CHARACTERIZATIONS OF SMOOTH AMBIGUITY BASED ON CONTINUOUS AND DISCRETE DATA

Stefania Minardi† and Andrei Savochkin‡

In the Anscombe and Aumann (1963) setup, we provide conditions for a collection of observations to be consistent with a well-known class of smooth ambiguity preferences (Klibanoff, Marinacci, and Mukerji, 2005). Each observation is assumed to take the form of an equivalence between an uncertain act and a certain outcome. We provide three results that describe these conditions for datasets of different cardinality. Our findings uncover surprising links between the smooth ambiguity model and classic mathematical results in complex and functional analysis.

KEYWORDS: smooth ambiguity, variational preferences, revealed preference, completely monotone functions, Afriat inequalities, moment problem.

1. INTRODUCTION

1.1. The objectives of the paper

In this work, we provide novel results about the smooth ambiguity model of Klibanoff, Marinacci, and Mukerji [26] — one of the most popular models of decision making under ambiguity. Since the celebrated work of Ellsberg [17], ambiguity refers to a decision maker’s lack of ability to assign a unique probability to some events. It has been a major subject in decision theory and has also attracted much attention in the applied literature.1

---

*This paper is an expanded version of an earlier work circulated under the title “A Functional Characterization of Smooth Ambiguity.” We thank an anonymous referee and an associate editor for very insightful comments that have led to a significant improvement of the paper. We are also very grateful to Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci, Efe Ok, Debraj Ray, Marco Scarsini, Tristan Tomala, and Nicolas Vieille for helpful conversations. Needless to say, all mistakes are our own.

†Economics and Decision Sciences Department, HEC Paris, minardi@hec.fr

‡New Economic School, asavochkin@nes.ru

1A comprehensive survey of the literature on ambiguity can be found in Gilboa and Marinacci [20].
The smooth ambiguity model stands alone in the field for its remarkable functional form and numerous applications to economics and finance. At the same time, the original analysis of Klibanoff et al. [26] requires an extension of the standard Anscombe and Aumann [2] framework. As is well known, this extension is demanding in terms of what the analyst needs to observe, making it more difficult to assess the ability of the model to accommodate agents’ observed preferences. In turn, the axiomatic characterization of the smooth ambiguity model under the constraints of the standard Anscombe-Aumann setup has been an open problem since its inception. We will review the model and discuss this challenge momentarily.

Before proceeding further, we summarize our contribution. The paper advances the agenda of establishing testable properties of the smooth ambiguity model in the Anscombe-Aumann setup. Our goal is to derive conditions for a collection of observations to be consistent with a special case of the model — namely, the smooth ambiguity preferences with constant absolute ambiguity aversion (henceforth, CASAP). The dataset on preferences is assumed to take the form of certainty equivalents: For any given uncertain prospect, the analyst has to elicit a sure outcome that the decision maker regards as equivalent to it. Our conditions are not intended to constitute any sort of axiomatic foundation of the model as, for instance, they clearly lack normative content. Nevertheless, the results of Section 3 and, partially, of Section 4 provide characterizations of the CASAP model: The conditions stated there must hold for any data generated by the model; conversely, if these conditions hold, then one can find the values of the parameters of the model that give rise to the data. Section 3 deals with the ultimate case in which the dataset is complete and contains certainty equivalents of all possible uncertain prospects. In this section, our result uncovers intrinsic properties of the CASAP model that are related, in particular, to smoothness, and demonstrates unexpected links between the model and certain notable
results in complex analysis. Section 4 is concerned with a more practical situation in which the dataset of observations on preferences is discrete. The results of this section are based on recent advances in functional analysis.

Our results are conceptually related to the decision-theoretic literature that originated with Afriat [1]. In his seminal paper, Afriat considers finite datasets describing the agent’s choices from budget sets, and shows that the property of cyclical monotonicity of a dataset is equivalent to the agent having a concave utility function. Subsequent works of different authors made similar inquiries for various other specifications of preferences and provided characterizing conditions for their consistency with a finite number of datapoints. Focusing on market behavior, Bayer, Bose, Polisson, and Renou [3] and Polisson, Quah, and Renou [30] develop tests for different models under ambiguity (including smooth ambiguity preferences) which involve solving a system of non-linear Afriat-type inequalities. Compared to these works, our results share the motivation but, mathematically, have a very different structure.

1.2. The Smooth Ambiguity model

To fix ideas, we recall first the smooth ambiguity model in the Anscombe-Aumann setting. Consider a state space $S$ endowed with an algebra $\Sigma$, a space $X$ of simple (i.e., finite support) probability distributions over a set of real outcomes containing the interval $[-1,1]$, and a set $\Delta$ of all probability measures $p : \Sigma \to [0,1]$. An act $f : S \to X$ is a simple $\Sigma$-measurable function.

According to the smooth ambiguity model studied by Klibanoff et al. [26], the functional

$$V(f) = \int_\Delta \phi \left( \int_S u(f(s)) \, dp(s) \right) \, d\mu(p)$$

Apart from these papers, we should mention that Epstein [18] and Echenique and Saito [16] provide revealed preference axioms to verify the consistency of market behavior with, respectively, probabilistic sophistication and risk-averse subjective expected utility.
represents the agent’s preferences \( z \) over acts — i.e., \( f \succ g \) if and only if \( V(f) \geq V(g) \) for any acts \( f \) and \( g \). In this representation, \( u \) is an affine von Neumann-Morgenstern utility index, \( \mu \) is a probability measure on \( \Delta \), and \( \phi \) is a strictly increasing function. In words, when evaluating an act \( f \), the decision maker uses a procedure that is based on a double mathematical expectation. First, she calculates the expected utility of \( f \) with respect to each probability measure \( p \) in \( \Delta \). Second, she aggregates those numbers by applying the “second-order utility function” \( \phi \) and taking a weighted average with respect to a probability measure \( \mu \) on \( \Delta \).

Compared to other models that incorporate ambiguity, such as the well-known maxmin expected utility preferences (Gilboa and Schmeidler [21]), the smooth ambiguity model has the important advantage of providing a separation between ambiguity — a characteristic of the agent’s subjective beliefs — and the attitude towards ambiguity — a characteristic relative to tastes and controlled by the curvature of \( \phi \). This distinction is useful when performing comparative statics exercises: for instance, it allows studying the effects of a change in the attitude toward ambiguity, while holding ambiguity fixed, by controlling only the curvature of \( \phi \). Moreover, it has the technical convenience of being smooth (differentiable) as long as \( \phi \) is smooth. All these features make the model highly tractable, as the classic techniques for performing comparative statics exercises in the risk domain can also be applied to the ambiguity domain. Indeed, it has been extensively adopted in applications to finance and economics.\(^3\)

As noted earlier, Klibanoff et al. [26] develop foundations of this model in a framework that is richer than the standard Anscombe-Aumann setup.\(^4\)

\(^3\)See Collard, Mukerji, Sheppard, and Tallon [11], Gollier [22], Ju and Miao [25], and Maccheroni, Marinacci, and Ruffino [27], among others.

\(^4\)In settings richer than the Anscombe-Aumann one, Nau [29], Ergin and Gul [19], and Seo [31] provide axiomatic foundations for representations similar to (1).
As in the Anscombe-Aumann setup, the decision maker’s core preferences are described by a binary relation $\succeq$ over first-order acts — functions $S \rightarrow X$ as above. In addition to that, Klibanoff et al. [26] posit a binary relation $\succeq^2$ over second-order acts — functions $\Delta \rightarrow X$ that constitute payoff profiles contingent on probabilities (not states).

To appreciate the differences between the first- and second-order acts, let us give an illustration. Learning the agent’s preferences over first-order acts may require observing her choice in response to the following question:

Do you prefer a bet paying $\$1$ if the winner in the next presidential elections in the US is a Democrat over a bet paying $\$1$ if the winner is a Republican?

By contrast, comparing two second-order acts may entail answering a question like:

Do you prefer a bet paying $\$1$ if the probability of a Democrat winning the next presidential election is between 0.45 and 0.46, over a bet paying $\$1$ if the probability of a Democrat winning the next presidential election is between 0.22 and 0.24?

The main result of Klibanoff et al. [26] lists a set of properties that $\succeq$ and $\succeq^2$ have to satisfy jointly and shows that they are equivalent to representation (1).

1.3. Overview of our Results

As announced earlier, we focus on the class of smooth ambiguity preferences that exhibit constant absolute attitude towards ambiguity. It is well known that, for this subset of smooth preferences, the function $\phi$ takes the exponential form $\phi(t) = -\frac{1}{\lambda}e^{-\lambda t}$ for some $\lambda > 0$ (see Klibanoff et al. [26, Proposition 2]). Moreover, within the general class of uncertainty averse preferences, Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [7]

---

5 This functional form is often adopted in applications of the smooth ambiguity model — see, for instance, Ju and Miao [25].
show that the CASAP model constitutes exactly the intersection between smooth ambiguity preferences and variational preferences (see their Corollary 22 and Theorem 23). Within the CASAP class, a well-known special case corresponds to the multiplier preferences introduced by Hansen and Sargent [23] and axiomatized by Strzalecki [33].

Our results answer the question about whether a given set of datapoints is consistent with the hypothesis that the agent’s preferences are CASAP. In the analyst’s baseline dataset, each datapoint consists of an act $f$ and its certainty equivalent — an outcome $x \in X$ such that the decision maker is indifferent between $f$ and $x$. However, to focus on the specifics of the CASAP model, we assume that the analyst has already analyzed the agent’s preferences restricted to constant acts and recovered her utility function $u : X \to \mathbb{R}$. Hence, for the purpose of stating our theorems, we take as a primitive a functional $I$ that is defined on the space $u(X)^S$ of utility acts and takes values in $u(X)$. The objective is to provide conditions describing whether the values of $I$ on a given set of points in $u(X)^S$ allow for the existence of some $\lambda > 0$ and a probability measure $\mu$ on the set of first-order probabilities $\Delta$ such that representation (1) holds.

Our results are organized by the cardinality of the dataset at the analyst’s disposal. The first one considers the case of $I$ known on its whole domain, which is equivalent to a complete observability of the agent’s preferences — the situation typically studied in axiomatic decision theory. Theorem 1 considers a functional $I$ that is known to belong to the variational class and claims that it also admits a smooth ambiguity representation if and only if the exponential of $I$ satisfies the following properties:

---

6For a recent work about the macroeconomic implications of modeling uncertainty using multiplier preferences, see Hansen and Sargent [24].

7These functionals are extensively studied for larger classes of preferences — see, e.g., Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [6] and Cerreia-Vioglio, Maccheroni, Marinacci, and Rustichini [5].
(i) it is completely monotone;
(ii) it can be extended to an entire function in the complex domain;
(iii) it satisfies a specific condition limiting its growth at infinity.

Hence, this result provides a characterization of the CASAP model in functional terms. As we discuss after stating the theorem, it also delivers a characterization of datasets consistent with the CASAP model in the cases of \( I \) being known on other types of sufficiently rich sets, such as any open set.

Although Theorem 1 provides a basic functional characterization of the CASAP model, its conditions cannot be verified algorithmically. Our subsequent results move in the direction of Afriat-type analysis. In particular, Theorem 2 and Proposition 3 suggest procedures to test for CASAP by relying only on observable data. From this perspective, the results in Section 4 represent a certain step towards a revealed preference analysis for this model (and, in turn, smooth ambiguity preferences).

Our Theorem 2 provides characterizing conditions for a countably infinite collection of observations of \( I \) on a discrete grid to be consistent with the CASAP model. This characterization problem appears to be closely related to the problem of representing a sequence of real numbers as the sequence of moments of some measure supported on a particular subset of \( \mathbb{R}^n \) — i.e., the so-called Problem of Moments, which was first advanced by famous mathematicians Stieltjes, Hausdorff, and Hamburger in late nineteenth–early twentieth century, and is still a topic of active research in mathematics.

These conditions of Theorem 2 amount to requiring positive semidefiniteness of certain matrices. Moreover, since each of these matrices is constructed by using a finite number of datapoints, the theorem can also be used to refute the hypothesis that a finite dataset can be rationalized by CASAP. However, the positive semidefinite conditions of Theorem 2 are not sufficient to guarantee the existence of a CASAP compatible with a given
finite dataset. To this end, Proposition 3 establishes a sufficient result by requiring (i) a finite number of certain matrices to be positive semidefinite; and, (ii) some of these matrices to have the same rank.

The rest of the paper is organized as follows. Section 2 introduces the decision-theoretic setup and recalls some basic results on variational preferences. Section 3 reviews the concepts that we need from complex analysis and states our functional characterization — Theorem 1. In Section 4 we present and prove our results based on discrete datasets — Theorem 2 and Proposition 3 — followed by examples of their use. In the Appendix we state and prove a general version of Theorem 1.

2. PRELIMINARIES

2.1. Decision-Theoretic Setup

We consider the classic Anscombe-Aumann setup with a finite state space. Let $S$ denote a finite set of states of the world and $X$ a set of outcomes, which is assumed to be a convex subset of a topological vector space. An act is an element $f$ of $X^S$; we denote by $f_s$ the consequence of act $f$ in state $s$ and by $F$ the set $X^S$ of all acts. Let $\Delta$ be the set of all probability distributions $p$ on $S$.

The decision maker’s preferences are described by a binary relation $\succsim$ on $F$.

Moreover, given a function $u : X \rightarrow \mathbb{R}$, a utility act is an element $\varphi$ of $u(X)^S$. We denote by $\varphi_s$ the $s$-th component of $\varphi$ and by $B$ the set $u(X)^S$ of all utility acts. In other words, $B = \{u \circ f : f \in F\}$.

We write $\varphi \succeq \psi$ if $\varphi_s \succeq \psi_s$ for all $s \in S$.

---

8Since $S$ is finite, $\Delta$ can be viewed as a subset of the Euclidean space $\mathbb{R}^{|S|}$ and, hence, has a metric.
2.2. Variational preferences

Here, we briefly review some basic features of variational preferences (Maccheroni, Marinacci, and Rustichini [28]) which will be useful to establish our results.

Let \( D \) be a nonempty subset of \( \mathbb{R}^S \). A functional \( I : D \rightarrow \mathbb{R} \) is:

- **normalized** if \( I(c1_S) = c \) for all \( c \in \mathbb{R} \) such that \( c1_S \in D \);
- **monotone** if \( \varphi \geq \psi \) implies \( I(\varphi) \geq I(\psi) \) for all \( \varphi, \psi \in D \);
- **vertically invariant** if \( I(\varphi + c1_S) = I(\varphi) + c \) for all \( \varphi \in D \) and \( c \in \mathbb{R} \) such that \( \varphi + c1_S \in D \).

Lemma 28 and Theorem 3 of Maccheroni et al. [28] guarantee that a binary relation \( \succsim \) on \( F \) is a variational preference if and only if there exist a nonconstant affine function \( u : X \rightarrow \mathbb{R} \) and a normalized, monotone, vertically invariant, quasiconcave, and continuous functional \( I : B \rightarrow \mathbb{R} \) such that \( f \succsim g \) if and only if \( I(u(f)) \geq I(u(g)) \).

We will refer to \( I \) as the representation of \( \succsim \). Moreover, we will say that a variational preference \( \succsim \) is **unbounded** if and only if \( u(X) = \mathbb{R} \).

3. A FUNCTIONAL CHARACTERIZATION

We start this section by recalling some notions in complex analysis which are needed to state our first result.

Let \( \mathbb{C}^n := \{(z_1, \ldots, z_n) : z_i \in \mathbb{C} \text{ for all } 0 \leq i \leq n\} \) be the complex Euclidean n-space and \( \Omega \subseteq \mathbb{C}^n \) be an open set.

A function \( f : \Omega \rightarrow \mathbb{C} \) is **complex-differentiable** at \( a \in \Omega \) if there exists a neighborhood \( U \) and a power series

\[
\sum_{\alpha \in \mathbb{N}_0^n} c_\alpha (z-a)^\alpha := \sum_{\alpha_1=0}^{\infty} \cdots \sum_{\alpha_n=0}^{\infty} c_{\alpha_1,\ldots,\alpha_n} (z_1-a_1)^{\alpha_1} \cdots (z_n-a_n)^{\alpha_n}
\]

\[9\]

From the revealed preference point of view, \( \succsim \) on \( F \) is a variational preference if and only if it satisfies Axioms A1–A6 of Maccheroni et al. [28].

\[10\]

Axiom A7 of Maccheroni et al. [28] guarantees that \( u(X) = \mathbb{R} \).
which converges to \( f(z) \) for any \( z \in U \).

A function \( f : \Omega \to \mathbb{C} \) is holomorphic (or analytic) if it is complex-differentiable in a neighborhood of every point of \( \Omega \). It is well known that a holomorphic function is infinitely differentiable.

A function \( f : \mathbb{C}^n \to \mathbb{C} \) is entire if it is holomorphic over the whole space \( \mathbb{C}^n \).

**Notation 1**  We will use the following notation of multidimensional calculus. For each \( k \)-dimensional object \( x \), \( x_j \) will stand for its \( j \)-coordinate for \( j = 1, \ldots, k \). Let \( x, y \) be vectors in \( \mathbb{R}^k \) and \( t \) be a multi-index in \( \mathbb{Z}^k_+ \). The inner product of \( x \) and \( y \) is denoted as \( xy \); for \( t \in \mathbb{Z}^k_+ \), \(|t| := t_1 + \cdots + t_k \), \( x^t := x_1^{t_1} \cdots x_k^{t_k} \), and \( t! := t_1! \cdots t_k! \). Furthermore, we denote by \( D^t \) the differentiation operator

\[
D^t := \frac{\partial^{|t|}}{\partial x_1^{t_1} \cdots \partial x_k^{t_k}}.
\]

Finally, if \( x \in \mathbb{R}^k \) and \( s \in \mathbb{R} \), then we think of \( (x, s) \) as an element in \( \mathbb{R}^{k+1} \).

The following definition of complete monotonicity in the multidimensional case will play a central role in stating our first result.

**Definition 2**  A function \( F : \mathbb{R}^k_+ \to \mathbb{R} \) is completely monotone if it has derivatives \( D^t F \) for all \( t \in \mathbb{Z}^k_+ \) and

\[
(-1)^{|t|} D^t F(\varphi) \geq 0,
\]

for all \( \varphi \in \mathbb{R}^k_+ \).

We can now state our functional characterization of the CASAP model.

**Theorem 1**  Let \( n := |S| \) and \( \succ \) be an unbounded variational preference relation on \( \mathcal{F} \) with the representation \( I : \mathbb{R}^n \to \mathbb{R} \). The following conditions are equivalent:
(a) There exists $\lambda > 0$ and a probability measure $\mu$ on $\Delta$ such that

$$I(\varphi) = -\frac{1}{\lambda} \ln \left( \int_{\Delta} e^{-\lambda \varphi} \, d\mu(p) \right),$$

for all $\varphi \in \mathbb{R}^n$.

(b) There exists $\theta > 0$ such that $F_{\theta} : \mathbb{R}^{n-1} \to \mathbb{R}$ defined as

$$F_{\theta}(\varphi) := e^{-\theta I(\varphi, 0)},$$

for all $\varphi \in \mathbb{R}^{n-1}$, satisfies the following conditions:

(i) $F_{\theta}$ is completely monotone;

(ii) $F_{\theta}$ can be extended to an entire holomorphic function $\mathbb{C}^{n-1} \to \mathbb{C}$;

(iii) $|F_{\theta}(\zeta_1, \ldots, \zeta_{n-1})| \leq Ce^{\max(-\Re \zeta_1, \ldots, -\Re \zeta_{n-1}, 0)}$ for some $C > 0$ and for all $(\zeta_1, \ldots, \zeta_{n-1}) \in \mathbb{C}^{n-1}$.

The above result assumes an unbounded variational representation $I$ and considers the auxiliary function $F_{\theta}$ defined in (3) as the negative exponential of $I$. The theorem provides necessary and sufficient conditions that $F_{\theta}$ has to satisfy in order for $I$ to be a smooth representation as in (2). Condition (i) requires $F_{\theta}$ to be completely monotone. Condition (ii) requires $F_{\theta}$ to be extendable to a holomorphic function with its domain being the entire $\mathbb{C}^{n-1}$ space. Condition (iii) corresponds to an inequality used in one of the well-known Paley-Wiener theorems, and it implies that $F_{\theta}$ has a particular rate of growth at infinity. A classic result in mathematics that relates complete monotonicity to the Laplace transform of a probability measure is the Bernstein Theorem.\textsuperscript{11} Therefore, Theorem 1 provides a surprising link between the smooth ambiguity model and some classic functional properties in the complex domain.

\textsuperscript{11}Let us mention that Caballé and Pomansky [4] investigate the behavioral properties of completely monotone utility functions in the risk setting, and characterize stochastic dominance for this class of functions.
Note that one of the essential conditions of the theorem is that $F_\theta$ is extendable to an entire holomorphic function. In complex analysis, there are many fundamental results showing that such functions are uniquely determined by their restrictions to certain types of smaller sets — for instance, to any open set. As a consequence, Theorem 1 provides a characterization of consistency with the CASAP model of a functional $I$ that is observed on any subset of $\mathbb{R}^n$ on which entire functions are known to have the uniqueness property.

We conclude this section by observing that, to the best of our knowledge, this is the first paper that presents a characterization of the smooth ambiguity model within the constraints of the Anscombe-Aumann setup.

There are very few related works in the literature. Some papers studying the smooth ambiguity model in extended settings are mentioned in the Introduction. Apart from these, Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [8, Theorem 6] provide an axiomatic foundation of smooth ambiguity preferences in a framework in which the set of (first-order) probabilities is a primitive of the model. In an earlier paper, Cerreia-Vioglio et al. [7] study the smooth ambiguity model within the Anscombe-Aumann setting under the umbrella provided by uncertainty averse preferences. Their Theorem 19 relates their uncertainty averse representation with a smooth ambiguity triplet $(u, \phi, \mu)$, where $\phi$ is concave.

4. CONSISTENCY OF THE MODEL WITH DISCRETE DATA

In Section 3, we provided a characterization of the CASAP model that leaves no doubt about whether or not a given preference relation belongs to this class. Naturally, our result there relies on the values of the certainty equivalent functional $I$ on its entire domain $\mathcal{B} = U^S$, where $U \subseteq \mathbb{R}$ is a nondegenerate interval. However, in many situations, such as experimental verification of subjects’ preferences in a laboratory, the set of points at
which the preferences are observed may be rather limited. Our objective in this section is to obtain conditions for datasets that have no accumulation points (or are finite) under which

1. the data are surely inconsistent with the hypothesis that the observations come from the studied class of preferences, or

2. the data do not contradict this hypothesis, in the sense that there exists at least one set of parameters of the model that can generate these datapoints.

As argued in the Introduction, this type of inquiry goes back to the seminal paper of Afriat [1] and is a subject of ongoing research. In particular, in the domain of ambiguity, Cerreia-Vioglio et al. [5, Theorem 2] provide a characterization of whether or not the values sampled at an arbitrary shaped finite set of points can come from a representation functional $I$ that is monotone and convex.

As is clearly evident from our Theorem 1, the characterizing properties of the CASAP functional are not so straightforward. To overcome the challenge presented by the model and move forward with our agenda, we make a simplifying assumption that $I$ is observed at a set of points that forms a square grid in $\mathbb{R}^n$.

We use the following notation to analyze the model in the discrete setting.

For $k \in \mathbb{N}$, let $(a_t)_{t \in \mathbb{Z}_+^k}$ be a multisequence and let $i : \mathbb{Z}_+^k \to \mathbb{N}$ be the enumerator of multi-indices by total degree then lexicographically, i.e., the successive numbers are assigned to $(0,0,\ldots,0)$, $(1,0,\ldots,0)$, $(0,1,\ldots,0)$, $\ldots$, $(0,0,\ldots,1)$, $(2,0,\ldots,0)$, $(1,1,\ldots,0)$, etc. For $r \in \mathbb{Z}_+$, we write $M_r(a)$ to denote the corresponding moment matrix — a symmetric $s \times s$ matrix, where $s = \begin{pmatrix} k+r \end{pmatrix}$, with elements

$$M_r(a)_{lm} := a_{i(t)} \quad \text{if} \quad l = i(t), \quad m = i(t'), \quad t, t' \in \mathbb{Z}_+^k.$$  

Note that $M_r(a)$ is constructed from elements $a_t$ for $|t| \leq 2r$. 

Date: 2016-03-31 Revision: 1.65
For any multisequence \((a_t)_{t \in \mathbb{Z}_{+}^n}\) we also define auxiliary multisequences 
\((a_t^{[j]})_{t \in \mathbb{Z}_{+}^n}\) and \((a_t^{[1-j]})_{t \in \mathbb{Z}_{+}^n}\) for \(j = 1, \ldots, k\), and 
\((a_t^{[C-\sigma]})_{t \in \mathbb{Z}_{+}^n}\) for \(C \in \mathbb{R}\) as

\[
\begin{align*}
  a^{[j]}(t_1, \ldots, t_k) &= a(t_1, \ldots, t_{j-1}, t_j + 1, t_{j+1}, \ldots, t_k), \\
  a^{[1-j]}(t_1, \ldots, t_k) &= a(t_1, \ldots, t_{j-1}, t_j + 1, t_{j+1}, \ldots, t_k), \\
  a^{[\sigma-C]}(t_1, \ldots, t_k) &= a(t_1 + 1, \ldots, t_k + 1) - Ca(t_1, \ldots, t_k).
\end{align*}
\]

### 4.1. Countably infinite data

The first result in this section provides a characterization of whether the preference data sampled at countably infinite number of points are consistent with the CASAP model.

**Theorem 2** Suppose that the functional \(I : \mathbb{R}^n \to \mathbb{R}\) can be computed by the CASAP equation (2) for some \(\lambda > 0\) and a probability measure \(\mu\) on \(\Delta\). Let \(F_\lambda : \mathbb{R}^{n-1} \to \mathbb{R}\) be constructed from \(I\) by (3), and let \((a_t)_{t \in \mathbb{Z}_{+}^{n-1}}\) be a multisequence defined by

\[
(4) \quad a_t = F_\lambda(ht)
\]

for some \(h > 0\). Then, the moment matrices \(M_r(a^{[j]}_t)\) and \(M_r(a^{[1-j]}_t)\) for all \(j = 1, \ldots, n - 1\) and \(M_r(a^{[\sigma-C]}_t)\) for \(C = e^{-h}\) are positive semidefinite for all \(r \in \mathbb{N}\).

Conversely, suppose that a real function \(I\) defined on the grid \(\{ht \mid t \in \mathbb{Z}_{+}^{n-1}\}\) for some \(h > 0\) is normalized and vertically invariant,\(^{12}\) \(F_\lambda\) is constructed from \(I\) by (3) for some \(\lambda > 0\), and \((a_t)_{t \in \mathbb{Z}_{+}^{n-1}}\) is given by (4). Furthermore, suppose that the moment matrices \(M_r(a^{[j]}_t)\) and \(M_r(a^{[1-j]}_t)\) for all \(j = 1, \ldots, n - 1\), and \(M_r(a^{[\sigma-C]}_t)\) for \(C = e^{-h}\), are positive semidefinite for all \(r \in \mathbb{N}\). Then, \(I\) can be extended to a functional \(\mathbb{R}_{+}^n \to \mathbb{R}\) such that (2) holds for some probability measure \(\mu\) on \(\Delta\).

---

\(^{12}\)Given \(h > 0\), a function \(I : \{ht' \mid t' \in \mathbb{Z}_{+}^{n-1}\} \to \mathbb{R}\) is vertically invariant if \(I(ht' + c1) = I(ht') + c\) for all \(t' \in \mathbb{Z}_{+}^{n-1}\) and \(c \in \mathbb{R}\) such that \(ht' + c1 \in \{ht \mid t \in \mathbb{Z}_{+}^{n-1}\}\).
Our proof strategy for this theorem is to establish a link between the problem at hand and the Moment Problem. In particular, we note that a (countable) sequence of observations is consistent with the CASAP model if and only if it is an exponential transformation of the sequence of moments of some probability measure supported on a particular compact set defined by polynomial inequalities.

**Proof of Theorem 2:** Suppose that \( I : \mathbb{R}^n \to \mathbb{R} \) can be computed by the CASAP equation (2) for some \( \lambda > 0 \) and a probability measure \( \mu \) on \( \Delta \). Let \( F_\lambda : \mathbb{R}^{n-1} \to \mathbb{R} \) be constructed from \( I \) by (3), and let \( (a_t)_{t \in \mathbb{Z}_+^{n-1}} \) be a multisequence defined by \( a_t = F_\lambda(ht) \) for some \( h > 0 \). Then, using the definition of \( F \) in (3), we have

\[
a_t = \int_P e^{-ht} \, d\mu(p) \quad \forall t \in \mathbb{Z}_+^{n-1},
\]

where \( P := \{(p_1, \ldots, p_{n-1}) \in \mathbb{R}_+^{n-1} : \sum_{k=1}^{n-1} p_k \leq 1\} \) is the set of all probability distributions on \( S \).

First, set \( x_i := e^{-hp_i} \) for all \( i = 1, \ldots, n-1 \) and consider a collection of polynomials of \( n-1 \) variables defined as

\[
g^{[j]}(x) = x_j, \quad g^{[1-j]}(x) = 1 - x_j \quad \text{for } j = 1, \ldots, n-1,
\]

\[
g^{[\sigma]}(x) = x_1 \ldots x_{n-1} - e^{-h}.
\]

By changing variables in (5), we obtain

\[
a_t = \int_{K_h} x^t \, d\mu(p) \quad \forall t \in \mathbb{Z}_+^{n-1},
\]

where

\[
K_h = \{ x \in \mathbb{R}^{n-1} : g_j(x) \geq 0, g_{1-j}(x) \geq 0 \text{ for } j = 1, \ldots, n-1, g_\sigma(x) \geq 0 \}.
\]

Hence, \( (a_t)_{t \in \mathbb{Z}_+^{n-1}} \) is a moment sequence of a Borel measure \( \mu \) with the support contained in \( K_h \), and note that the set \( K_h \) is compact.
Next, for each polynomial \( g \in \{g^{[j]}, g^{[1-j]}, g^{[\sigma]}\} \), let \( g_{\gamma} \) for \( \gamma \in \mathbb{Z}^{n-1}_+ \) denote its coefficients, and \( M_r(g,a) \) for \( r \in \mathbb{N} \) be a so called localizing matrix — a symmetric \( s \times s \) matrix, where \( s = \binom{n-1+r}{r} \), defined by

\[
M_r(g,a)_{lm} := \sum_{\gamma \in \mathbb{Z}^{n-1}_+} g_{\gamma} a_{\gamma+t+t'} \quad \text{if} \quad l = i(t), \quad m = i(t'), \quad t, t' \in \{0, 1, \ldots, r\}^{n-1}.
\]

It is easy to verify that these matrices coincide with the ones listed in the statement of our theorem:

\[
M_r(g_j, a) = M_r(a^{[j]}), \quad M_r(g_{1-j}, a) = M_r(a^{[1-j]}) \quad \text{for} \quad j = 1, \ldots, n-1;
M_r(g_\sigma, a) = M_r(a^{[\sigma-C]}) \quad \text{for} \quad C = e^{-h}.
\] (9)

Now, we are almost ready to conclude that the matrices \( M_r(a^{[j]}) \) and \( M_r(a^{[1-j]}) \) for all \( j = 1, \ldots, n-1 \) and \( M_r(a^{[\sigma-C]}) \) for \( C = e^{-h} \) are positive semidefinite for all \( r \in \mathbb{N} \) by applying the characterization of moment sequences given by Cominetti, Facchinei, and Lasserre [12, Theorem 2.14(b)]. For that, we need to verify the requirement of that theorem on the set \( K_h \).

Observe that the polynomial \( Q \) defined as \( Q = \sum_{j=1}^{n-1} x_j^2 g^{[j]} + \sum_{j=1}^{n-1} (x_j + 2)^2 g^{[1-j]} \) is a linear combination of polynomials defining \( K_h \) multiplied by squares of polynomials, and, since \( Q \) can be rewritten as \( Q = \sum_{j=1}^{n-1} (4 - 3x_j^2) \), the set \( \{x \in \mathbb{R}^{n-1} : Q(x) \geq 0\} \) is compact. Hence, the conditions of the theorem are satisfied, and the “if” part of our theorem is now proven.

Conversely, suppose that a real function \( I \) defined on the grid \( \{ht \mid t \in \mathbb{Z}^{n-1}_+\} \) for some \( h > 0 \) is normalized and vertically invariant, \( F_\lambda \) is constructed from \( I \) by (3) for some \( \lambda > 0 \), and \( (a_t)_{t \in \mathbb{Z}^{n-1}_+} \) is given by (4). Furthermore, suppose that the moment matrices \( M_r(a^{[j]}) \) and \( M_r(a^{[1-j]}) \) for all \( j = 1, \ldots, n-1 \), and \( M_r(a^{[\sigma-C]}) \) for \( C = e^{-h} \), are positive semidefinite for all \( r \in \mathbb{N} \). Then, given the equalities in (9), we can apply the “if” part of Cominetti et al. [12, Theorem 2.14(b)] to obtain that \( (a_t)_{t \in \mathbb{Z}^{n-1}_+} \) is the moment sequence of a measure \( \mu \) with support contained in \( K_h \), that is, \( a_t = \int_{K_h} x^t d\mu(x) \). The normalization of \( I \) ensures that \( \mu \) is a probability measure. By setting \( x_i = e^{-h p_i} \) for all \( i = 1, \ldots, n-1 \), we have
\[ a_t = \int_P e^{-hpt} \, d\mu(p). \] Hence, \( F_\lambda \) can be extended to a function \( \mathbb{R}^{n-1} \to \mathbb{R} \) by letting \( F_\lambda(\varphi) = \int_P e^{-p\varphi} \, d\mu(p) \) for all \( \varphi \in \mathbb{R}^{n-1} \); in turn, we can let

\[ I(\varphi_1, \ldots, \varphi_n) = -\frac{1}{\lambda} \ln \left( \int_{p \in \Delta} e^{-\lambda \varphi_n - \sum_{j=1}^{n-1} \lambda p_j (\varphi_j - \varphi_n)} \, d\mu(p) \right) \]

in agreement with the values of \( I \) on \( \{ht \mid t \in \mathbb{Z}^{n-1}_+ \} \) that are given. Q.E.D.

4.2. Finite data

The theorem of the previous subsection lists the conditions for a collection of values \( \{I(ht) \mid t \in \mathbb{Z}^n_+\} \) to be consistent with the CASAP model. Those conditions constitute a sequence of requirements of positive semidefiniteness of certain matrices for each \( r \in \mathbb{N} \). Moreover, each of those matrices uses only a finite number of “datapoints” that grows with \( r \) — namely, \( M_r(a[j]) \) and \( M_r(a^{[1-j]}) \) (where \( j = 1, \ldots, n-1 \)) are constructed using the values of \( I(ht) \) for \( t \in \mathbb{Z}^n_+ \) such that \(|t| \leq 2r + 1\), and \( M_r(a^{[\sigma-C]}) \) uses the values for \(|t| \leq 2r + n - 1\). Therefore, the conditions of Theorem 2 can still be used if the analyst has only finitely many datapoints. Indeed, if one of these matrices for some \( r \in \mathbb{N} \) is not positive semidefinite, then one can immediately reject the hypothesis that the data come from the CASAP model.

On the flip side of this test, the analyst may be interested to know whether, for a given finite collection of observed datapoints, there exists a prior \( \mu \) on \( \Delta \) that generates it. Unfortunately, the answer to this question is not immediate: While violations of positive semidefiniteness of one of the abovementioned matrices imply that such \( \mu \) does not exist, the positive semidefiniteness of finitely many matrices cannot guarantee the existence of a suitable prior.\(^{13}\)

\(^{13}\)This is a general property of truncated moment problems, which ask for conditions on a finite (multi-)sequence to be represented as a sequence of moments of some measure up to a finite order. Stochel [32, Theorem 4] shows that the truncated moment problem is more general than the full moment problem. An extensive study of truncated moment problems.
The next result provides an additional condition that guarantees the existence of a prior compatible with a given finite dataset.\(^{14}\)

**Proposition 3** Fix \( r \in \mathbb{N} \), \( r \geq \lceil \frac{n-1}{2} \rceil \), and \( h > 0 \), and suppose that a function \( I : \{ht \mid t \in \mathbb{Z}_n^+, |t| \leq 2r \} \to \mathbb{R} \) is normalized and vertically invariant, \( F_\lambda \) is constructed from \( I \) by (3) for some \( \lambda > 0 \), and \((a_t)_{t \in \mathbb{Z}_n^+, |t| \leq 2r} \) is given by (4). Furthermore, suppose that the moment matrices \( M_q(a) \), \( M_q(a[j]) \), and \( M_q(a[1-j]) \) for all \( j = 1, \ldots, n - 1 \), and \( M_q(a^{[\sigma-C]}) \), where \( C = e^{-h} \) and \( q = r - \lceil \frac{n-1}{2} \rceil \), are positive semidefinite, and \( \text{rank} M_r(a) = \text{rank} M_q(a) \). Then, there exists a probability measure \( \mu \) on \( \Delta \) such that the representation (2) of \( I \) holds.

**Proof:** The claim follows from Curto and Fialkow [13, Theorem 1.1].\(^{15}\)

Using Curto and Fialkow’s terminology, \( M_{q+1}(a) \) is a flat extension of \( M_q(a) \), and, as follows from their Theorem 2.19, \( M_{q+2}(a) \), \( \ldots \), \( M_r(a) \) are successive flat extensions that are unique given \( M_{q+1}(a) \). Since \( M_q(a[j]) \) and \( M_q(a[1-j]) \) for all \( j = 1, \ldots, n - 1 \), and \( M_q(a^{[\sigma-C]}) \), where \( C = e^{-h} \), are the localizing matrices for the system of polynomials (6) that defines the support (8) of the measure we seek (see the proof of Theorem 2), and those matrices are positive semidefinite by assumption, then the conditions of that theorem hold.

We conclude that there exists a measure \( \mu \) such that the integral representation (7) holds, which, by changing variables, gives (2). The normalization of \( I \) ensures that \( \mu \) is a probability measure. \( Q.E.D. \)

Proposition 3 suggests the following procedure for analyzing subjects’ problems can be found in the works of Curto and Fialkow — see, e.g., Curto and Fialkow [13; 14].

\(^{14}\)Hence, this condition also guarantees that a given collection \( \{I(ht) \mid t \in \mathbb{Z}_n^+, |t| \leq 2r \} \), where \( r \in \mathbb{N} \), can be extended to \( \{I(ht) \mid t \in \mathbb{Z}_n^+ \} \) such that the conditions of Theorem 2 hold.

\(^{15}\)For a re-statement of this theorem, see Cominetti et al. [12, Theorem 2.17].
preferences. The analyst starts with setting the size of the sample he wishes to study by fixing some \( r \in \mathbb{N} \); then, he chooses some \( h_0 > 0 \), and collects the data \( \{ I(h_0 t) \mid t \in \mathbb{Z}^+, |t| \leq 2r \} \). The first property to check is vertical invariance. Then, he can set \( h := h_0 \lambda \) and construct matrices \( M_q(a), M_q(a_{[j]}), \) and \( M_q(a_{[1-j]}) \) for all \( j = 1, \ldots, n-1 \), and \( M_q(a_{[\sigma-C]}) \), where \( C = e^{-h} \) and \( q = r - \lceil \frac{n-1}{2} \rceil \), as functions of \( \lambda \). If, for each \( \lambda > 0 \), at least one of these matrices is not positive semidefinite, then the sample is not consistent with the CASAP model.\(^\text{16}\) If all matrices are positive semidefinite for some set of values of \( \lambda \), the analyst proceeds by checking whether \( \text{rank} M_r(a) = \text{rank} M_q(a) \) for one of these values. If this condition holds, then the data are confirmed to be consistent with the model, and it is possible to find a measure \( \mu \) on \( \Delta \) that generates the data. If the rank condition does not hold, then, the available data are inconclusive and do not allow the analyst to either reject or confirm the hypothesis. To obtain a conclusive answer, the analyst may consider increasing the size of his dataset.\(^\text{17}\)

We illustrate the above procedure with two examples.

**Example 1** Let \( n = 2 \). Suppose that the analyst obtains the preference data \( \{ I(\phi_1, \phi_2) \mid \phi_1, \phi_2 \in \{0, \ldots, 5\}, \phi_1 + \phi_2 \leq 4 \} \) (hence, \( h_0 = 1 \) and \( r = 2 \) in this example). Assume also that \( I \) satisfies the vertical invariance property.

To make the example concrete, let the data be

\[
I(\cdot, 0) = (0, 0.53437444, 1.0569788, 1.5738683, 2.0941030).
\]

after reducing its dimensionality. To illustrate the construction of the mo-

\(^\text{16}\)Note that the positive semidefiniteness of \( M_r(a) \) is not explicitly mentioned in Theorem 2, however, it is also clearly necessary for the representation (1)–(3)–(4).

\(^\text{17}\)If increasing the size of the dataset is not possible, then, as a last resort, the analyst can try to check if there exist real numbers \( (a_t)_{2r < |t| < 2r'} \) for some \( r' > r \) that extend the observed sequence and such that the rank and positive semidefiniteness conditions hold. This can also provide evidence that the observed data may arise from the CASAP model.
ment matrices, note that the matrix $M_2(a)$ in this case is

$$
\begin{pmatrix}
1.000000 & e^{-0.53437444\lambda} & e^{-1.05697888\lambda} \\
e^{-0.53437444\lambda} & e^{-1.05697888\lambda} & e^{-1.57386833\lambda} \\
e^{-1.05697888\lambda} & e^{-1.57386833\lambda} & e^{-2.09410300\lambda}
\end{pmatrix}
$$

and $M_1(a^{[1-1]})$ is

$$
\begin{pmatrix}
1.000000 & e^{-0.53437444\lambda} & -e^{-1.05697888\lambda} + e^{-0.53437444\lambda} \\
-e^{-1.05697888\lambda} + e^{-0.53437444\lambda} & e^{-1.57386833\lambda} & -e^{-1.05697888\lambda} + e^{-1.57386833\lambda}
\end{pmatrix}.
$$

Now we observe that $\det M_2(a) < 0$ for all $\lambda > 0$: Indeed, $\det M_2(a) < -e^{-3.1477367\lambda}(1 + z^{226} + z^{347} - 2z^{261} - z^{50})$, where $z = e^{-0.0033451\lambda/50}$, and it can be verified that the polynomial term in the expression for the determinant is positive for all $0 < z < 0.999938$ and, hence, for $\lambda \geq 1$. For $0 < \lambda < 1$, the sign of the determinant can be inferred from its Taylor expansion around $\lambda = 0$. We conclude that the given data are not consistent with the CASAP model for any parameters $\lambda$ and $\mu$.

A certain difficulty in the above analysis arises from the need to verify the positive semidefiniteness of matrices for all $\lambda \in \mathbb{R}_{++}$. However, in certain cases the analyst may have some exogenous constraints on this parameter in representation (1). For instance, given the interpretation of the function $\phi$ in Klibanoff et al. [26], the analyst may believe that, for a reasonable decision maker, $\lambda$ should not exceed, for example, twenty (or one hundred). If the range of the values of $\lambda$ is bounded, then testing the hypothesis about preferences becomes analytically simpler, and can also be aided by conventional numerical methods.

**Example 2** Let $n = 3$. Assume that the observed data satisfy the vertical invariance property and, after reducing the dimensionality, $\{I(\varphi_1, \varphi_2, 0) |$
\( \varphi_1, \varphi_2 \in \{0, \ldots, 4\}, \varphi_1 + \varphi_2 \leq 4 \) is given by

\[
\begin{pmatrix}
0 & 0.2684203989 & 0.5319232006 & 0.7903983908 & 1.043775559 \\
0.3139951847 & 0.5853962381 & 0.8523665482 & 1.114790653 \\
0.6099489358 & 0.8832835339 & 1.152679453 \\
0.8878379286 & 1.162086330 \\
1.148017719
\end{pmatrix},
\]

where the first coordinate is the row number and the second coordinate is the column number.

Again, to illustrate, \( M_2(a) \) in this case is

\[
\begin{pmatrix}
1 & e^{-0.3139951847t} & e^{-0.2684203989t} & e^{-0.6099489358t} & e^{-0.5853962381t} & e^{-0.5319232006t} \\
e^{-0.3139951847t} & e^{-0.6099489358t} & e^{-0.5853962381t} & e^{-0.8878379286t} & e^{-0.8832835339t} & e^{-0.8523665482t} \\
e^{-0.2684203989t} & e^{-0.5853962381t} & e^{-0.5319232006t} & e^{-0.8832835339t} & e^{-0.8523665482t} & e^{-0.7903983908t} \\
e^{-0.6099489358t} & e^{-0.8878379286t} & e^{-0.8832835339t} & e^{-1.148017719} & e^{-1.162086330t} & e^{-1.152679453t} \\
e^{-0.5853962381t} & e^{-0.8832835339t} & e^{-1.148017719} & e^{-1.162086330t} & e^{-1.152679453t} & e^{-1.114790653t} \\
e^{-0.5319232006t} & e^{-0.8523665482t} & e^{-0.7903983908t} & e^{-1.152679453t} & e^{-1.114790653t} & e^{-1.043775559t}
\end{pmatrix}
\]

and \( M_1(a^{[1-2]}) \) is

\[
\begin{pmatrix}
1 - e^{-0.2684203989t} & -e^{-0.5853962381t} + e^{-0.3139951847t} & -e^{-0.5319232006t} + e^{-0.2684203989t} \\
-e^{-0.5853962381t} + e^{-0.3139951847t} & -e^{-0.8832835339t} + e^{-0.6099489358t} & -e^{-0.8523665482t} + e^{-0.5853962381t} \\
-e^{-0.5319232006t} + e^{-0.2684203989t} & -e^{-0.8523665482t} + e^{-0.5853962381t} & -e^{-0.7903983908t} + e^{-0.5319232006t}
\end{pmatrix}.
\]

To illustrate the next step — finding the suitable values of \( \lambda \) — we plot the determinants of matrices \( M_1(a) \) and \( M_2(a) \) as functions of \( \lambda \) in Figure 1. In order to apply Proposition 3 with \( r = 2 \) and \( q = r - \left[ \frac{a-1}{2} \right] \equiv 1 \), it can be verified that, for \( \lambda = 1 \), all matrices \( M_1(a) \), \( M_1(a^{[1]}) \), \( M_1(a^{[2]}) \), \( M_1(a^{[1-1]}) \), \( M_1(a^{[1-2]}) \), and \( M_1(a^{[\sigma^{-C}]}) \) are positive semidefinite. Moreover, for \( \lambda = 1 \), we also have \( \text{rank } M_2(a) = \text{rank } M_1(a) = 3 \). Therefore, it is confirmed that the observed data are consistent with the CASAP model, and there exists some probability measure \( \mu \) on \( \Delta \) that generates the data.
The results of Curto and Fialkow [13], in particular, Theorems 1.1–1.2, also provide a way to obtain the generating measure $\mu$ (and, hence, a CASAP representation for $I$) constructively. We compute a basis in the subspace $\{v \in \mathbb{R}^6: M_2(a)v = 0\}$:
The corresponding system of polynomials
\[
\begin{align*}
0.5571224114x_1 - 0.5910356618x_2 - 0.2670338212x_1^2 - 0.1741580493x_1x_2 + 0.4885216242x_2^2 &= 0 \\
-0.04488589616x_1 + 0.1644972343x_2 + 0.4686681419x_1^2 - 0.841862059x_1x_2 + 0.2062628475x_2^2 &= 0 \\
-0.3044150965x_1^2 - 0.1557632068x_2x_1 + 0.5856851231x_1 - 0.2430028164x_2^2 + 0.5346593118x_2 - 0.4417366416 &= 0
\end{align*}
\]
has three real solutions: \(x^{(1)} = (0.88250, 0.77880)\), \(x^{(2)} = (0.71653, 0.71653)\), \(x^{(3)} = (0.60653, 0.84648)\). These points constitute the support of the generating measure in variables \(x_i = e^{-h_0\lambda p_i}\). By applying the inverse — the function \(p = -\ln(x)\) — to these values, we obtain the support of the generating measure on the \((n-1)\)-dimensional simplex as \((0.12500, 0.25000), (0.33333, 0.33333),\) and \((0.50000, 0.16667)\). The corresponding weights \((w_1, w_2, w_3)\) can be computed by solving
\[
\begin{pmatrix}
1 & 1 & 1 \\
x_1^{(1)} & x_2^{(2)} & x_3^{(3)} \\
x_1^{(1)} & x_2^{(2)} & x_3^{(3)}
\end{pmatrix}
\begin{pmatrix}
w_1 \\
w_2 \\
w_3
\end{pmatrix}
= \begin{pmatrix}
e^{-f(0,0)} \\
e^{-f(0,1)} \\
e^{-f(0,1)}
\end{pmatrix},
\]
which gives the values of \((0.25, 0.5, 0.25)\). We conclude that the measure \(\mu \) on \(\Delta = \{(p_1,p_2,p_3) \in \mathbb{R}^3_+ : p_1 + p_2 + p_3 = 1\} \) in Representation (2) is \(\mu = \frac{1}{4}\delta((1/8,1/4,5/8)) + \frac{1}{2}\delta((1/3,1/3,1/3)) + \frac{1}{4}\delta((1/2,1/6,1/3))\), where \(\delta\) denotes the Dirac measure.

In the above example, if the decision maker’s preferences are truly CASAP with the second-order prior \(\mu\) as determined, then the additional data will
not reveal any new information. In practice, however, the analyst may proceed without the prior knowledge of $\mu$ and, given also the possibility of measurement errors, will benefit from as much data as he can collect.

As a final remark, we note that our results and the procedure outlined above do not yield a unique value of $\lambda$ in the case the test confirms that the data originate from a CASAP representation. In the setting of this section, $\lambda$ cannot be identified uniquely for a natural reason — the finite amount of data. However, beyond that, the CASAP representation is not unique if the analyst is limited to observing only choices among first-order acts. For instance, the CASAP preference relation with $\lambda = 1$ and $\mu$ as computed in Example 2 can be described equally well by another CASAP representation with $\lambda = 2$ and $\mu = \frac{\delta(1/8,1/4,5/8)}{10} + \frac{\delta(1/3,1/3,1/3)}{10} + \frac{\delta(1/2,1/6,1/3)}{10} + \frac{\delta(11/14,7/24,23/48)}{4} + \frac{\delta(5/16,5/24,23/48)}{8} + \frac{\delta(5/12,1/4,1/3)}{4}$. With respect to this type of non-uniqueness, we note that if the analyst takes the value of $\lambda$ for which the rank condition holds for the smallest value of $r$, then our procedure delivers the measure $\mu$ with the support of the smallest cardinality.

APPENDIX: PROOF OF THEOREM 1

In this appendix we prove a more general version of Theorem 1 in which we drop the unboundedness assumption. In order to deal with the case in which the range of $u$ is bounded, we will work with an extension of $I$ from $B$ to a larger subset $D \subseteq \mathbb{R}^S$. Let

$$D = \{ \varphi - x \mathbb{1}_S \mid \varphi \in B, x \in u(X) \}.$$ 

Since $\succsim$ is assumed to be variational, $I$ has a unique extension to $D$: for an element $\varphi - x \mathbb{1}_S \in D$, it assigns the value $I(\varphi) - x$. Extending Lemma 28 in Maccheroni et al. [28], it can be shown that vertical invariance of the preference relation implies that this value does not depend on the choice of $\varphi \in B$ and $x \in u(X)$. Assuming without loss of generality that $0 \in u(X)$,
it can be seen that, by taking \( x = 0 \), the newly defined functional agrees with \( I \) on \( \mathcal{B} \). Therefore, we will continue to use the same notation \( I \) for the extended functional.

**Theorem 4**  Let \( I : \mathcal{D} \to \mathbb{R} \) be a normalized, monotone, vertically invariant, quasiconcave, and continuous functional. The following conditions are equivalent:

(a) There exist \( \lambda > 0 \) and a probability measure \( \mu : \Delta \to [0,1] \) such that

\[
I(\varphi) = -\frac{1}{\lambda} \ln \left( \int_{\Delta} e^{-\lambda p \varphi} \, d\mu(p) \right),
\]

for all \( \varphi \in \mathcal{D} \).

(b) There exists \( \theta > 0 \) such that \( F_\theta : \mathbb{R}^{n-1} \to \mathbb{R} \) defined as

\[
F_\theta(\varphi) := e^{-\theta I(\frac{1}{\theta} \varphi, 0)},
\]

for all \( \varphi \in \mathbb{R}^{n-1} \), satisfies the following conditions:

(i) \( (-1)^{|t|} D^t F_\theta(\varphi) \geq 0 \) for all \( t \in \mathbb{Z}_+^{n-1} \) and \( \varphi \in \mathbb{R}_+^{n-1} \);

(ii) \( F_\theta \) can be extended to an entire holomorphic function \( \mathbb{C}^{n-1} \to \mathbb{C} \);

(iii) \( |F_\theta(\zeta_1, \ldots, \zeta_{n-1})| \leq C e^{\max(-\Re \zeta_1, \ldots, -\Re \zeta_{n-1}, 0)} \) for some \( C > 0 \) and for all \( (\zeta_1, \ldots, \zeta_{n-1}) \in \mathbb{C}^{n-1} \).

**Proof:** Suppose that \( I : \mathcal{D} \to \mathbb{R} \) is a normalized, monotone, vertically invariant, quasiconcave, and continuous functional. For each \( \theta > 0 \), let \( F_\theta : \mathbb{R}^{n-1} \to \mathbb{R} \) be defined as \( F_\theta(\varphi) := e^{-\theta I(\frac{1}{\theta} \varphi, 0)} \).

(a) implies (b). Suppose there exist a scalar \( \lambda > 0 \) and a probability measure \( \mu : \Delta \to [0,1] \) such that \( I : \mathcal{D} \to \mathbb{R} \) can be represented as in (10).

Then, by setting \( \theta = \lambda \), it follows that

\[
F(\varphi) = \int_{\mathbb{P}} e^{-\theta p \varphi} \, d\mu(p),
\]

for all \( \varphi \in \mathbb{R}^{n-1} \), where \( \mathbb{P} \in \mathbb{R}_+^{n-1} \) is the set of all probability distributions on \( S \), \( \mathbb{P} := \{ (p_1, \ldots, p_{n-1}) \in \mathbb{R}_+^{n-1} : \sum_{k=1}^{n-1} p_k \leq 1 \} \). Now it is evident that \( F \) is
a restriction to \( \mathbb{R}^{n-1} \) of the Laplace transform of a probability measure \( \mu \) with support \( P \).

Probability measure \( \mu \) also can be thought of as a distribution (a generalized function) of order 0, as it induces a linear functional on the space of infinitely differentiable test functions with compact support, and this functional is continuous in the sup-norm. Also, \( \mu \) has a compact support. Therefore, its Fourier transform \( \mathcal{F}(\mu) \) is an entire function \( \mathbb{C}^{n-1} \rightarrow \mathbb{C} \) (see Donoghue [15, §30]). Since the Laplace transform \( \mathcal{L}(\mu)(z) = (2\pi)^{n/2} \mathcal{F}(\mu)(-iz) \) for all \( z \in \mathbb{C}^{n-1} \), it follows that \( F \) can be extended to an entire function. This proves property (ii). Moreover, as follows from the Taylor series expansion theorems for holomorphic functions, this extension is unique.

Define a function \( H : \mathbb{R}^{n-1} \rightarrow \mathbb{R} \) as

\[
H(\eta_1, \ldots, \eta_{n-1}) := \max(\eta_1, \ldots, \eta_{n-1}, 0).
\]

Note that \( H \) is the support function for the set \( P \):

\[
H(\eta) = \max_{p \in P} pq \quad \text{for all } \eta \in \mathbb{R}^{n-1},
\]

so by the Paley-Wiener Theorem (Donoghue [15, §43, p. 211]) the Fourier transform of \( \mu \) satisfies the inequality

\[
|\mathcal{F}(\mu)(z)| \leq Ce^{H(\text{Im}z)}
\]

for some \( C > 0 \), which translates into property (iii) for the holomorphic extension of \( F \).

Finally, since \( F \) is the Laplace transform of a probability measure, then property (i) follows from the multidimensional extension of Bernstein Theorem (Choquet [9, §43–47]; Choquet [10, Theorem 10]; for a modern exposition, see Thomas [34, §4]).

(b) implies (a). Suppose that there exists a scalar \( \theta > 0 \) such that \( F_\theta \) satisfies conditions (i)–(iii). We will show that there exist a scalar \( \lambda \) and a
probability distribution $\mu$ on $\Delta$ such that the functional $I : D \rightarrow \mathbb{R}$ can be represented as in (10).

First, from the definition of certainty equivalent, $F(0) = 1$. By the multidimensional version of Bernstein Theorem, it follows that there exists a probability measure $\mu$ on $\mathbb{R}^{n-1}$ such that

$$F(\varphi) = \int_{\mathbb{R}^{n-1}_+} e^{-\nu \varphi} d\mu(p)$$

for all $\varphi \in \mathbb{R}^{n-1}$.

Moreover, for $z \in \mathbb{C}^{n-1}$, $F(iz)$ is an entire function; using property (iii) and the converse claim of the Paley-Wiener Theorem (Donoghue [15, §43, p. 213]), it follows that $F(iz)$ is the Fourier transform of some distribution $\mu'$ with support contained in $P$. Since the Fourier transform is an injection on the class of distributions that includes all measures of finite total mass (Donoghue [15, §28–29]), it has to be that $\mu = \mu'$ and, hence, the support of $\mu$ is contained in $P$.

Using the definition of $F$ in (11), we obtain

$$I(\varphi_1, \ldots, \varphi_{n-1}, 0) = -\frac{1}{\lambda} \ln \left( \int_{\Delta} e^{-\sum_{k=1}^{n-1} \lambda p_k \varphi_k} d\mu(p) \right)$$

for all $(\varphi_1, \ldots, \varphi_{n-1}) \in \mathbb{R}^{n-1}$. Since $I$ is vertically invariant, it follows that

$$I(\varphi_1, \ldots, \varphi_{n-1}, \varphi_n) = \varphi_n - \frac{1}{\lambda} \ln \left( \int_{\Delta} e^{-\sum_{k=1}^{n-1} \lambda p_k \varphi_k - \lambda(1-\sum_{k=1}^{n-1} p_k) \varphi_n} d\mu(p) \right) =$$

$$= -\frac{1}{\lambda} \ln \left( \int_{\Delta} e^{-\sum_{k=1}^{n} \lambda p_k \varphi_k} d\mu(p) \right)$$

for all $(\varphi_1, \ldots, \varphi_n) \in D$. Q.E.D.

**REFERENCES**


1575–1605.

