlyeq^{\mathcal{D}}$ is defined for functions $f : \Omega \to I \subset \mathbb{R}$ 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 $f \succcurlyeq^{\mathcal{D}} g$.

32

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) $|f^n(\omega)| \leq b(\omega)$ and $|g^n(\omega)| \leq b(\omega)$, 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 $f : \Omega \to I \subset \mathbb{R}$. 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 (19) using the u related to $\succcurlyeq$, and now consider $\succcurlyeq^{\mathcal{D}}$ with another linear function $u' : \Delta(C) \to \mathbb{R}$. If we use u' and $\succcurlyeq^{\mathcal{D}}$ as in (20), 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 (19), it is enough to show that $\succcurlyeq$ and $\succcurlyeq^{\mathcal{D}}$ have the same null sets. Let $f, g, h \in \mathcal{F}$. Then, (19) and (20) imply that:

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

∎

A.3 Statement of subjective ergodic theorems

The proof of Theorem 1 is based on two theorems about preferences defined for real-valued functions, whose statements and proofs are given below.

Consider a preference $\succeq$ defined on the set $\mathcal{D}$ of all Σ-measurable functions $f : \Omega \to I \subset \mathbb{R}$, satisfying all the assumptions listed in Corollary A.1. In this section, we will simplify notation by writing $\succeq$ instead of $\succeq^{\mathcal{D}}$. No confusion should arise since we consider no other preference here.

Theorem A.8 (The Subjective Ergodic Theorem) *The following conditions are equivalent:*

1. *$\succeq$ is τ-invariant.*

2. *For every act f, the empirical limit $f^\star$ is well-defined off a $\succeq$-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 $\succeq$ is τ-ergodic, then $f^\star$ is constant except in a $\succeq$-null set.

This theorem was proved in De Castro and Al-Najjar (2009). Since the proof is long and technical, we include it as supplemental material.

For stating the next theorem, we need some notation. Let $\Delta(\Omega)$ is the set of all probability measures in Ω, endowed with its usual weak*-topology. Let $\mathcal{P}_\tau^{er} \subset \Delta(\Omega)$ denote the set of all 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 *Assume that Ω is compact. If $\succeq$ is τ-invariant, then there exists a decomposition map $\vartheta : \Omega \to \Theta$ such that (Θ, ϑ) is sufficient for $\succeq$.*

---

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

A.4 Proof of Theorem A.9

Consider the sup-norm in $\mathcal{D}$:

$$\|f\| = \sup_{\omega \in \Omega} |f(\omega)|.$$

A subset $\tilde{\mathcal{D}} \subset \mathcal{D}$ is separable if there exists a countable dense subset, that is, a countable set $H \subset \tilde{\mathcal{D}}$ such that for every $f \in \tilde{F}$ and $\epsilon > 0$, there exists $h \in H$ such that $\|f - h\| < \epsilon$.

> **Comments:** Our first difficulty in proving Theorem A.9 is that the set $\mathcal{D}$ is not separable in general. This creates a number of difficulties. Therefore, we need to find a suitable subset $\tilde{\mathcal{D}} \subset \mathcal{D}$ which is separable *and* has sufficiently nice property as to imply things about $\mathcal{D}$.

We need the following:

Lemma A.10 *Let S be a Polish space and $\mathcal{S}$, its Borel σ-algebra. Then, there exists a countable algebra $\mathcal{S}^\circ$ of subsets of S that generates $\mathcal{S}$.*

Proof: Since S is Polish, by (Royden 1968, Theorem 8, p. 326) it is Borel isomorphic to (i) $[0,1]$; (ii) $\mathbb{N}$; or (iii) a finite set. Sets A and B are Borel isomorphic if there is a measurable bijective map $h : A \to B$, with measurable inverse. Consider first the case where S is Borel isomorphic to $[0,1]$. Let $\mathcal{A}$ denote the collection of finite unions of intervals of the form $[a, b)$, for $a, b \in \mathbb{Q} \cap [0,1]$. It is easy to see that $\mathcal{A}$ is a countable algebra that generates the Borel field of $[0,1]$. Since intersections and set difference is preserved under the inverse of a function, then $\mathcal{S}^\circ \equiv h^{-1}(\mathcal{C})$ is also a countable algebra. Since $h^{-1}(\sigma(\mathcal{C})) = \sigma(h^{-1}(\mathcal{C}))$, then $\mathcal{S}^\circ$ generates the Borel field $\mathcal{S}$.

If S is Borel isomorphic to $\mathbb{N}$, take as $\mathcal{A}$ the algebra of all singletons of $\mathbb{N}$, which generates its Borel field. Then, repeat the ideas above. Finally, if S is Borel isomorphic to a finite set, take as $\mathcal{A}$ the power set of this finite set and repeat the same arguments. ∎

Lemma A.11 *Let $\mathcal{S}^\circ$ as in the above lemma. Then $B(S, \mathcal{S}^\circ)$ is separable.*

35

Proof: The set $B(S, \mathcal{S}^\circ)$ is the closed linear span of a countable set, namely, $\cup_{n \in \mathbb{N}} \{\sum_{i=1}^{n} \alpha_i 1_{A_i} : \alpha_i \in \mathbb{Q}; A_i \in \mathcal{S}^\circ, \text{ for } i = 1, 2, ..., n\}$. Therefore, it is separable. ∎

Let Σ° be a countable algebra that generates Σ, which exists by Lemma A.10. Thus the set $B(\Omega, \Sigma^\circ) \subset \mathcal{D}$ is clearly separable.

Lemma A.12 *Let* $\tilde{\mathcal{D}} \subset \mathcal{D}$ *be separable. Then, there exists* $\Omega' \subset \Omega$ *with* $\succcurlyeq$-*null complement, such that the limit*

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

exists for all $f \in \tilde{\mathcal{D}}$ *and* $\omega \in \Omega'$.

Proof: Let $H = \{h_n : n \in \mathbb{N}\}$ be a countable dense set of $\tilde{\mathcal{D}}$. By Theorem A.8, there exist a set Ω_n such that $\Omega \setminus \Omega_n$ is $\succcurlyeq$-null, and the limit

$$\lim_{N \to \infty} \frac{1}{N} \sum_{j=0}^{N-1} h_n\left(\tau^j \omega\right)$$

exists for all $\omega \in \Omega^n$. Define $\Omega' \equiv \cap_n \Omega^n$ and, for each $f \in \tilde{\mathcal{D}}$ and $N \in \mathbb{N}$, $f^N(\omega) \equiv \frac{1}{N} \sum_{j=0}^{N-1} f\left(\tau^j \omega\right)$. By Lemma A.1, $\Omega \setminus \Omega'$ is $\succcurlyeq$-null. Therefore, it is sufficient to show that for all $\omega \in \Omega'$ there exists $\lim_{N \to \infty} f^N(\omega)$ or, equivalently, that $\{f^N(\omega)\}$ is Cauchy. Given $\varepsilon > 0$, choose h_n such that $\|f - h_n\| < \frac{\varepsilon}{3}$, which can be done because $\{h_n\}$ is dense. Now, choose n_ϵ such that $N, M > n_\epsilon$ implies $|h_n^N(\omega) - h_n^M(\omega)| < \frac{\varepsilon}{3}$. Therefore:

$$
\begin{aligned}
|f^N(\omega) - f^M(\omega)| \;\leqslant\; & |f^N(\omega) - h_n^N(\omega)| + |h_n^N(\omega) - h_n^M(\omega)| \\
& + |h_n^M(\omega) - f^M(\omega)| \\
\leqslant\; & \frac{1}{N} \sum_{j=0}^{N-1} |f\left(\tau^j \omega\right) - h_n(T^j \omega)| + \frac{\epsilon}{3} \\
& + \frac{1}{M} \sum_{j=0}^{M-1} |f\left(\tau^j \omega\right) - h_n(T^j \omega)| \\
\leqslant\; & \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon,
\end{aligned}
$$

36

that is, $\{f^N(\omega)\}$ is Cauchy for all $\omega \in \Omega'$. This shows that $\lim_{N \to \infty} f^N(\omega)$ exists for all $\omega \in \Omega'$ and $f \in \tilde{\mathcal{D}}$. ∎

For each $A \in \Sigma$ and $\omega \in \Omega$, define $I(\omega, A) = 1$ if $\omega \in A$ and 0 otherwise. Occasionally, it will be more convenient to write $I(\omega, A)$ as $I_A(\omega)$ (to see it as a function of ω for some fixed A). We will also be interested in $I_\omega(A)$, to see it as a function of A, for a fixed ω. No confusion should arise from this abuse of notation. The following comes directly from Lemmas A.10, A.12 and the Subjective Ergodic Theorem (Theorem A.8):

Corollary A.2 *There exists Ω' such that for all $\omega \in \Omega'$ and all $A \in \Sigma^\circ$, the limit $I_A^\star(\omega) \equiv \lim_{N \to \infty} \frac{1}{N} \sum_{j=0}^{n-1} I_A(\tau^j \omega)$ exists. Moreover, $I_A \sim I_A^\star$.*

Let the set Ω' given by the lemma above be fixed from now on.

Definition A.13 *Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be measurable spaces. A Markov kernel (or a stochastic kernel) is a mapping $k : X \times \mathcal{Y} \to [0, 1]$ satisfying the following two properties.*

1. *For each $x \in X$, the set function $k(x, \cdot) : \mathcal{Y} \to [0, 1]$ is a (countably additive) probability measure.*

2. *For each $A \in \mathcal{Y}$, the mapping $k(\cdot, A) : X \to [0, 1]$ is $\mathcal{X}$-measurable.*

Of course, $I^\star(\omega, A)$ is well-defined only for $\omega \in \Omega'$ and $A \in \Sigma^\circ$. Our objective will be to extend it to $I^\star : \Omega \times \Sigma \to [0, 1]$ and show that it is a Markov kernel. This will be established through a series of lemmas.

Lemma A.14 *For each $\omega \in \Omega'$, $I_\omega^\star : \Sigma^\circ \to [0, 1]$ is a finitely additive and monotonic set function.*

Proof: If $A_1, ... A_n$ are disjoint events in Σ°, then $I_{\cup_{i=1}^n A_i}^\star(\omega)$ is equal to:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{j=0}^{N-1} I_{\cup_{i=1}^n A_i}(\tau^j \omega) = \lim_{N \to \infty} \frac{1}{N} \sum_{j=0}^{N-1} \sum_{i=1}^n I_{A_i}(\tau^j \omega) = \sum_{i=1}^n I_{A_i}^\star(\omega), \quad (23)$$

37

for all $\omega \in \Omega'$. This shows that the limit $I_A^\star(\omega)$ exists and it is unique for all $\omega \in \Omega'$ and all $A = \cup_{i=1}^n A_i$, for $A_1, ... A_n$ disjoint events in Σ°, that is, for all $A \in \Sigma^\circ$. For each $A \in \Sigma^\circ$, define $\nu(A) \equiv I_A^\star(\omega)$ (for any $\omega \in \Omega'$). If $A, B \in \Sigma^\circ$ and $A \subset B$, then $I_A^\star(\omega) \leqslant I_B^\star(\omega)$, that is, ν is monotone. (23) shows that ν is finitely additive. $\blacksquare$

Lemma A.15 *For each $\omega \in \Omega'$, $I_\omega^\star : \Sigma^\circ \to [0, 1]$ is countably additive.*

Proof: Consider a decreasing sequence of sets $A_n \in \Sigma^\circ$, $A_n \downarrow \emptyset$. By (Billingsley 1995), Example 2.10, p. 25, it is sufficient to prove that $I_\omega^\star(A_n) \to I_\omega^\star(\emptyset) = 0$. Suppose otherwise. Then there exist $\epsilon > 0$ and a subsequence A_{n_j} such that $I_\omega^\star(A_{n_j}) \geqslant \epsilon$, which means that $I_{A_{n_j}}^\star \succcurlyeq \epsilon$. It is clear that $I_{A_{n_j}}^\star$ converges to 0 pointwise and, therefore, $I_{A_{n_j}}^\star \to 0$. Pointwise continuity implies that $0 \succcurlyeq \epsilon$, but this contradicts the non-triviality assumption. $\blacksquare$

Lemma A.16 *For each $\omega \in \Omega'$, there exists a unique extension $\nu_\omega : \Sigma \to [0, 1]$ of $I_\omega^\star : \Sigma^\circ \to [0, 1]$.*

Proof: By the Caratheodory extension theorem (see (Royden 1968, Theorem 8, p. 257)), the following outer measure is the unique extension of $I_\omega^\star$ to all $A \in \Sigma$:

$$\nu_\omega(A) \equiv \inf \left\{ \sum_{n=1}^\infty I_\omega^\star(A_n) : A_n \in \Sigma^\circ, A \subset \cup_{n=1}^\infty A_n \right\}. \tag{24}$$

$\blacksquare$

Recall that for each $\omega \in \Omega'$ and $A \in \Sigma$, $I_\omega^\star(A)$ exists. The above result established that $I_\omega^\star(A) = \nu_\omega(A)$ only for $A \in \Sigma^\circ$. However, $\nu_\omega(\cdot)$ and $I_\omega^\star(\cdot)$ actually agree for more general sets, as the following lemma establishes.

Lemma A.17 *For any $A \in \Sigma$ and $\omega \in \Omega'$, $I_\omega^\star(A) = \nu_\omega(A)$.*

Proof: Take $A_i \in \Sigma^\circ$, such that $A \subset \cup_{i=1}^n A_i$. It is clear that $I_A(\omega) \leqslant \sum_{i=1}^n I_{A_i}(\omega)$ and, therefore,

$$\bar{I}_A(\omega) \equiv \limsup_{N \to \infty} \frac{1}{N} \sum_{j=0}^{N-1} I_A(\tau^j \omega) \leqslant \sum_{i=1}^n I_{A_i}^\star(\omega) = \sum_{i=1}^n I_\omega^\star(A_i).$$

38

Since this is valid for any n, (24) implies that $\bar{I}_A(\omega) \leqslant \nu_\omega(A)$. Similarly,

$$\underline{I}_A(\omega) \equiv \liminf_{N \to \infty} \frac{1}{N} \sum_{j=0}^{N-1} I_A(\tau^j \omega) \leqslant \sum_{i=1}^{n} I^\star_{A_i}(\omega) = \sum_{i=1}^{n} I^\star_\omega(A_i),$$

which proves that $\underline{I}_A(\omega) \leqslant \nu_\omega(A)$. It is easy to see that $\bar{I}_A(\omega) = 1 - \underline{I}_{A^c}(\omega)$. Assume that for some A and $\omega \in \Omega'$, we have $\bar{I}_A(\omega) < \nu_\omega(A)$. Then, $\underline{I}_{A^c}(\omega) = 1 - \bar{I}_A(\omega) > 1 - \nu_\omega(A) = \nu_\omega(A^c)$, but this contradicts $\underline{I}_{A^c}(\omega) \leqslant I^\star_\omega(A^c)$. This shows that $\bar{I}_A(\omega) = \nu_\omega(A)$. Similarly, $\underline{I}_A(\omega) = \nu_\omega(A)$, for all $\omega \in \Omega'$, which shows that the limit $I^\star_A(\omega)$ exists and it is equal to $\nu_\omega(A)$ for all $\omega \in \Omega'$. ∎

Fix the $\nu : \Omega' \times \Sigma \to [0, 1]$ defined above.

Corollary A.3 *For any $\omega \in \Omega'$, the set function $\nu_\omega : \Sigma \to [0, 1]$ is τ-invariant and τ-ergodic.*

Proof: We want to prove that $\nu_\omega(\tau^{-1}(A)) = \nu_\omega(A)$ for every $A \in \Sigma$. This comes directly from Lemma A.17 since for every $\omega \in \Omega'$ and $A \in \Sigma$, $I^\star_{\tau^{-1}(A)}(\omega) = I^\star_A \circ \tau(\omega) = I^\star_A(\omega)$, where the last inequality holds because $I^\star_A$ is τ-invariant, by Theorem A.8.

Now let $A \in \mathcal{E}_\tau$, that is, $\tau^{-1}(A) = A$. Then $I_A(\omega) = I_A(\tau^j \omega)$ for every j. Therefore, $I^\star_A(\omega) = 1$ if $\omega \in A$ and 0 otherwise. In other words, ν_ω assumes only the values 0 or 1 for any invariant A, that is, ν_ω is τ-ergodic. ∎

Lemma A.18 *Fix a Σ-measurable bounded $f : \Omega \to \mathbb{R}$ and let $g : \Omega \to \mathbb{R}$ be a Σ-measurable function which satisfies $g(\omega) = \int f(\tilde{\omega}) \nu(\omega, d\tilde{\omega}), \forall \omega \in \Omega'$. Then, $f \sim g$.*

Proof: Let $A_1, ..., A_n \in \Sigma$. By Lemma A.17, $\sum_{i=1}^{n} \alpha_i I^\star_{A_i}(\omega) = \sum_{i=1}^{n} \alpha_i \nu_\omega(A_i)$ for all $\omega \in \Omega'$. Since $\Omega \setminus \Omega'$ is $\succcurlyeq$-null,

$$\sum_{i=1}^{n} \alpha_i I_{A_i} \sim \sum_{i=1}^{n} \alpha_i I^\star_{A_i} \sim \sum_{i=1}^{n} \alpha_i \nu_\omega(A_i),$$

that is, $f \sim g$ if f is a simple function and g is as above.

Now, given a Σ-measurable bounded f, let $B > 0$ be a bound for f, that is, $f(\omega) \in [-B, B]$, $\forall \omega \in \Omega$. For each $j \in \{-2^n, -2^n + 1, ..., 2^n\}$ and $n \in \mathbb{N}$, let $A_{j,n} \equiv \{\omega : f(\omega) \in [\frac{jB}{2^n}, \frac{(j+1)B}{2^n})\}$. Define:

$$f^n(\omega) = \sum_{j=-2^n}^{2^n} \frac{jB}{2^n} 1_{A_{j,n}}(\omega),$$

that is, $f^n(\omega)$ is valued $\frac{jB}{2^n}$, whenever $\frac{jB}{2^n} \leqslant f(\omega) < \frac{(j+1)B}{2^n}$. It is easy to see that f^n is a sequence of simple functions that converge pointwise to f, $f^n \uparrow f$.

Now, fix g as in the statement above and define $g^n : \Omega \to \mathbb{R}$ as follows:

$$g^n(\omega) = \begin{cases} \int f^n(\tilde{\omega})\nu(\omega, d\tilde{\omega}), & \text{if } \omega \in \Omega' \\ g(\omega) & \text{if } \omega \in \Omega \setminus \Omega' \end{cases}$$

Fix $\omega \in \Omega'$ and recall that $\nu_{\omega'}(\cdot)$ is a countably additive probability measure. By the Lebesgue Monotone Convergence Theorem, we have $g^n(\omega) = \int f^n(\tilde{\omega})\nu(\omega, d\tilde{\omega}) \to \int f(\tilde{\omega})\nu(\omega, d\tilde{\omega}) = g(\omega)$. Since this is valid for all $\omega \in \Omega'$ and $g^n(\omega) = g(\omega)$ if $\omega \notin \Omega'$, $\{\omega : \lim_{n\to\infty} g^n(\omega) \neq g(\omega)\} = \emptyset$. This shows that by $g^n \to g$. Since f^n is simple, in the beginning we have established that $f^n \sim g^n$. Therefore, by continuity $f \sim g$. ■

At this point, it is useful to recall the following standard definitions and notation. Let X be a complete separable metrizable (Polish) space and $\mathcal{X}$ denote the σ-field of its Borel sets. The set $\Delta(X)$ denotes all probability measures which are defined on $\mathcal{X}$. Let $C_b(X)$ denote the set of continuous and bounded real-valued functions $f : X \to \mathbb{R}$. The set $\Delta(X)$ is endowed with its usual weak*-topology, that is, the $\sigma(\Delta(X), C_b(X))$-topology. It is well known that $\Delta(X)$ is also Polish. Let $\mathcal{P}_\tau$ and $\mathcal{P}_\tau^{er}$ denote, respectively, the sets of τ-invariant and τ-ergodic measures in $\Delta(\Omega)$. Also, let $\mathcal{P}_\tau^{ex}$ denote the set of extreme points of $\mathcal{P}_\tau$, that is, the set of .

Lemma A.19 *If Ω is compact, then the set $\mathcal{P}_\tau^{ex}$ is a G_δ subset of $\Delta(\Omega)$.*

Proof: The set $\Delta(\Omega)$ is compact if and only if Ω is compact—see (Aliprantis and Border 2006, Theorem 15.11). The statement then follows from (Aliprantis and Border 2006, Lemma 7.63)—see also (Phelps 2001, Proposition 1.3). ■

It is useful to observe that the above result is false without the assumption that Ω is compact. (Jayne and Rogers 1977) construct examples of closed convex bounded subsets of Banach spaces whose set of extreme points are not Borel measurable. The conclusion above in particular implies that $\mathcal{P}_\tau^{ex}$ is isomorphic to a Borel set of the real line. This conclusion is important in the proof of Lemma A.24 below, but the assumption that Ω is compact is needed nowhere else.

Lemma A.20 *The set of τ-ergodic measures is equal to the set of extreme points of τ-invariant measures, that is, $\mathcal{P}_\tau^{ex} = \mathcal{P}_\tau^{er}$.*

Proof: This comes from (Aliprantis and Border 2006, Theorem 19.25). ∎

This allows us to define the set of parameters:

Definition A.21 *Let $\mathcal{P}_\tau^{ex} = \mathcal{P}_\tau^{er}$ be written as $\{P^\theta\}_{\theta \in \Theta}$, where Θ is a set of parameters, which is identified with $\mathcal{P}_\tau^{ex} = \mathcal{P}_\tau^{er} \subset \Delta(\Omega)$ and, therefore, inherits the topology and measurable structure of $\Delta(\Omega)$.*

The following definition is standard.

Definition A.22 *Let X and Y be separable metrizable spaces. A Markov transition (or simply a transition) from X to Y is a Borel measurable function $M : X \to \Delta(Y)$.*

Lemma A.23 *The function $I^\star : \Omega' \times \Sigma \to [0,1]$ is a Markov kernel and $\vartheta : \Omega' \to \Delta(\Omega)$ defined by $\vartheta(\omega)(\cdot) = \nu_\omega(\cdot)$ is a Markov transition and takes values in $\mathcal{P}_\tau^{er}$.*

Proof: The set Ω' is clearly separable metrizable. Corollary A.3 shows that $\nu_\omega \in \mathcal{P}_\tau^{er}$ and the fact that $I_A^\star$ is τ-invariant for each $A \in \Sigma$ implies that ϑ above defined is measurable, by Aliprantis and Border (2006, Theorem 19.7, p. 627). This establishes that ϑ is a Markov transition, which is equivalent to $I^\star : \Omega' \times \Sigma \to [0,1]$ is a Markov kernel—see Aliprantis and Border (2006)'s Theorem 19.12 and 19.13, p. 630. ∎

Lemma A.24 *There is an extension of $I^\star : \Omega' \times \Sigma \to [0,1]$ to $\nu : \Omega \times \Sigma \to [0,1]$ which is a Markov kernel. Moreover, there is a Markov transition $\vartheta : \Omega \to \Theta$ such that $\vartheta(\omega)(\cdot) = I^\star_\omega(\cdot)$, for every $\omega \in \Omega'$.*

Proof: Let $I^\star : \Omega' \times \Sigma \to [0,1]$ and $\vartheta : \Omega' \to \Delta(\Omega)$ be as in Lemma A.23. By lemmas A.19 and A.20, the set of ergodic measures Θ is a G_δ set and, therefore, isomorphic to a Borel subset of the real line. By (Shortt 1983, Theorem 1), there exists an extension ϑ of ϑ' from Ω' to Ω . By (Aliprantis and Border 2006, Theorem 19.12), the function $\nu : \Omega \times \Sigma \to [0,1]$ defined by $\nu(\omega, \cdot) \equiv \vartheta(\omega)(\cdot)$ is a Markov kernel. Since ϑ extends ϑ', then $\nu(\omega, A) = I^\star(\omega, A)$ for every $(\omega, A) \in \Omega' \times \Sigma$, that is, ν is an extension of $I^\star$. This completes the proof. ∎

From now on, let $\nu : \Omega \times \Sigma \to [0,1]$ and $\vartheta : \Omega \to \Theta$ be as defined above.

Lemma A.25 *For each $f \in \mathcal{D}$, we have:*

$$f(\cdot) \sim \int f(\tilde{\omega}) \ P^{\vartheta(\cdot)}(d\tilde{\omega}), \tag{25}$$

Proof: Define $g : \Omega \to \mathbb{R}$ by $g(\omega) = \int f(\tilde{\omega}) \nu(\omega, d\tilde{\omega})$. Observe that we can we write $P^{\vartheta(\omega)} = \vartheta(\omega)$, and equivalently define $g(\omega) = \int f(\tilde{\omega}) \ P^{\vartheta(\omega)}(d\tilde{\omega}), \forall \omega \in \Omega$. By (Aliprantis and Border 2006, Theorem 19.7), g is Σ-measurable. Thus, Lemma A.18 implies that $f \sim g$, as we wanted to show. ∎

Corollary A.4 *Let the set $A \in \Sigma$ be such that $\mu(A) = 0$ for all $\mu \in \mathcal{P}^{er}_\tau$. Then A is $\succcurlyeq$-null.*

Proof: Let f, g and h be arbitrary acts and let fAh denote the act that takes value $f(\omega)$ if $\omega \in A$ and $h(\omega)$ if $\omega \notin A$. By Lemma A.25, $fAh \sim \int (fAh)(\tilde{\omega}) \ P^{\vartheta(\cdot)}(d\tilde{\omega})$. Since $P^{\vartheta(\omega)} \in \mathcal{P}^{er}_\tau$ and $\mu(A) = 0$ for all $\mu \in \mathcal{P}^{er}_\tau$, then $\int (fAh)(\tilde{\omega}) \ P^{\vartheta(\omega)}(d\tilde{\omega}) = \int h(\tilde{\omega}) \ P^{\vartheta(\omega)}(d\tilde{\omega}), \forall \omega \in \Omega$. Therefore, $fAh \sim \int h(\tilde{\omega}) \ P^{\vartheta(\cdot)}(d\tilde{\omega}) \sim h$. Repeating the argument for g, we conclude $fAh \sim h \sim gAh$, that is, A is $\succcurlyeq$-null.

Definition A.26 *A map $U : \mathcal{D} \to \mathcal{D}$ is a τ-universal conditional expectation if it satisfies the following two conditions:*

1. *$U(f)$ is $\mathcal{E}_\tau$-measurable for every $f \in \mathcal{D}$;*

2. *for any τ-invariant measure $\mu \in \mathcal{P}_\tau$, $f \in \mathcal{D}$ and $A \in \mathcal{E}_\tau$,[22]*

$$\int_A f \; d\mu = \int_A U(f) \; d\mu. \tag{26}$$

Whenever τ is clear from the context, we will refer only to universal conditional expectation.

Lemma A.27 *The map $M^\vartheta : \mathcal{D} \to \mathcal{D}$ defined by:*

$$M^\vartheta(f)(\cdot) \equiv \int f(\tilde{\omega}) \; P^{\vartheta(\cdot)}(d\tilde{\omega}), \tag{27}$$

is a universal conditional expectation.

Proof: This map is the specialization of the general map defined by (Varadarajan 1963) for the semigroup generated by τ. As such, it is a universal conditional expectation. ∎

Corollary A.5 *If $\tilde{\vartheta} : \Omega \to \Theta$ is such that*

$$f(\cdot) \sim \int f(\tilde{\omega}) \; P^{\tilde{\vartheta}(\cdot)}(d\tilde{\omega}), \tag{28}$$

is satisfied and the map $M^{\tilde{\vartheta}} : \mathcal{D} \to \mathcal{D}$ defined by

$$M^{\tilde{\vartheta}}(f)(\cdot) \equiv \int f(\tilde{\omega}) \; P^{\tilde{\vartheta}(\cdot)}(d\tilde{\omega}), \tag{29}$$

is a universal conditional expectation, then $D = \{\omega \in \Omega : \vartheta(\omega) \neq \tilde{\vartheta}(\omega)\}$ is $\succcurlyeq$-null.

Proof: Since M^ϑ and $M^{\tilde{\vartheta}}$ are both universal conditional expectations, $\mu(D) = 0$ for all $\mu \in \mathcal{P}_\tau$—see Varadarajan (1963, section 4). By Corollary A.4, D is $\succcurlyeq$-null. ∎

---

[22]It is sufficient to require (26) to hold for $f = 1_A$, for any $A \in \Sigma$ since this implies (26) for all $f \in \mathcal{D}$.

B Proofs

B.1 Proof of Theorem 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 $u : \Delta(C) \to \mathbb{R}$ that represents $\succcurlyeq$ on $\Delta(C)$. Analogously, let $\mathbb{P}^{\mathcal{D}}$ denote the set of preferences on real-valued functions $\mathcal{D}$ satisfying the properties stated in Corollary A.1. 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 (19) and, conversely, to each $\succcurlyeq^{\mathcal{D}} \in \mathbb{P}^{\mathcal{D}}$ and linear u corresponds a $\succcurlyeq \in \mathbb{P}$, as defined by (20).

Fix $f \in \mathcal{F}$ and a linear function $v : \Delta(C) \to \mathbb{R}$, and define the sets:

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

and

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

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

Since C is finite, $\Delta(C)$ is finite dimensional and we can find a countable set $\mathcal{U}$ of linear functions $u : \Delta(C) \to \mathbb{R}$ 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 $\tilde{A} \equiv \cap_{u \in \mathcal{U}} \tilde{A}_u$. 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 $u(f) \sim^{\mathcal{D}} \int u(f) \, dP^{\vartheta(\cdot)} = u\left(\int f \, dP^{\vartheta(\cdot)}\right)$, because u is linear. By (20), $f \sim \int f \, dP^{\vartheta(\cdot)}$ 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. ∎

Note that the assumption that C is finite was used in the proof of the above theorem only to guarantee that the convergence of $v(f(\omega))$ for all linear $v : \Delta(C) \to \mathbb{R}$ implies the convergence of $f(\omega)$. If C is not finite, the above proof still establishes the following:

Theorem B.1 *Assume that $\succcurlyeq$ satisfies assumptions 1-5, Ω is compact Polish, C is a Polish, and $\succcurlyeq$ is τ-invariant. Then there exists a decomposition map $\vartheta : \Omega \to \Theta$, where Θ is the set of τ-ergodic parameters, and an affine utility function $u : \Delta(C) \to \mathbb{R}$ which represents $\succcurlyeq$ on $\Delta(C)$ such that:*

1. *(Θ, ϑ) is sufficient for $\succcurlyeq$.*

2. *for any $f \in \mathcal{F}$, the set $\{\omega \in \Omega : \nexists \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} u\left(f\left(\tau^j \omega\right)\right)\}$ is $\succcurlyeq$-null.*

B.2   Proof of Theorem 2

We first prove the equivalence of (2) and (3) without using the semi-group Γ. Then, we define Γ and using this Γ, we show the equivalence of (1) and (3).

(2) $\Rightarrow$ (3): Assume that $(\overline{\Theta}, \overline{\vartheta})$ is a sufficient parametrization of $\succcurlyeq$. We want to prove that $A \equiv \vartheta^{-1}(\Theta \setminus \overline{\Theta})$ is $\succcurlyeq$-null. Assume otherwise. Then there exist $f, g, h \in \mathcal{F}$ such that the acts

$$f' \equiv \left[\begin{array}{ll} f(\omega), & \text{if } \omega \in A \\ h(\omega), & \text{if } \omega \notin A \end{array} \right] \text{ and } g' \equiv \left[\begin{array}{ll} g(\omega), & \text{if } \omega \in A \\ h(\omega), & \text{if } \omega \notin A \end{array} \right]$$

are not indifferent. Fix $\omega \in \Omega$ and let $\theta = \overline{\vartheta}(\omega) \in \overline{\Theta}$. Then,

$$\int_\Omega f' \, dP^{\overline{\vartheta}(\omega)} = \int_{\vartheta^{-1}(\theta)} f' \, dP^\theta = \int_{\vartheta^{-1}(\theta)} h \, dP^\theta,$$

where the last inequality comes from the fact that $\vartheta^{-1}(\theta) \subset A^c$. A similar equality holds also for g, for any ω. This implies that the integrals $\int f' \, dP^{\overline{\vartheta}(\cdot)}$ and $\int g' \, dP^{\overline{\vartheta}(\cdot)}$ are equal and thus indifferent. But since $(\overline{\Theta}, \overline{\vartheta})$ is a sufficient parametrization of $\succcurlyeq$, this means that $f' \sim g'$, which is a contradiction.

(3) $\Rightarrow$ (2): Assume that $A \equiv \vartheta^{-1}(\Theta \setminus \overline{\Theta})$ is $\succcurlyeq$-null. Define $\overline{\vartheta}(\omega) = \vartheta(\omega)$ if $\omega \in \vartheta^{-1}(\overline{\Theta})$ and $\overline{\vartheta}(\omega) = \theta \in \overline{\Theta}$ (arbitrarily) if $\omega \in A$. Since (Θ, ϑ) is a

sufficient parametrization and A is $\succ$-null, it is clear that $f \sim \int f\, dP^{\bar{\vartheta}(\cdot)}$, that is, $(\overline{\Theta}, \bar{\vartheta})$ is a sufficient parametrization of $\succ$.

We now define Γ as follows. For each $\tilde{\omega} \in \vartheta^{-1}(\overline{\Theta})$, define the transformation $\gamma_{\tilde{\omega}} : \Omega \to \Omega$ by:

$$\gamma_{\tilde{\omega}}(\omega) = \begin{cases} \omega, & \text{if } \omega \in \vartheta^{-1}(\overline{\Theta}) \\ \tilde{\omega}, & \text{if } \omega \in A = \vartheta^{-1}(\Theta \setminus \overline{\Theta}) \end{cases}$$

That is, the transformation $\gamma_{\tilde{\omega}}$ takes every point in A to $\tilde{\omega}$, while leaving every other point fixed. Let $\Gamma \equiv \{\gamma_\omega : \omega \in \vartheta^{-1}(\overline{\Theta})\}$. It is easy to see that this defines a semi-group.

$(1) \Rightarrow (3)$: Now assume that $\succ$ is weakly Γ-invariant, that is, $f \sim f \circ \gamma$ for all $\gamma \in \Gamma$. We want to prove that A is $\succ$-null. Assume otherwise. Then there exist $f, g, h \in \mathcal{F}$ such that the acts

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

are not indifferent. Fix some $\tilde{\omega} \notin A$. Since $f' \sim f' \circ \gamma_{\tilde{\omega}}$, $g' \sim g' \circ \gamma_{\tilde{\omega}}$ and $f' \circ \gamma_{\tilde{\omega}}(\omega) = g' \circ \gamma_{\tilde{\omega}}(\omega), \forall \omega$, then $f' \sim g'$, which is a contradiction.

$(3) \Rightarrow (1)$: If $\vartheta^{-1}(\Theta \setminus \overline{\Theta})$ is $\succ$-null, then $f \sim f \circ \gamma_\omega, \forall \omega \in \vartheta^{-1}(\overline{\Theta})$. ∎

B.3 Proof of Theorem 3

Although the assumption $u(\Delta(C)) = \mathbb{R}$ rules out C finite, this is not a problem for this and related results. See Theorem B.1.

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

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

which establishes (1).

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

Fix $\gamma \in \Gamma$ and denote $\Theta_\gamma \equiv \{\theta \in \Theta_\Gamma^c : P^\theta \neq P^\theta \circ \gamma^{-1}\}$. Observe that $A_\gamma = \cup_{\theta \in \Theta_\gamma} \vartheta^{-1}(\theta) = \vartheta^{-1}(\Theta_\gamma)$. 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 $\tilde{A}_\gamma \equiv \cup_{\theta \in \Theta_\gamma} B^\theta$ 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

$$
f' \equiv \left[\begin{array}{ll} f(\omega), & \text{if } \omega \in \tilde{A}_\gamma \\ g(\omega), & \text{if } \omega \notin \tilde{A}_\gamma \end{array} \right] \text{ and } g
$$

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 $\vartheta(\omega) = \vartheta(\gamma(\omega)) = \theta \in \Theta_\gamma$,[23]

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

---

[23] Recall that ϑ is Γ-invariant.

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

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

Also, in this case,

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

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

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 $A_\gamma \setminus \tilde{A}_\gamma = \cup_{\theta \in \Theta_\gamma} \left[\vartheta^{-1}(\theta) \setminus B^\theta\right]$ 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.4 Proof of Proposition 4.1

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.[24] Define $M^\vartheta : \mathcal{F} \to \mathcal{F}_\Gamma$ by:

$$M^\vartheta(f)(\omega) = \int_\Omega f \, dP^{\vartheta(\omega)}. \tag{30}$$

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:

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

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

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 (7).

Given an aggregator $\succcurlyeq$ satisfying assumptions 1-5, define $\succcurlyeq$ on $\mathcal{F}_\Gamma$ by (7). 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, (7) also holds. Again, it is easy to see that $\succcurlyeq$ satisfies assumptions 1-5. ∎

References

AL-NAJJAR, N. I., AND L. DE CASTRO (2010): "Observability and "Second-Order Acts"," Northwestern University.

AL-NAJJAR, N. I., L. DE CASTRO, AND M. PAI (2010): "A Subjective Framework for Testing," Northwestern University.

ALIPRANTIS, C., AND K. BORDER (2006): *Infinite dimensional analysis: a hitchhiker's guide.* Springer Verlag.

BILLINGSLEY, P. (1995): *Probability and Measure*, Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edn., A Wiley-Interscience Publication.

CERREIA, S., F. MACCHERONI, M. MARINACCI, AND L. MONTRUCCHIO (2008): "Uncertainty Averse Preferences," working paper; Collegio Carlo Alberto.

CHEW, S., AND J. SAGI (2008): "Small worlds: Modeling attitudes toward sources of uncertainty," *Journal of Economic Theory*, 139(1), 1–24.

DAWID, A. (1982): "Intersubjective statistical models," *Exchangeability in Probability and Statistics*, pp. 217–232.

DE CASTRO, L., AND N. I. AL-NAJJAR (2009): "A Subjective Foundation of Objective Probability," Northwestern University.

DE FINETTI, B. (1937): "La prévision: ses lois logiques, ses sources subjectives," *Annales de l'Institut Henri Poincaré*, 7, 1–68.

DIACONIS, P., AND D. FREEDMAN (1980): "de Finetti's theorem for Markov chains," *Ann. Probab*, 8(1), 115–130.

——— (1984): "Partial exchangeability and sufficiency," *Proceedings of the Indian Statistical Intitute Golden Jubilee International conference on Statistics: Applications and New Directions*, pp. 205–236.

DYNKIN, E. (1978): "Sufficient statistics and extreme points," *The Annals of Probability*, 6(5), 705–730.

EPSTEIN, L., AND K. SEO (2010): "Symmetry of Evidence without Evidence of Symmetry," *Theoretical Economics*, 5(313-368).

ERGIN, H., AND F. GUL (2009): "A Subjective Theory of Compound Lotteries," *Journal of Economic Theory*, pp. 899–929.

GHIRARDATO, P., F. MACCHERONI, M. MARINACCI, AND M. SINISCALCHI (2003): "A Subjective Spin on Roulette Wheels," *Econometrica*, 71(6), 1897–1908.

GILBOA, I., F. MACCHERONI, M. MARINACCI, AND D. SCHMEIDLER (2008): "Objective and Subjective Rationality in a Multiple Prior Model," *Econometrica (forthcoming)*.

GRANT, S., B. POLAK, AND T. STRZALECKI (2009): "Second-Order Expected Utility," Discussion paper, Harvard University.

HERSTEIN, I., AND J. MILNOR (1953): "An axiomatic approach to measurable utility," *Econometrica, Journal of the Econometric Society*, 21(2), 291–297.

JACKSON, M. O., E. KALAI, AND R. SMORODINSKY (1999): "Bayesian Representation of Stochastic Processes under Learning: de Finetti Revisited," *Econometrica*, 67(4), 875–893.

JAYNE, J. E., AND C. A. ROGERS (1977): "The extremal structure of convex sets," *Journal of Functional Analysis*, 26(3), 251 – 288.

KLIBANOFF, P., M. MARINACCI, AND S. MUKERJI (2005): "A smooth model of decision making under ambiguity," *Econometrica*, 73(6), 1849–1892.

KLIBANOFF, P., S. MUKERJI, AND K. SEO (2010): "Relevance and Symmetry," Northwestern University.

LAURITZEN, S. (1984): "Extreme Point Models in Statistics," *Scandinavian Journal of Statistics*, pp. 65–91.

NAU, R. (2001): "Uncertainty Aversion with Second-Order Utilities and Probabilities," in *2nd International Symposium on Imprecise Probabilities and Their Applications, Ithaca, New York, June 2001*.

——— (2006): "Uncertainty Aversion with Second-Order Utilities and Probabilities," *Management Science*, 52(1), 136.

NEILSON, W. (1993): "Ambiguity aversion: An axiomatic approach using second order probabilities," Discussion paper, mimeo.

——— (2009): "A simplified axiomatic approach to ambiguity aversion," *Journal of Risk and Uncertainty*, forthcoming.

PHELPS, R. (2001): *Lectures on Choquet's theorem.* Springer.

ROYDEN, H. L. (1968): *Real Analysis.* MacMillan Publishing Co., Inc., New York, 2 edn.

SAVAGE, L. J. (1954): *The foundations of statistics.* John Wiley & Sons Inc., New York.

SHORTT, R. M. (1983): "The Extension of Measurable Functions," *Proceedings of the American Mathematical Society*, 87(3), 444–446.

STRZALECKI, T. (2010): "Axiomatic Foundations of Multiplier Preferences," *Econometrica (forthcoming)*.

VARADARAJAN, V. (1963): "Groups of automorphisms of Borel spaces," *Trans. Amer. Math. Soc*, 109(2), 191–220.