Conditional Preference for Flexibility: Eliciting Beliefs from Behavior*

Philipp Sadowski†

January 18, 2009

Abstract

Following Kreps (1979), we consider a decision maker with uncertain beliefs about her own future taste. This uncertainty leaves the decision maker with preference for flexibility: When choosing among menus containing alternatives for future choice, she weakly prefers larger menus. Existing representations accommodating this choice pattern cannot distinguish tastes (indexed by a subjective state space) and beliefs (a probability measure over the subjective states) as different concepts, making it impossible to relate parameters of the representation to choice behavior. We allow choice among menus to depend on exogenous states, interpreted as information. Our axioms yield a representation that uniquely identifies beliefs, provided the impact of information on choice is *rich*. The result is suggested as a choice theoretic foundation for the assumption, commonly made in the incomplete contracting literature, that contracting parties, who know each other's ranking of contracts, also share beliefs about each others future tastes in the face of unforeseen contingencies.

1. Introduction

Choice among opportunity sets or menus, which have alternatives for future choice as elements, allows attitudes towards the future to be reflected in static choice. In particular, uncertain beliefs about future tastes suggest preference for flexibility: a decision maker (DM) may prefer not to constrain future choice today. Choice among menus will therefore favor larger menus. Let $\succ$ be DM's preference ranking of menus and let A and B be menus. Then $A \cup B \succcurlyeq A$. This property is called *Monotonicity.*

*Preliminary. I thank Roland Benabou and Wolfgang Pesendorfer for their invaluable advice and support. I am also grateful to Faruk Gul, Mark Machina, Eric Maskin and Tymon Tatur for helpful lessons and discussions.

†Department of Economics, Duke University, p.sadowski@duke.edu

Kreps (1979) assumes Monotonicity and establishes a representation theorem for preference for flexibility. The representation can be written as

$$V(A) = \sum_{s \in S} \mu(s) \left(\max_{x \in A} u_s(x) \right),$$

where a future taste, $s \in S$, is captured by the utility function over items, u_s. Given taste s, a menu is evaluated according to its most preferred element. The probability distribution μ on S is interpreted as DM's beliefs that particular tastes occur.

Even though the representation suggests an appealing interpretation of choice satisfying Monotonicity, its parameters cannot be related to choice behavior: consider a different probability distribution $\widehat{\mu}(s)$ on S and rescaled utilities

$$\widehat{u}_s(x) = u_s(x) \frac{\mu(s)}{\widehat{\mu}(s)}.$$

Then

$$\sum_{s \in S} \mu(s) \left(\max_{x \in A} u_s(x) \right) \equiv \sum_{s \in S} \widehat{\mu}(s) \left(\max_{x \in A} \widehat{u}_s(x) \right).$$

Observation of choice does not allow the distinction of beliefs and utilities.

Our work combines preference for flexibility with the arrival of publicly observable information $i \in I$. Timing is the following: in period 1, DM chooses among menus. Between periods 1 and 2 information arrives and subjective uncertainty about taste resolves. In period 2 DM chooses from the menu determined in period 1. The menu chosen in period 1 may be contingent on information. In analogy to the terminology in the classical work by Savage (1954,) we call the mapping from information to menus an act.[1] Acts are objects of choice.

We have in mind a representation, where DM has an expectation about the arrival of information and the only impact of information is the updating of beliefs about her taste. We answer two main questions:

- Does the assumption that information only leads to updated beliefs constrain period 1 choice behavior? Theorem 1 establishes that it does not.
- Can the parameters of such a representation be related to choice behavior in the sense that beliefs and utilities are conceptually distinct? Theorem 2 identifies an axiom on choice behavior, which is equivalent to the unique identification of beliefs and utilities.

[1] The notion of "contingent menus" appears in Epstein (2006). Nehring (1999) calls acts with menus as outcomes "opportunity acts".

Example: As an example of an act, consider a contract between a retailer and a supplier, which conditions on information about consumer confidence. Consumer confidence is either high (H) or low (L). This information becomes publicly available before the retailer's order is finalized. The idiosyncratic demand facing the retailer depends on indescribable contingencies and is either high (h) or low (l), allowing the sale of quantities q_h and q_l, respectively. Idiosyncratic demand may be correlated with consumer confidence.

A contract g specifies whether the retailer retains flexibility and can choose between the two quantities (in this case the contract can be viewed as incomplete), or whether he is committed to order a particular one. This may depend on the information about consumer confidence: $g(i) \subseteq \{q_h, q_l\}$ for $i \in \{H, L\}$.

The retailer maximizes profits and a shortage or an oversupply are costly: facing high demand, he values q_h at $v_h(q_h)$ and q_l at

$$v_h(q_l) < v_h(q_h).$$

Facing low demand,

$$v_l(q_l) > v_l(q_h).$$

The relative value of those utilities is private, unless it is revealed through choice. The retailer can assign subjective probabilities $\mu(h|H)$ and $\mu(h|L)$, based on his knowledge about the underlying indescribable contingencies and conditional on information. Choosing among contracts, the retailer's objective function is

$$V(g) = \sum_{i \in \{H,L\}} \phi(i) \left[\sum_{s \in \{h,l\}} \mu(s|i) \max_{q \in g(i)} (v_s(q)) \right],$$

where ϕ is the objective probability distribution over information, $i \in \{H, L\}$.[2]

Our domain allows any finite (or metric) information space, I, and any finite prize space. While in the example h and l have a clear interpretation, the taste space S is endogenous and may in general only have an interpretation in terms of DM's period 2 preferences. Kreps (1992) points out that an endogenous taste space naturally accounts for contingencies, that are not just indescribable, but unforeseen, at least by the observer.

As the taste space is endogenous, it cannot simply be assumed to be finite. To make sense of a continuous taste space, we follow Dekel, Lipman and Rustichini (2001), henceforth

[2] Objective probabilities correspond to observed frequencies.

DLR.[3] As they do, we consider lotteries over the finite prize space as items on a menu. An act g is then a mapping from information to contingent menus, $g(i)$, of lotteries over prizes.

Theorem 1 identifies axioms on choice over acts, under which it can be represented by a straight forward generalization of the representation in the example: if I is finite, then

$$V(g) = \sum_{i \in I} \phi(i) \left[\int_S \left(\max_{\alpha \in g(i)} U_s(\alpha) \right) d\mu(s|i) \right]$$

represents choice, where $\mu(s|i)$ is a subjective probability measure on the taste space, S. $\mu(s|i)$ is interpreted as the belief that taste s occurs, conditional on information i. Tastes are realized von Neumann-Morgenstern utility functions, U_s. Information only leads to updated beliefs, so the scaling of U_s does not depend on information.[4] The distribution ϕ corresponds to objective probabilities over information.

Theorem 2 identifies an additional axiom on choice over acts that requires ex post preferences to depend enough on information, and under which the parameters of the representation V are unique.[5] Clearly there are other representations of $\succ$, in particular any monotone transformation of V. Our uniqueness result is analogous, for example, to the uniqueness in von Neumann-Morgenstern's (1944) work, where there are many possible representations, but only one has the functional form of expected utility.

The usual choice theoretic approach is to take our representation as a description only of period 1 choice, where DM behaves in period 1, *as if* she held beliefs about possible tastes that might govern period 2 choice. Theorem 2 relates beliefs, which are parameters of the representation, to period 1 choice behavior.

However, our context features a second period, in which another choice is made. If DM has knowledge about the contingencies underlying the formation of her taste, then the natural inductive step is to employ the beliefs about future tastes to forecast period 2 choice behavior. Doing so implies that the representation V is interpreted as a map of the decision making process. This cuts two ways: on the one hand it requires evaluating the appropriateness of the representation for a particular application,[6] on the other hand the model can be refuted, if its forecasts do not agree with observation.

Being able to forecast behavior is relevant in strategic situations. For example, if two

[3] Dekel, Lipman, Rustichini and Sarver (2007), henceforth DLRS, is a relevant corrigendum.

[4] The scaling of each vNM utility is endogenous, but it is fixed across information.

[5] In order to make DLR's domain more versatile, Ozdenoren (2002) also allows for objective states of the world, which correspond to our information. The work assumes that ex-post preferences over menus are independent of the objective state.

[6] The three main modelling choices are: The expected utility criterion is used to evaluate uncertainty, information impacts only beliefs and, ultimately, only the chosen item on a menu generates utility.

parties write a contract in the face of unforeseen or indescribable contingencies, which are relevant for future utility-payoffs, then the optimal contract is incomplete as in the example. Incomplete contracting models usually have to assume that both parties *share* a common belief about the probability of future utility-payoffs, when writing the contract.

However, if those contingencies are more relevant for one party, it seems natural that this same party can also foresee them better, leading to asymmetric information. In a survey on incomplete contracts, Tirole (1999) speculates that *"... there may be interesting interaction between "unforeseen contingencies" and asymmetric information. There is a serious issue as to how parties [...] end up having common beliefs ex ante."*

Without doing the game theoretic complexity of the contracting problem justice, beliefs that are elicited from a party's preferences over contracts give choice theoretic substance to the assumption of common beliefs.[7]

Section 1.1 give a more detailed overview over our results. Section 1.2 demonstrates the generic identification of beliefs in the example above. Section 2 lays out the model and establishes Theorems 1 and 2 for a finite as well as a metric information space.[8] Section 3 contains Theorem 3, which combines the two results. Section 4 elaborates the example to consider both contracting parties and comments on generalizations. Section 5 concludes.

1.1. Overview of Results

In general probabilities of information are subjective. Throughout the paper we use π to denote subjective probabilities of information, while ϕ is used in the special case, where subjective probabilities coincide with objective probabilities. The representation V then becomes

$$V(g) = \sum_{i \in I} \pi(i) \left[\int_S \left(\max_{\alpha \in g(i)} U_s(\alpha) \right) d\mu(s \mid i) \right].$$

To address the question, whether the assumption that information only leads to updated beliefs constrains period 1 choice behavior, note that V is a special case of a representation,

[7] Dekel, Lipman and Rustichini (1998-a) note that *"... there are very significant problems to be solved before we can generate interesting conclusions for contracting [...] while the Kreps model (and its modifications) seems appropriate for unforeseen contingencies, [...] there are no meaningful subjective probabilities. A refinement of the model that pins down probabilities would be useful."*

[8] In the finite case, the richness assumption limits the degree to which DM can exhibit preference for flexibility. This is not the main thrust of the axiom, as can be seen in the case of a metric information space.

where information may also impact the scaling of utilities,

$$\widetilde{V}(g) = \sum_{i \in I} \pi(i) \left[\int_S \left(\max_{\alpha \in g(i)} U_{s,i}(\alpha) \right) d\mu(s|i) \right].$$

If $\succ$ is a binary relation on acts that can be represented by $\widetilde{V}$, it has to satisfy the following standard axioms: it is a *Preference* ranking. It is *Continuous* in an appropriate topology. *Independence* holds for an appropriately defined convex combination of acts and, because we constrain S to exclude the trivial taste, $\succ$ must also satisfy *Nontriviality.* Comparing acts which differ only in the menu they assign under one particular information, i, but agree on some default menu otherwise, induces a ranking of menus, $\succ_i$. It must satisfy *Monotonicity.* Because V is a special case of $\widetilde{V}$, those five axioms are necessary for the existence of the representation V, as well.

Theorem 1 states that the axioms are not only necessary, but also sufficient for the existence of the representation V. Thus, whenever $\widetilde{V}$ represents $\succ$, so does V. Hence, the assumption that information only leads to updated beliefs does not constrain period 1 choice.

To address the question, whether the parameters of the representation V can be related to choice behavior, suppose a representation V exists, where subjective probabilities π coincide with objective probabilities ϕ. Theorem 2 states that in this representation, subjective probabilities of tastes $\mu(s|i)$ and utilities U_s representing those tastes are generically identified by $\succ$. Thus, beliefs and utilities are conceptually distinct.

The genericity assumption in Theorem 2 requires the impact of changing information on the ranking of menus to be *rich* enough: whenever there is preference for flexibility with respect to two menus under information i, $A \cup B \succ_i A$, then those menus are not indifferent under every information; there is $j \in I$, such that $A \nsim_j B$. This assumption is trivially satisfied, if $A \nsim_i B$. Otherwise, $A \cup B \succ_i A \sim_i B$ implies that the best lottery is sometimes in A and sometimes in B. *Richness* requires that changing information can make one or the other case more relevant.

Both V and $\widetilde{V}$ evaluate uncertainty according to the expected utility criterion, involving subjective beliefs. However, just as in Kreps' representation, utilities and beliefs are not two distinct concepts in the representation $\widetilde{V}$. This indeterminacy is the central drawback of existing representations of preference for flexibility.[9] Theorem 2 is the main contribution of this work.

If Richness fails only partially, then parameters of the representation V are not identified uniquely. However, Proposition 2 gives bounds on those parameters, which depend on the

[9] In addition to Kreps (1979) and DLR, see also Nehring (1999) and Epstein and Seo (2007).

extent to which Richness is violated.

While Theorem 1 establishes existence of a representation based on some subjective probability distribution π over information, Theorem 2 makes a uniqueness statement for representations based on the objective probability distribution ϕ over information. A combined result that gives conditions under which the representation V exists, the subjective probability distribution over information π is unique and $l(s)$ and $\mu(s|i)$ are unique in the appropriate sense, is not implied by the axioms imposed so far. Theorem 3 fills this gap. It requires *Partial Information Independence*, which is a weakened version of the state independence assumption suggested by Anscombe and Aumann (1963, henceforth AA.)

If π is not unique, a representation based on ϕ may still not exist. Proposition 3 gives a robustness result: if there is a representation based on a distribution π, which disagrees only "slightly" with ϕ, then a representation based on ϕ exists under the technical condition that I is not too large. In this case the probabilities $\mu(s|i)$ of tastes, elicited using ϕ, are at least good estimates of those based on π.

1.2. Illustration of Identification of Beliefs

First Continuation of Example: Recall the retailer's objective function:

$$V(g) = \sum_{i\in\{H,L\}} \phi(i) \left[\sum_{s\in\{h,l\}} \mu(s|i) \max_{q\in g(i)} (v_s(q)) \right].$$

• Specification 1: The retailer believes that he faces high demand with higher probability, if consumer confidence is high, $\mu(h|H) = \frac{2}{3}$ and $\mu(h|L) = \frac{1}{3}$.

The ranking of acts induced by V satisfies Richness: $\{q_h\} \cup \{q_l\} \succ_H \{q_l\}$, because if demand is high, q_h is preferred over q_l. If also $\{q_h\} \sim_H \{q_l\}$, then $\mu(h|H) > \mu(h|L)$ implies $\{q_l\} \succ_L \{q_h\}$.[10]

Hence, beliefs should be identified. Suppose, to the contrary, that there was another representation of the same ranking of contracts with beliefs $\widehat{\mu}(s|i)$ and tastes $\widehat{v}_l(q_l) - \widehat{v}_l(q_h)$ and $\widehat{v}_h(q_h) - \widehat{v}_h(q_l)$:

$$\widehat{V}(g) = \sum_{i\in\{H,L\}} \phi(i) \left[\sum_{s\in\{h,l\}} \widehat{\mu}(s|i) \max_{\alpha\in g(i)} (\widehat{v}_s(q)) \right].$$

[10] In the example we *start* from a representation, which is based on only two subjective states. Therefore, we do not need to consider lotteries over prizes, in order to identify beliefs. Richness is clearly satisfied, even if we do consider lotteries.

$V(g)$ and $\widehat{V}(g)$ have to generate the same ranking of contracts, in particular of those, which disagree only under information i. It may be intuitive that the two representations must agree on the relative "overall weight" given to a particular taste, where this weight is comprised of the probability and the scaling of the utility:[11]

$$\frac{\widehat{\mu}(h|i)(\widehat{v}_h(q_h)-\widehat{v}_h(q_l))}{\widehat{\mu}(l|i)(\widehat{v}_l(q_h)-\widehat{v}_l(q_l))}=\frac{\mu(h|i)(v_h(q_h)-v_h(q_l))}{\mu(l|i)(v_l(q_h)-v_l(q_l))}. \qquad (*)$$

$\widehat{\mu}(s|i)$ must be a probability measure. Therefore, for some $\eta>0$,

$$r_h:=\eta\frac{v_h(q_h)-v_h(q_l)}{\widehat{v}_h(q_h)-\widehat{v}_h(q_l)} \text{ and } r_l:=\eta\frac{v_l(q_l)-v_l(q_h)}{\widehat{v}_l(q_l)-\widehat{v}_l(q_h)}$$

must satisfy

$$\begin{aligned}1 &= \mu(h|H)r_h+\mu(l|H)r_l=\frac{2}{3}r_h+\frac{1}{3}r_l\\ 1 &= \mu(h|L)r_h+\mu(l|L)r_l=\frac{1}{3}r_h+\frac{2}{3}r_l.\end{aligned}$$

The two equations are linearly independent, and therefore $r_h=r_l=1$ is the unique solution. Then, $\widehat{\mu}(s|i)=\mu(s|i)$ for $i\in\{H,L\}$ and $s\in\{h,l\}$ follows immediately from $(*)$. Hence V is unique up to a common linear transformation of $v_h(q_h)-v_h(q_l)$ and $v_l(q_l)-v_l(q_h)$.

• Specification 2: Now suppose that, instead, $\mu(h|H)=\mu(h|L)=\frac{2}{3}$, so the retailer's beliefs are independent of information about consumer confidence. Clearly the ranking of contracts induced by V under this specification does not satisfy Richness. To see that beliefs are not identified, note that now r_h and r_l have to satisfy

$$\begin{aligned}1 &= \mu(h|H)r_h+\mu(l|H)r_l=\frac{2}{3}r_h+\frac{1}{3}r_l\\ 1 &= \mu(h|L)r_h+\mu(l|L)r_l=\frac{2}{3}r_h+\frac{1}{3}r_l.\end{aligned}$$

These equations are identical, and therefore r_h and r_l are not determined uniquely, and neither is $\mu(s|i)$ for $i\in\{H,L\}$ and $s\in\{h,l\}$.

This reasoning can be generalized: whenever a representation generates at least as many independent conditions, indexed by $i\in I$, on the scaling of vNM utilities as there are relevant tastes, then the scaling of utilities is uniquely identified up to a common linear transformation, and consequently beliefs are uniquely identified. For the proof of Theorem

[11] The generalization of this condition is shown in DLR and also follows from the proof of Theorem 2.

2, however, no particular representation is given. The crucial step is to show that every ranking satisfying Richness must be represented this way.

2. The Model

We consider a two-stage choice problem, where public information becomes available between the two stages. In period 2, DM chooses a lottery over prizes. This choice is not modelled explicitly. The lotteries available to her in period 2 may depend on choice in period 1 and on the information. Period 1 choice is described as choice of an act, which specifies a set of lotteries (a menu) that is contingent on information and contains the feasible choices for period 2.

Let Z be a finite prize space with cardinality k and typical elements x, y, z. $\Delta(Z)$ is the space of all lotteries over Z with typical elements α, β, γ. We write explicitly $\alpha = \langle \alpha(x), x; \alpha(y), y; ...\rangle$, where $\alpha(x)$ is the probability α assigns to $x \in Z$ etc. Let $\mathcal{A}$ be the collection of all compact subsets of $\Delta(Z)$ with menus A, B, C as elements.[12]

Endow $\mathcal{A}$ with the topology generated by the Hausdorff metric

$$d_h(A, B) = \max\left\{\max_A \min_B d_p(\alpha, \beta), \max_B \min_A d_p(\alpha, \beta)\right\}$$

where d_p is the Prohov metric, which generates the weak topology, when restricted to lotteries.

Further let I be an exogenous state space with elements i, j. Call elements of I "information". Information is observable upon realization. Let $\mathcal{F}$ be a σ-algebra on I. Two cases have to be distinguished. If I is finite, $\mathcal{F}$ is assumed to be the σ-algebra generated by the power set of I. If I is a generic metric space, then $\mathcal{F}$ is the Borel σ-algebra with respect to the induced topology.

Let G be the set of all acts. An act is a measurable function $g : I \to \mathcal{A}$. After information i realizes, DM chooses an object from the menu $g(i) \in \mathcal{A}$. This choice is not explicitly modeled. $\succ$ is a binary relation on $G \times G$. $\succcurlyeq$ and $\sim$ are defined the usual way.

G can be viewed as a product space generated by the index set I, $G = \prod_{i \in I} \mathcal{A}$. Thus, it can be endowed with the product topology, based on the topology defined on $\mathcal{A}$.

[12] Compactness is not essential. However, items on a menu are alternatives for future choice and choice from a non-compact set is difficult to interpret. If menus were not compact, maximum and minimum would have to be replaced by supremum and infimum, respectively, in all that follows.

Definition 1: *The convex combination of menus is defined as*

$$pA + (1-p)\,B := \{\gamma = p\alpha + (1-p)\,\beta\,|\alpha \in A, \beta \in B\}\,.$$

The convex combination of acts is defined, such that

$$(pg + (1-p)\,g')\,(i) := pg\,(i) + (1-p)\,g'\,(i)\,.$$

The first four axioms are standard:

Axiom 1 *(Preference): $\succ$ is asymmetric and negatively transitive.*

Axiom 2 *(Continuity): The sets $\{g\,|g \succ h\}$ and $\{g\,|g \prec h\}$ are open in the topology defined on G for all $h \in G$.*

Axiom 3 *(Independence): If for $g, g' \in G$, $g \succ g'$ and if $p \in (0,1)$, then*

$$pg + (1-p)\,h \succ pg' + (1-p)\,h$$

for all $h \in G$.

Axiom 4 *(Nontriviality): There are g,$h \in G$, such that $g \succ h$.*

If a convex combination of menus were defined as a lottery over menus, then the motivation of Independence in our setup would be the same as in more familiar contexts. Uncertainty would be resolved before DM consumes an item from one of the menus. However, following DLR and Gul and Pesendorfer (2001), we define the convex combinations of menus as the menu containing all the convex combinations of their elements. The uncertainty generated by the convex combination is only resolved after DM chooses an item from this new menu. Gul and Pesendorfer term the additional assumption needed to motivate Independence in our setup *"indifference as to when uncertainty is resolved."*[13]

The next definition considers acts that give a menu A in event D and some default menu in the event not D. The default menu is the whole prize space, Z, which takes the form $\{\langle 1, z\rangle\,|z \in Z\}$, when written in terms of lotteries. Comparing those acts induces a ranking $\succ_D$ over menus:

[13] Both DLR and Gul and Pesendorfer elaborate this argument.

Definition 2: *For $D \in \mathcal{F}$ and $A \in \mathcal{A}$, define g_D^A by*

$$g_D^A(i) := \begin{cases} A \text{ for } i \in D \\ \{\langle 1, z\rangle \,|z \in Z\} \text{ otherwise} \end{cases}.$$

Let $\succ_D$ be the induced binary relation on $\mathcal{A} \times \mathcal{A}$, $A \succ_D B$, if and only if $g_D^A \succ g_D^B$. $\succsim_D$ and $\sim_D$ are defined the usual way. An event $D \in \mathcal{F}$ is nontrivial, if there are $A, B \in \mathcal{A}$ with $A \succ_D B$.

Ranking menus according to $\succ_D$, DM may have preference for flexibility. This is captured by the central axiom in Kreps, which states that larger menus are weakly better than smaller menus:

Axiom 5 *(Monotonicity): $A \cup B \succsim_D A$ for all $A, B \in \mathcal{A}$ and $D \in \mathcal{F}$.*

Corollary 1: *If $\succ$ satisfies Axioms 1-5, then $\succ_D$ is a preference relation and satisfies the appropriate variants of Continuity, Independence and Monotonicity for all $D \in \mathcal{F}$. Further, there is a nontrivial event $D \in \mathcal{F}$.*

The proof is immediate.

Theorem DLRS *(Theorem 2 in DLRS): For $D \in \mathcal{F}$ nontrivial, $\succ_D$ is a preference that satisfies Continuity, Independence and Monotonicity, if and only if there is a subjective state space S_D, a positive countably*[14] *additive measure $\mu_D(s)$ on S_D and a set of non-constant, continuous expected utility functions $U_{s,D} : \Delta(Z) \to \mathbb{R}$, such that*

$$V_D(A) = \int_{S_D} \max_{\alpha \in A} U_{s,D}(\alpha)\, d\mu_D(s)$$

represents $\succ_D$ and every cardinal ranking of prizes $x \in Z$ corresponds to at most one state in S_D.

Because $U_{s,D}(\alpha)$ are expected utility functions,

$$V_D(pA + (1-p)B) = pV_D(A) + (1-p)V_D(B).$$

This linearity allows the following convention for expressing $V_{D'}$ for all $D' \in \mathcal{F}$. First, fix

[14] See footnote 3 in DLRS.

$U_s(\alpha) := U_{s,D}(\alpha)$ for some $D \in \mathcal{F}$. Then, for all $D' \in \mathcal{F}$, $\succ_{D'}$ can be represented with $U_{s,D'}(\alpha) = U_s(\alpha)$ by simply rescaling $\mu_{D'}(s)$ to compensate. Second, normalize $\mu_{D'}(s)$ to be a probability measure for all $D' \in \mathcal{F}$.

From now on, consider the "taste space", given the prize space Z with cardinality k:

Definition 3:

$$S := \left\{ s \in \mathbb{R}^k_+ \middle| |s| = 1, \min_{z \in Z} s_z = 0 \right\}$$

is the taste space.

Up to normalization this is the state space DLRS use in their proof of Theorem DLRS. Every taste in S is a vector with k components. Every entry can be thought of as specifying the (relative) utility associated with the corresponding prize, given the taste. Thus, every ranking of lotteries over prizes corresponds to a unique future taste.[15] The cardinal ranking (or utility), given taste s, is then fully determined by a bounded intensity $l(s)$ associated with that taste, and the value of the worst prize, $\underline{u}_s$. Prize $x \in Z$ is evaluated by $\underline{u}_s + l(s) s_x$, where s_x is the entry of s, which corresponds to prize $x \in Z$. Then, for a lottery $\langle \alpha(x), x; \alpha(y), y; ... \rangle$,

$$U_s(\alpha) = \underline{u}_s + \sum_{x \in Z} l(s) \alpha(x) s_x.$$

Normalize $\underline{u}_s = 0$ for all $s \in S$ and write

$$U_s(\alpha) = l(s)(\alpha \cdot s).$$

We do not consider the taste where DM is indifferent over all prizes, implicitly assuming nontriviality of the ex-post preference over prizes.[16] Hence $l(s) > 0$ for all $s \in S$.[17]

[15] In Theorem DLRS, as in the theorems that follow, there is clearly always a larger state space, also allowing a representation of $\succ_D$, in which multiple states represent the same ranking of lotteries.

[16] For a nontrivial event $D \in \mathcal{F}$, the trivial taste is not required to obtain the representation in Theorem DLRS. As seen below, a representation based on S does not require the assumption of nontriviality for each $D \in \mathcal{F}$.

[17] DLR further establish that, for the smallest taste space S_D, which allows a representation of $\succ_D$, $closure(S_D)$ is unique. If we wanted this closure to agree across information, we would have to require that information is not exhaustive: If $D, D' \in \mathcal{F}$ and $A, B \in \mathcal{A}$, then $A \cup B \not\sim_D A$ implies $A \cup B \not\sim_{D'} A$.

2.1. A finite information space

Let T be the cardinality of I. Let $\mathcal{F}$ be be the σ-algebra generated by the power set of I. Without risk of confusion we write $i \in I$ to denote elementary events $\{i\} \in \mathcal{F}$.

Definition 4: *Let $\overline{\mathcal{A}}$ be the collection of all convex subsets of $\Delta(Z)$. Let $\overline{G}$ be the collection of all acts: $g : I \to \overline{\mathcal{A}}$. Call $g \in \overline{G}$ a convex act.*

Proposition 1: *$\succ$ constrained to $\overline{G}$ satisfies Axioms 1-3, if and only if there are continuous linear functions $v_i : \overline{\mathcal{A}} \to \mathbb{R}$, such that $v : \overline{G} \to \mathbb{R}$ with*

$$v(g) = \sum_{i \in I} v_i(g(i))$$

represents $\succ$ on $\overline{G}$.

Moreover, if there is another collection of continuous linear functions, $v_i' : \overline{\mathcal{A}} \to \mathbb{R}$, such that

$$v'(g) = \sum_{i \in I} v_i'(g(i))$$

represents $\succ$ on $\overline{G}$, then there are constants $a > 0$ and $\{b_i \, | i \in I\}$, such that $v_i' = b_i + a v_i$ for each $i \in I$.

Proof: The collection of convex acts $\overline{G}$ together with the convex combination of acts as a mixture operation is a mixture space. Proposition 1 is an application of the Mixture Space Theorem (Theorem 5.11 in Kreps (1988)),[18] where additive separability across I follows from the usual induction argument. ■

Corollary 2: If $i \in I$ is nontrivial, then $V_i(A)$ and $v_i(A)$ agree on $\overline{\mathcal{A}}$ up to positive affine transformations.

Proof: Evaluating $v\left(g_i^A\right)$ implies that v_i represents $\succ_i$ on $\overline{\mathcal{A}}$. v_i is linear. The Mixture Space Theorem states that any other linear representation of $\succ_i$ agrees with v_i, up to a positive affine transformation. According to Theorem DLRS, $V_i(A)$ is linear and represents $\succ_i$ on $\mathcal{A}$. □

[18] Axiom 2 (Continuity) is stronger than von Neumann-Morgenstern Continuity on $\overline{G}$, which requires that for all $g \succ g' \succ g''$ there are $p, q \in (0,1)$, such that $pg + (1-p)g'' \succ g' \succ qg + (1-q)g''$.

Theorem 1: $\succ$ *satisfies Axioms 1-5, if and only if there are a bounded set of positive numbers* $\{l(s) | s \in S\}$, *a conditional probability measure* $\mu(s|i)$ *and a probability measure* π *on* I, *such that the function* $V : G \to \mathbb{R}$,

$$V(g) = \sum_{i \in I} \pi(i) \left[\int_S l(s) \max_{\alpha \in g(i)} (\alpha \cdot s) \, d\mu(s|i) \right]$$

represents $\succ$.

Proof: See Appendix.

From Corollary 2, the proof establishes that setting $V \equiv v$ on $\overline{G}$ determines V on all of G, because $\overline{G}$ is order-dense in G with respect to $\succ$. On $\overline{G}$, V_i (as determined by Theorem DLRS) agrees up to scaling with v_i. The scaling is absorbed by $\pi(i)$, which is then normalized to be a probability distribution. Thus, an act is evaluated by

$$V(g) = \sum_{i \in I} \pi(i) V_i(g(i)).$$

To illustrate the theorem, note that this is AA's representation, where our acts have menus as outcomes, while Anscombe-Aumann acts have lotteries as outcomes.[19] Indeed, Axioms 1-3 imply AA's axioms.

According to Theorem DLRS, $V_i(A)$ has the form

$$V_i(A) = \int_S \max_{\alpha \in A} (U_{s,i}(\alpha)) \, d\mu_i(s),$$

where $U_{s,i}$ are vNM utility functions.[20] And, indeed, Axioms 1-5 imply DLR's axioms, according to Corollary 1. Combining the two:

$$\widetilde{V}(g) = \sum_{i \in I} \pi(i) \left[\int_S \max_{\alpha \in g(i)} (U_{s,i}(\alpha)) \, d\mu(s|i) \right]$$

The representation V in Theorem 1 is a special case of $\widetilde{V}$, where exogenous states are

[19] In terms of the contracting example, Anscombe-Aumann acts correspond to complete contracts.

[20] In terms of the contracting example, menus correspond to contracts that do not condition on information.

information, which impacts only probabilities $\mu(s|i)$, but not the intensity $l(s)$ of taste s:

$$V(g) = \sum_{i \in I} \pi(i) \left[\int_S l(s) \max_{\alpha \in g(i)} (\alpha \cdot s) \, d\mu(s|i) \right]$$

The intensity of each taste is endogenous, but it is fixed across information. Theorem 1 states that this particular combination of AA and DLRS allows representation of $\succ$ under their combined axioms. The assumption that information impacts only beliefs does, therefore, not constrain period 1 choice.

So far the axioms have not capitalized on the richness of the domain provided by the information space, I. In fact, it might still be that information is irrelevant for the decision maker. The next axiom specifies the way it matters.

Axiom 6 *(Richness of Information)*: *If $A \cup B \nsim_i B$ for some $i \in I$, then there is $j \in I$ with $A \nsim_j B$.*

To roughly paraphrase Axiom 6: whenever there is preference for flexibility with respect to two menus, then there is also some information dependence in their ranking.

Throughout, the interpretation is that, ultimately, only the chosen item matters for the value of a menu. If $A \nsim_i B$, then Axiom 6 is empty. If $A \sim_i B$, then $A \cup B \nsim_i B$ implies that under i, the chosen item must sometimes be in A and sometimes in B. Axiom 6 requires that changing information can make either one or the other case more relevant, namely that there is $j \in I$ with $A \nsim_j B$.

In terms of the representation established in Theorem 1, there must be a change of information, that makes the tastes where an item in A is preferred over all items in B more likely.[21] Axiom 6 is not a strong assumption in the sense that it is local: it only requires breaking indifference.

As before, let ϕ denote objective probabilities of information.

Theorem 2: *If, given $\phi : I \to \mathbb{R}_+$,*

$$V(g) = \sum_{i \in I} \phi(i) \left[\int_S l(s) \max_{\alpha \in g(i)} (\alpha \cdot s) \, d\mu(s|i) \right]$$

[21] If there are multiple such tastes, only the utility-weighted aggregate probability has to increase.

represents $\succ$,[22] *then statements i and ii below are equivalent and imply statement iii:*
i) $\succ$ *satisfies Axiom 6,*
ii) the measure $\mu(s\,|i)$ *is unique for all* $i \in I$ *and up to* ϕ*-measure zero changes,* $l(s)$ *is unique for all* $s \in S$ *up to common linear transformations and changes on a set* $S' \subset S$ *with* $\sum_{i \in I} \phi(i) \left[\int_{S'} d\mu(s\,|i) \right] = 0$,[23]
iii) the cardinality of $S^* := \bigcup_{i \in I} support(\mu(s\,|i))$ *is bounded above by* T*.*

Proof: See Appendix.

If a decision maker acts, *as if* she had preference for flexibility, updated her beliefs when learning information and otherwise maximized expected utility according to objective probabilities, then her preferences satisfy Axiom 6, if and only if the subjective probabilities that a Kreps-style representation assigns to future tastes are determined uniquely.

The unique identification of probabilities and utilities gives meaning to the description of beliefs and tastes as two distinct concepts. The lack of this distinction is the central drawback of previous work on preference for flexibility, starting with Kreps' (1979) seminal paper. Theorem 2 is the main contribution of our work.

Further, Axiom 6 generates finiteness of S (statement iii in Theorem 2). Hence, Axiom 6 is not innocuous: $\mathcal{F}$ must be rich enough to distinguish between any two menus, for which DM might have preference for flexibility. This implies that only finitely many lotteries can be appreciated in any menu.[24] In this sense a limit on the cardinality of S limits the degree to which DM can have preference for flexibility. Section 2.2 considers an amendment to Axiom 6, which guarantees sufficient tightness of information to allow I to be a metric space, thereby lifting the constraint on the cardinality of S.

If our theory is viewed as purely descriptive of period 1 choice, then considering only the class of representations V is not an assumption, in the sense that it does not constrain choice behavior. It merely singles out an intuitive representation, where parameters can be related to period 1 choice. For illustration, consider the following simple comparative statics: Information, $i \in I$, may have the form of a market outlook. There are three prizes, $Z = \{m, x, 0\}$, interpreted as money, a risky asset and nothing. Money generates the same value (compared to nothing) under every taste. Uncertainty about the value of the risky asset is subjective. Thus, every taste identifies the value of the risky asset in terms of

[22] In particular every event with $\phi(i) = 0$ must be trivial, if this representation exists.

[23] The expected utilities, $U_s(\alpha) = \underline{u}_s + \sum_{x \in Z} l(s)\, \alpha_x s_x$, are then only unique up to the common linear transformation of $l(s)$ and the addition of a constant.

[24] Kopylov (2007) turns this around and generates finiteness of S in the absence of an exogenous state space, by basically assuming that the number of lotteries DM can appreciate in any given menu is limited.

money. Renormalizing the taste space, such that $s = (1, s_x, 0)$ with $s_x \in \mathbb{R}$ and $l(s) = 1$ for all $s \in S$, shows that S is isomorphic to $\mathbb{R}$. $\succ_i$ and $\succ_j$ can then be compared in terms of the distributions $\mu(s_x|i)$ and $\mu(s_x|j)$ on $\mathbb{R}$. The following statement is straight forward: $\mu(s_x|i)$ has a higher mean than $\mu(s_x|j)$, if and only if a chance to get the risky asset is compared more favorably to a chance to get money under information i than under information j:

$$\{\langle q, x; 1-q, 0\rangle\} \succsim_j \{\langle p, m; 1-p, 0\rangle\} \implies \{\langle q, x; 1-q, 0\rangle\} \succ_i \{\langle p, m; 1-p, 0\rangle\}.$$

Remark: Let $\widetilde{S} = S \times \mathbb{R}_+$ be the subjective state space that collects all pairs of vNM utilities and intensities. Then

$$\sum_{i \in I} \phi(i) \left[\int_{\widetilde{S}} \max_{\alpha \in g(i)} \left(U_{\widetilde{s}}(\alpha)\right) d\widetilde{\mu}(\widetilde{s}|i) \right]$$

includes as a special case

$$\sum_{i \in I} \phi(i) \left[\int_{S} \max_{\alpha \in g(i)} \left(U_{s,i}(\alpha)\right) d\mu(s|i) \right].$$

Theorem 2 implies that there is a unique probability measure, $\widetilde{\mu}(\widetilde{s}|i)$, on this larger subjective state space, $\widetilde{S}$, that allows representation of $\succ$ and has the smallest support, $\bigcup_{i \in I} \text{support}(\widetilde{\mu}(\widetilde{s}|i)) \subset \widetilde{S}$. It is the only measure that allows representation and for which every taste, $s \in S$, corresponds to at most one state, $\widetilde{s} \in \widetilde{S}$, in its support.

Thus, the restriction to representations based on the taste space S, where information impacts only beliefs, is equivalent to considering those representations based on the subjective state space $\widetilde{S}$, which require only a minimal amount of subjective states in the sense of DLR.

What can be learned about period 2 choice? If DM has information about the formation of her future taste, then the natural inductive step is to employ the beliefs about future tastes elicited in Theorem 2 to forecast period 2 choice. Doing so implies that the representation V is interpreted as a map of the decision making process and that period 1 choice, which satisfies Axiom 6, constrains period 2 choice frequencies.

Thus, confining attention to the class of representations V becomes an assumption with behavioral implications. On the one hand this requires justification of the assumption for a particular application, on the other hand it allows the assumption to be refuted, if observed behavior is not in line with the model's predictions. The particular restriction to representations where information leads only to updated beliefs has a similar character to the restriciton underlying the uniqueness results in AA and Savage. There, beliefs over exogenous states are identified based on the restriction that a state-independent ordinal ranking implies a state-independent cardinal ranking of prizes.[25][26]

The ability to forecast period 2 choice frequencies is relevant in strategic situations, for example in the context of incomplete contracts. Section 4 elaborates.

Both types of exogenous uncertainty in our domain are essential for the uniqueness result: on the one hand, DLR find that menus over lotteries alone do not allow to distinguish utility levels $l(s)$ and probabilities $\mu(s)$. There has to be some channel through which to vary one, but not the other. In the representation V, information impacts only probabilities, $\mu(s|i)$. On the other hand, Nehring (1999) finds that acts with menus of prizes as outcomes do not allow to distinguish utility levels $l(s)$ and probabilities $\mu(s)$ in the axiomatic setup developed by Savage (1954).[27] To establish the uniqueness result, the payoff a menu gives under an individual taste has to be varied. This is possible, only because DM can be offered lotteries over prizes.

At the outset, the proof of the sufficiency statement in Theorem 2 shows that in case of a finite information space I, Axiom 6 implies that any representation is based on a finite taste space:

$$V(g) = \sum_{i \in I} \phi(i) \left[\sum_{S^*} l(s) \left(\max_{\alpha \in g(i)} (\alpha \cdot s) \right) \mu(s|i) \right].$$

with $S^* \subset S$ finite. To make the next argument easier, suppose $z \in Z$ is the worst prize under every taste $s \in S^*$. Let $K_{\{z\}}$ be an appropriately constructed reference menu. For each $s \in S^*$ and ε small enough, it is possible to construct a menu that is worse than $K_{\{z\}}$ by $c(\varepsilon) l(s)$ under taste s but equally good under all other tastes, where $c(\varepsilon)$ is a strictly increasing continuous function with $c(0) = 0$. This construction is a central step in the

[25] AA's representation can be viewed as a special case of ours, where there is only one taste. The restricition, as in our case, is that the scaling of the vNM utility indexed by this taste is independent of information.

[26] Karni and co-authors (Karni and Mongin [2000], Grant and Karni [2005], Karni [2006]) point out that, if one were to take AA's identification of unique subjective probabilities over objective states as a foundation for Bayesian decision making in the sense that subjective probabilities are interpreted as DM's true beliefs, then the assumption that exogenous states do not impact the scaling of the vNM utility would be problematic. The problem would be that beliefs in AA's model have no implications for observable behavior. Hence, the model could not be measured against the quality of its predictions.

[27] Following Nehring (1996), a companion paper to the one cited above, Epstein and Seo (2007) consider a domain of random menus, which are lotteries with menus as outcomes. On this domain they tease out unique induced probability distributions over ex post upper contour sets as the strongest possible uniqueness statement.

proof. It is illustrated in the Appendix. Let K_ρ^ε be the convex combination of those menus, where $\rho(s)$ is the weight given to the menu, which is worse than $K_{\{z\}}$ under taste s. Then

$$V\left(g_i^{K_{\{z\}}}\right) - V\left(g_i^{K_\rho^\varepsilon}\right) = \phi(i)\, c(\varepsilon) \sum_{S^*} \rho(s)\, l(s)\, \mu(s\,|i).$$

Now suppose there is another representation

$$\widehat{V}(g) = \sum_{i \in I} \phi(i) \left[\sum_{\widehat{S^*}} \widehat{l}(s) \left(\max_{\alpha \in g(i)} (\alpha \cdot s) \right) \widehat{\mu}(s\,|i) \right].$$

According to this representation the same type of construction yields a set $K_{\widehat{\rho}}^\xi$, such that

$$\widehat{V}\left(g_i^{K_{\{z\}}}\right) - \widehat{V}\left(g_i^{K_{\widehat{\rho}}^\xi}\right) = \phi(i)\, \widehat{c}(\xi) \sum_{\widehat{S^*}} \widehat{\rho}(s)\, \widehat{l}(s)\, \widehat{\mu}(s\,|i).$$

Choose $\rho(s) \propto \frac{1}{l(s)}$. Then

$$\frac{V\left(g_i^{K_{\{z\}}}\right) - V\left(g_i^{K_\rho^\varepsilon}\right)}{V\left(g_j^{K_{\{z\}}}\right) - V\left(g_j^{K_\rho^\varepsilon}\right)} = \frac{\phi(i)}{\phi(j)}$$

for all $i, j \in I$. For $\widehat{\rho}(s) \propto \frac{1}{\widehat{l}(s)}$, the same holds for $K_{\widehat{\rho}}^\xi$ in terms of $\widehat{V}$ and, as V and $\widehat{V}$ represent the same preference, also in terms of V. This implies that, for ε and ξ such that $K_\rho^\varepsilon \sim_j K_{\widehat{\rho}}^\xi$ for some $j \in I$, it must be true that $K_\rho^\varepsilon \sim_i K_{\widehat{\rho}}^\xi$ for all $i \in I$. Suppose, contrary to the Theorem, that Axiom 6 holds and $S^* \neq \widehat{S^*}$ or $l(s) \neq \widehat{l}(s)$ for $s \in S^*$. Then there is $s \in S^*$ with $\rho(s) \neq \widehat{\rho}(s)$. But then there must be $j \in I$, such that, according to V, $K_\rho^\varepsilon \cup K_{\widehat{\rho}}^\xi \succ_j K_\rho^\varepsilon$. Axiom 6 then implies that there is $k \in I$, such that $K_\rho^\varepsilon \nsim_k K_{\widehat{\rho}}^\xi$, a contradiction. Thus $S^* = \widehat{S^*}$ and $l(s) = \widehat{l}(s)$ for all $s \in S^*$. It is then a variation of standard results, that $\mu(s\,|i) = \widehat{\mu}(s\,|i)$ for all $s \in S^*$ and $i \in I$.[28]

If Axiom 6 fails completely, in the sense that information is irrelevant to the decision maker, clearly there are no bounds on the range of probability measures $\mu(s\,|i)$, which allow a representation. This is the same indeterminacy first encountered by Kreps.

But how much indeterminacy is implied by a partial failure of Axiom 6? Suppose there is a representation of $\succ$ as in Theorem 2. Further suppose there is a pair of menus, $A, B \in \mathcal{A}$, such that $A \cup B \nsim_i B$ for some $i \in I$, but $A \sim_j B$ for all $j \in I$. This means there is some preference for flexibility in having both A and B available, but their comparison is independent of information. To say this more precisely:

[28] Up to $\phi(i)$-measure zero changes.

Definition 5:

$$c_{A,B}(s) := l(s)\left(\max_{\alpha\in A}(\alpha\cdot s) - \max_{\alpha\in B}(\alpha\cdot s)\right)$$

is the cost of having to choose from $B \in \mathcal{A}$ instead of $A \in \mathcal{A}$ under taste $s \in S$.

$A \cup B \nsim_i B$ implies that $c_{A,B}(s)$ cannot be zero for all s and $A \sim_i B$ implies that it cannot be any other constant. Still, $A \sim_j B$ for all $j \in I$ means

$$\int_S c_{A,B}(s)\, d\mu(s\,|j) = 0$$

for all $j \in I$. This suggests that with $\widehat{\mu}(s\,|i) = \frac{(1+\eta c_{A,B}(s))\mu(s|i)}{\sum_{S^*}(1+\eta c_{A,B}(s))\mu(s|i)}$ and $\widehat{l}(s) = \frac{l(s)}{1+\eta c_{A,B}(s)}$ we can construct another representation, $\widehat{V}(g)$, if η is small enough, such that $(1+\eta c_{A,B}(s)) > 0$ for all $s \in S$. The proof of the necessity of Axiom 6 in Theorem 2 shows how this is done in detail. The following proposition is based on the same idea.

Proposition 2: *Suppose*

$$V(g) = \sum_{i\in I}\phi(i)\left[\int_S l(s)\max_{\alpha\in g(i)}(\alpha\cdot s)\, d\mu(s\,|i)\right]$$

represents $\succ$. Then the following two conditions are equivalent:
i) *there is a pair of menus $A, B \in \mathcal{A}$, such that $A\cup B \nsim_i B$ for some $i \in I$, but $A \sim_j B$ for all $j \in I$,*
ii) *there is a continuum of representations based on $\widehat{\mu}(s\,|i) = \frac{(1+\eta c_{A,B}(s))\mu(s|i)}{\sum_{S^*}(1+\eta c_{A,B}(s))\mu(s|i)}$ and $\widehat{l}(s) = \frac{l(s)}{1+\eta c_{A,B}(s)}$, parametrized by $\eta > -\frac{1}{c_{A,B}(s)}$.*

If there is another pair of menus $A', B' \in A$ satisfying i), then they add another set of possible representations, if and only if

$$\frac{c_{A',B'}(s)}{c_{A',B'}(s')} \neq \frac{c_{A,B}(s)}{c_{A,B}(s')}$$

for some $s, s' \in S$.

Proof: See Appendix.

2.2. A metric information space

Let I be a generic metric space and $\mathcal{F}$ the Borel σ-algebra with respect to the induced topology. The goal is to establish appropriate versions of Theorems 1 and 2 for this case.

In addition to Axioms 1-5, the result corresponding to Theorem 1 requires that $\succ_D$ does not change "too much" for "small" changes in D. As pointed out before, Axiom 6 has to be strengthened to establish a result as in Theorem 2.

Definitions and results that correspond to ones in the previous section, but only after appropriate adjustments, are distinguished by a prime on their label.

Definition 6: *Let $\{D_t \,|t \in \{1, .., T\}\}$ be a finite partition of I with $D_t \in \mathcal{F}$. $\{D_t\}$ denotes a generic partition of this type. Further let $G_{\{D_t\}}$ be the collection of acts where the outcome depends only on the event $D \in \{D_t\}$. Let*

$$G^* := \bigcup_{\{D_t\}} G_{\{D_t\}}$$

be the set of simple acts. $\overline{G} \cap G^*$ is the collection of all simple convex acts.

The support of $g \in G_{\{D_t\}}$ is a finite subset of $\mathcal{A}$.

Proposition 1': *$\succ$ constrained to $\overline{G} \cap G^*$ satisfies Axioms 1-3, if and only if there are continuous linear functions $v_D : \overline{\mathcal{A}} \to \mathbb{R}$, such that $v : \overline{G} \cap G^* \to \mathbb{R}$ with*

$$v\left(g\right) = \sum_{t=1}^{T} v_{D_t}\left(g\left(D_t\right)\right)$$

for $g \in \overline{G} \cap G_{\{D_t\}}$, represents $\succ$.

Moreover, if there is another collection of continuous linear functions, $v'_D : \overline{\mathcal{A}} \to \mathbb{R}$, such that

$$v'\left(g\right) = \sum_{t=1}^{T} v'_{D_t}\left(g\left(D_t\right)\right)$$

represents $\succ$ on $\overline{G} \cap G^*$, then there are constants $a > 0$ and $\{b_D \,|D \in \mathcal{F}\}$, such that $v'_D = b_D + a v_D$ for each $D \in \mathcal{F}$.

Proof: See Appendix.

The expectation for probability measures on $\mathcal{F}$ can only be calculated directly for simple functions, as in the previous case. For general functions it is defined as an appropriate limit:

Definition 7 *(Based on Definition 10.2 in Fishburn (1970)): For a countably additive probability measure* π *on* $\mathcal{F}$ *and a bounded measurable function* $g : I \rightarrow \mathbb{R}$, *let* $\langle \varphi_n \rangle$ *be a sequence of simple functions,* $\varphi_n : I \rightarrow \mathbb{R}$,[29] *that converge from below to* φ. *Then define*

$$E_\pi [\varphi] := \sup \{E_\pi [\varphi_n] \,| n = 1, 2, ...\}$$

to be the expectation of φ *under* π.

Fishburn establishes that this expectation is well defined. To apply the definition, the better set $\{A \,| A \succ_D B\}$ must not change too much for a small change in D. For this purpose denote by $d_{\mathcal{A}} (\mathcal{A}', \mathcal{A}'')$ the Hausdorff metric with respect to the Hausdorff metric on $\mathcal{A}$, where $\mathcal{A}', \mathcal{A}'' \subseteq \mathcal{A}$. For $D, D' \in \mathbf{2}^I$, let $d_{\mathbf{2}^I} (D, D')$ be the Hausdorff metric with respect to the metric on I.

Axiom 7 *(Continuity in I): If the sequence* $\langle D^n \rangle$ *in* $\mathcal{F}$ *converges in terms of* $d_{\mathbf{2}^I}$, *then* $\{A \,| A \succ_{D^n} B\}$ *converges in terms of* $d_{\mathcal{A}}$.

Theorem 1': $\succ$ *satisfies Axioms 1-5 and 7, if and only if there are a bounded set of positive numbers* $\{l(s) \,| s \in S\}$, *a conditional probability measure* $\mu (s \,| i)$, *continuous in* I, *and a probability measure* π *on* I, *such that the function* $V : G \rightarrow \mathbb{R}$,

$$V(g) = E_\pi \left[\int_S l(s) \max_{\alpha \in g(i)} (\alpha \cdot s) \, d\mu (s \,| i) \right]$$

represents $\succ$.

Proof: See Appendix.

Straight forward changes to the proof of Theorem 1 establish the result for $\succ$ constrained to all simple acts, G^*. In addition, the simple acts are shown to be dense in G in the topology defined on G. Ensuring that Definition 7 applies completes the proof.

Recall that, if the information space I is finite, Axiom 6 is not an innocuous assumption,

[29] As for simple acts, the value of a simple function depends only on some finite and measurable partition $\{D_t \,| t \in \{1, .., T\}\}$ of I . $E_\pi [\varphi_n] := \sum_{t=1}^{T} \pi (D_t) \varphi_n (D_t)$.

because it limits the cardinality of the taste space, S. For the case where I is a metric space, Axiom 6 is amended in order to guarantee sufficient tightness to deal with a continuous S.

The next definition provides a measure of how much a set A is preferred over set B in terms of how much the menu corresponding to the whole prize space, $\{\langle 1,z\rangle \,|z\in Z\}$, is preferred over the worst menu.

Definition 8: *Given $D\in\mathcal{F}$, let $\underline{z}$ be a prize, such that $A\succeq_D \{\langle 1,\underline{z}\rangle\}$ for all $A\in\mathcal{A}$.*[30] *For $A,B\in\mathcal{A}$, define $p_{A,B}(D)\in(-1,1)$, such that*
i) for $A \succsim_D B$, $p=p_{A,B}(D)$ solves

$$\frac{1}{1+p}A+\frac{p}{1+p}\{\langle 1,\underline{z}\rangle\}\sim_D \frac{1}{1+p}B+\frac{p}{1+p}\{\langle 1,z\rangle\,|z\in Z\},$$

ii) for $B\succ_D A$, $p_{A,B}(D)=-p_{B,A}(D)$.
Call $p_{A,B}(D)$ the cost of having to choose from B instead of A under event D.

Note that $p_{A,B}(D)\neq 0$ implies that D is nontrivial for any representation as in Theorem 1', because otherwise $A\sim_D B$ for all $A,B\in\mathcal{A}$.

If two sequences of menus, $\langle A_n\rangle$ and $\langle B_n\rangle$, approach each other, then the cost of having to choose from B_n rather than A_n vanishes under every event. However, the ratio of such costs may have a well defined limit.

Axiom 6' *(Richness and Tightness of Information): If $\langle A_n\rangle,\langle B_n\rangle,\langle C_n\rangle\subseteq\mathcal{A}$ converge in the Hausdorff topology, then*

$$\frac{p_{C_n,A_n\cup B_n}(D)}{p_{C_n,B_n}(D)}\nrightarrow 1$$

for some $D\in\mathcal{F}$ implies that there is $D'\in\mathcal{F}$, such that

$$\frac{p_{C_n,A_n}(D')}{p_{C_n,B_n}(D')}\nrightarrow 1.$$

Axiom 6' implies Axiom 6, where i is substituted by D. To see this, note that Axiom 6 holds trivially unless there is $D\in\mathcal{F}$, such that $A\cup B\nsim_D B$ and $A\sim_D B$. Define the constant sequences $A_n:=A$ and $B_n:=B$ and let $C_n:=C\succ_D A$. Then $p_{C_n,A_n}(D)=p_{C,A}(D)$ must be satisfied. This implies $p_{C,B}(D)=p_{C,A}(D)$ and $p_{C,A\cup B}(D)\neq p_{C,B}(D)$ or, in terms of the constant sequences,

$$\frac{p_{C_n,A_n\cup B_n}(D)}{p_{C_n,B_n}(D)}\nrightarrow 1.$$

[30] This prize exists, because Z is finite and because of *Monotonicity* (Axiom 5).

Thus, according to Axiom 6', there is $D' \in \mathcal{F}$ with

$$\frac{p_{C_n,A_n}(D')}{p_{C_n,B_n}(D')} \nrightarrow 1.$$

Hence $A \nsim_{D'} B$, and Axiom 6 is satisfied.[31]

Again, let ϕ represent objective probabilities of events.

Theorem 2': *If, given* $\phi : I \to \mathbb{R}_+$,

$$V(g) = E_\phi \left[\int_S l(s) \max_{\alpha \in g(i)} (\alpha \cdot s) \, d\mu(s|i) \right]$$

represents $\succ$*, then statements i and ii below are equivalent:*
i) $\succ$ *satisfies Axiom 6',*
ii) the measure $\mu(s|D) := E_\phi[\mu(s|i)|D]$ *is unique for all* $D \in \mathcal{F}$ *and up to* ϕ*-measure zero changes,* $l(s)$ *is unique for all* $s \in S$ *up to common linear transformations and changes on a set* $S' \subset S$ *with* $E_\phi \left[\int_{S'} d\mu(s|i) \right] = 0$.

Proof: See Appendix.

The discussion of Theorem 2 applies here, including the implications of a partial failure of Axiom 6 and Axiom 6', respectively.

In the proof of Theorem 2, the individual taste $s \in S^*$ is identified by two menus: one is preferred over the other under taste s, but they generate the same payoff under every other taste. If S is continuous, the complication is that making a menu preferred less by a finite amount under one taste will invariably make it worse under similar tastes,[32] too. Therefore, individual tastes can only be identified in the limiting case, where the less preferred and the more preferred menu approach each other. In this limiting case the cost of having to choose from the less preferred menu instead of the more preferred menu tends to zero. Axiom 6' allows statements about the ratio of these costs for two different pairs of menus in the limit. The formal argument is much more tedious, but the main ideas of the proof are the same as for Theorem 2.

[31] If $p_{C_n,B_n}(D) \nrightarrow 0$, Axiom 6 trivially implies Axiom 6'. Thus, Axiom 6' is only stronger than Axiom 6 for $p_{C_n,B_n}(D) \to 0$.

[32] When tastes are viewed as vectors in $\mathbb{R}^k_+$.

3. Subjective versus Objective Probabilities of Information

Section 2 answers two questions about choice on our domain: under what conditions does a decision maker with preference for flexibility, à la Kreps, behave like a standard agent in every other way? In particular DM should choose *as if* lotteries were evaluated according to expected utility and there should be a subjective probability measure over information, according to which acts are also evaluated by expected utility. Further, information should only lead to updated beliefs. Theorems 1 (meaning Theorem 1 and 1') give the answer to this question by establishing the representation

$$V(g) = E_{\pi}\left[\int_S l(s) \max_{\alpha \in g(i)} (\alpha \cdot s)\, d\mu(s\,|i)\right]$$

for some subjective probability measure π on I.

Now suppose DM's preferences can be represented thus, where $\pi \equiv \phi$. ϕ is the objective probability distribution, which corresponds to the true observed frequencies of information. When is the distinction of probabilities and utilities meaningful? Theorems 2 give a condition under which $l(s)$ and $\mu(s\,|i)$ in the representation V are unique in the appropriate sense.

The two questions can be answered independently of each other. Still, there is a gap between Theorems 1 and Theorems 2: a combined result that gives conditions under which the representation V exists, π is unique and $l(s)$ and $\mu(s\,|i)$ are unique in the appropriate sense, has yet to be established.

If π is unique in Theorems 1, then Theorems 2 can be applied to establish uniqueness of $l(s)$ and $\mu(s\,|i)$, based on this unique π. Determining π uniquely in Theorems 1 is analogous to the classical problem addressed by AA. Their unique identification of probabilities of exogenous states is based on the assumption of a *state independence* of the ranking of outcomes. The difference is that they consider acts with lotteries (instead of menus of lotteries) as outcomes, so there is no room for preference for flexibility in their setup. The combination of state independence and Axiom 6 would rule out any preference for flexibility. Thus, the assumption of state independence has to be weakened, to be useful in or context.

Definition 9: *Let $\Omega \subseteq Z$ denote a non-degenerate set of prizes and $\Delta(\Omega)$ the set of all lotteries with support in Ω. Let $\Psi(\Delta(\Omega)) \subseteq \mathcal{A}$ be the set of all menus of lotteries that have support in Ω.*

Axiom 8 *(Partial Information Independence): There is $\Omega \subseteq Z$, such that for $A, B \in \Psi(\Delta(\Omega))$, $A \succ_D B$ for some event $D \in \mathcal{F}$ implies $A \succ_{D'} B$ for all nontrivial $D' \in \mathcal{F}$. If $\succ$*

satisfies the same condition for $\Omega' \subseteq Z$, then also for $\Omega \cup \Omega'$.

To illustrate Axiom 8, consider $\Omega = \{m, 0\}$ to consist of the prizes "money" and "nothing". The first part of Axiom 8 then requires all menus that consist only of lotteries over money and nothing to be ranked independent of information. To motivate the requirement, it is sufficient, that the value of money (versus nothing) is assumed to be independent of information.

Having assumed state independent rankings, AA move on to assume cardinally state independent rankings (or state independent utilities). This cannot be assumed in terms of an axiom. Instead it is a constraint on the class of representations for which they establish their uniqueness result.[33] In our representation it would amount to requiring that $\int_S l(s) \max_{\alpha \in A} (\alpha \cdot s)\, d\mu(s\,|i)$ is independent of $i \in I$ for all $A \in \Psi$. But if $\Psi \subset \mathcal{A}$ is a generic collection of menus, then this might not be consistent with $\succ$, which applies to all of G.[34] Thus, the requirement must be limited to *particular* collections of menus $\Psi(\Delta(\Omega)) \subseteq \mathcal{A}$, as defined in Definition 9.

There clearly is a representation, where the value of menus in $\Psi(\Delta(\Omega))$ does not depend on information $i \in I$, and which is consistent across $\mathcal{A}$. To construct it, let the utility $l(s)\, s_x$ of prizes $x \in \Omega$ be independent of the taste, $s \in S$. AA's argument implies that $\pi(D)$ is unique for every event $D \in \mathcal{F}$. Given this probability measure π, Theorems 2 imply that $l(s)$ is unique in the appropriate sense. Thus, $l(s)\, s_x$ is independent of s for all $x \in \Omega$ for *all* of the representations, for which the value of menus in $\Psi(\Delta(\Omega))$ does not depend on information $i \in I$. The appropriate uniqueness of $\mu(s\,|i)$ also follows from Theorems 2.

Once AA restrict attention to representations with state independent utilities, there is no arbitrariness in their model. In our model, there may be: preference for flexibility implies $\Omega \subset Z$. Hence, there could be $\Omega' \subset Z$, for which $\succ$ also satisfies our assumption, while for $\Omega \cup \Omega'$ it does not. Either the prizes in Ω or those in Ω' could, then, be assigned a cardinal ranking, which is independent of information. While there is no inherent argument to favor one over the other, the two assumptions clearly lead to different representations. This arbitrariness would render the uniqueness result meaningless. The second part of Axiom 8 rules out this scenario, suggesting the following definition:

Definition 10: *If $\succ$ satisfies Axiom 8, let $\Omega^* \subseteq Z$ be the largest set, for which it does.*

[33] Compare the discussion of Theorem 2.

[34] For a simple example of such inconsistency consider $\Psi = \{\{\alpha\}, \{\beta\}, \{\gamma\}\}$ but, for some $p \in (0,1)$ and $D, D' \in \mathcal{F}$, $\{p\alpha + (1-p)\gamma\} \succ_D \{\beta\} \succ_{D'} \{p\alpha + (1-p)\gamma\}$. Since $\int_S l(s) \sup_{\alpha \in A} (\alpha \cdot s)\, d\mu(s\,|i)$ is linear, it can not be independent of $i \in I$.

Theorem 3:
$\succ$ satisfies Axioms 1-6 (6' and 7 if I is metric) and Axiom 8, if and only if $\succ$ can be represented by

$$V(g) = E_\pi \left[\int_S l(s) \max_{\alpha \in g(i)} (\alpha \cdot s) \, d\mu(s|i) \right]$$

as in Theorems 1, where the evaluation of menus in $\Psi(\Delta(\Omega^*))$ is independent of information. For this representation π is unique, $\mu(s|D)$ is unique for all $D \in \mathcal{F}$ and up to π-measure zero changes and $l(s)\, s_x$ is constant on S for all $x \in \Omega^*$ and up to changes on a set S' with $E_\pi \left[\int_{S'} d\mu(s|i) \right] = 0$.

Proof: In the class of representations, where $\int_S l(s) \max_{\alpha \in A} (\alpha \cdot s) \, d\mu(s|i)$ does not depend on $i \in I$, the uniqueness of π follows in complete analogy to the corresponding result in AA. Given a unique π, Theorem 2 implies the appropriate uniqueness of $\mu(s|D)$ and $l(s)$. Because a representation with $l(s)\, s_x$ constant on S for all $x \in \Omega^*$ clearly exists, the unique representation must have this feature. ■

While AA are specifically interested in a situation without objective probabilities of exogenous states, we are not. As before, let π be the subjective probabilities DM assigns to information and ϕ the objective probabilities corresponding to observed frequencies of information. It is easy to modify Axiom 8, such that it implies $\pi \equiv \phi$ in Theorem 3.[35]

Even if π in the representation V is not unique, a representation based on ϕ may not exist. For example an event $D \in \mathcal{F}$ that is trivial according to π but not according to ϕ rules out a representation based on ϕ.

Proposition 3: *Suppose V based on π represents $\succ$ and I and S^* both have the same cardinality. If $\succ$ satisfies Axiom 6, then there is a neighborhood of π in $\mathbb{R}^T$, such that for*

[35] Strengthen Axiom 8 the following way. *(Objective Probabilities): There is $\Omega \subseteq Z$, such that for $A, B \in \Psi(\Delta(\Omega))$ and nontrivial $D, D' \in \mathcal{F}$,*

$$\frac{\phi(D')}{\phi(D) + \phi(D')} h_D^A + \frac{\phi(D)}{\phi(D) + \phi(D')} h_{D'}^B \sim \frac{\phi(D)}{\phi(D) + \phi(D')} h_{D'}^A + \frac{\phi(D')}{\phi(D) + \phi(D')} h_D^B.$$

If $\succ$ satisfies the same condition for $\Omega' \subseteq Z$, then also for $\Omega \cup \Omega'$.

This implies Axiom 8. It also implies that $V(g_D^A) - V(g_D^B) = \left(V(g_{D'}^A) - V(g_{D'}^B)\right) \frac{\phi(D)}{\phi(D')}$ for $A, B \in \Psi(\Delta(\Omega))$.

ϕ in this neighborhood, there is a representation

$$\widehat{V}(g) = E_{\phi}\left[\sum_{S^*} \widehat{l}(s) \max_{\alpha \in g(i)} (\alpha \cdot s)\, d\widehat{\mu}(s\,|i)\right],$$

where $\widehat{l}(s)$ and $\widehat{\mu}(s\,|i)$ are "close" to $l(s)$ and $\mu(s\,|i)$.

Proof: See Appendix.

This robustness result is interesting for applications where beliefs are used as a forecast of period 2 choice: if the information space is just large enough to identify the representation uniquely, given objective probabilities of information, and if the observer and the decision maker disagree slightly in their perception of those "objective" probabilities, then Theorem 2 can be applied[36] and the unique subjective probabilities of future tastes provided by Theorem 2 are at least a good approximation of DM's true beliefs.

The following lemma is at the heart of the proof of Proposition 3:

Lemma 1: *If I is finite, V based on π and $\widehat{V}$ based on ϕ both represent $\succ$, then*

$$\frac{\phi(i)}{\phi(j)} = \frac{\pi(i)}{\pi(j)} \frac{\int_S \frac{l(s)}{\widehat{l}(s)} d\mu(s\,|i)}{\int_S \frac{l(s)}{\widehat{l}(s)} d\mu(s\,|j)}$$

has to hold for all nontrivial $i, j \in I$.

In Proposition 3, I and S^* both have cardinality T. To illustrate the proof of the proposition, let $\mu(s) \in \mathbb{R}_+^T$ denote the vector with i-th component $\mu(s\,|i)$ and let $\pi \odot \mu(s) \in \mathbb{R}_+^T$ be the component wise product of this vector with π. On S^* there is a positive solution $\widehat{l}(s)$ of the system of equations

$$\phi(i) \propto \pi(i) \sum_{S^*} \frac{l(s)}{\widehat{l}(s)} \mu(s\,|i),$$

whenever ϕ is in the positive linear span of the T vectors $\pi \odot \mu(s)$.[37] It can be shown that those T vectors are linearly independent. Thus, they span $\mathbb{R}^T$. π can be expressed

[36] There is a representation based on ϕ, even if DM truly believes π.

[37] The positive linear span of a collection of vector $\{v_s\,|s \in S^*\}$ is the set of vectors $\left\{\sum_{S^*} k_s v_s\,|k_s \in \mathbb{R}_+\right\}$.

as a linear combination, which assigns unit weight to the vector $\pi \odot \mu(s)$ for each $s \in S^*$. Therefore, π is in the interior of the positive linear span of those vectors, implying that there is a neighborhood of π in $\mathbb{R}^T_+$, which is also in their positive linear span.[38] The solution of a linear system of equations is continuous in all parameters. Hence, for ϕ in a small enough neighborhood of π, $\frac{l(s)}{\hat{l}(s)} \approx 1$, which implies $\frac{\mu(s|i)}{\hat{\mu}(s|i)} \approx 1$.

4. Incomplete Contracts

As illustrated by the example in the introduction, our domain accommodates incomplete contracts very naturally: consider I as the event space. At the time that two parties write a contract, I is describable. Further, there are unforeseen or indescribable contingencies, which may impact parties' preferences over prizes.

Information about those contingencies may be asymmetric. To illustrate the problem this poses for two parties trying to agree on an efficient contract, reconsider the example:

Second Continuation of Example: In order to agree on an efficient contract both the retailer and the supplier must be able to rank contracts which may condition on information about consumer confidence.

The retailer understands the indescribable contingencies determining her idiosyncratic demand: she can assign probabilities $\mu(h|H)$ and $\mu(h|L)$, conditional on information. Comparing contracts, the objective function of the retailer is

$$V(g) = \sum_{i \in \{H,L\}} \phi(i) \left[\sum_{s \in \{h,l\}} \mu(s|i) \max_{q \in g(i)} (v_s(q)) \right],$$

as stated above.

Suppose that only the retailer's valuations depend on the idiosyncratic demand she faces. The possible valuations of the supplier only depends on consumer confidence, $i \in \{H, L\}$. The supplier values supplying q_h and q_l at $w_i(q_h)$ and $w_i(q_l)$, respectively. Further suppose that the supplier does not understand the contingencies determining the retailer's idiosyncratic demand. He can, therefore, not assign them probabilities. Because of this asymmetry in information, the supplier's objective function is only specified conditional on the retailer's

[38] Theorem 2 states that S^* has at most cardinality T. This argument makes clear why Proposition 3 requires it to have at least cardinality T: If not, then the T vectors $\pi \odot \mu(s)$ can not span $\mathbb{R}^T$. Hence the interior of their positive span is not open in $\mathbb{R}^T$.

beliefs:

$$W(g) = \sum_{i\in\{H,L\}} \phi(i) \left[\sum_{s\in\{h,l\}} \mu(s|i)\, w_i\left(q^*_{g(i)}(s)\right) \right]$$

where $q^*_A(s) := \arg\max_{q\in A}(v_s(q))$. The supplier's ability to rank contracts depends on learning the retailer's beliefs.

More generally, consider a principal and an agent. For simplicity, suppose that only the principal's valuations depend on indescribable contingencies, which are unforeseen only by the agent. Let S denote the principal's taste space.[39]

Actions are observable, so there is no moral hazard. An action pair specifies actions to be taken by the principal and the agent, respectively. Each action pair corresponds to a lottery (potentially dependent on the event $i \in I$) over relevant prizes.[40]

Because there is no uncertainty about the agent's payoff, which resolves between writing the contract and taking an action and on which the contract cannot condition, an efficient contract will assign some control rights to the principal: it will specify a collection of action pairs for every describable event $i \in I$, from which the principal can choose at a later time.

The reduced form of such an incomplete contract, g, specifies a menu of lotteries over prizes for every event $i \in I$, $g : I \to \mathcal{A}$. The principal may choose from $g(i)$, if state i realizes. The interpretation is that by the time he chooses, the uncertainty about the contingencies that determine his taste over prizes, $s \in S$, has been resolved. Thus, from the principal's point of view, the incomplete contract is nothing more than an act in the domain considered in the previous sections.[41]

To agree on an efficient contract, both parties must be able to rank all contracts. Let $\alpha_s(A) := \arg\max_{\alpha\in A}(\alpha \cdot s)$.[42] According to the previous sections, a principal satisfying our

[39] This simplifying assumption can be relaxed. There could be a space S_{agent} and another space $S_{principal}$ to capture possible future preferences of the respective parties. Further, information might be less asymmetric. However, it is crucial that each party's information about its own future tastes encompasses all the information the other party might have. This constraint seems reasonable in many contexts: Contingencies that are meaningful to both parties should be part of the describable event space.

[40] If contingencies have impact on the effect of actions on the probabilities of prizes, they are relevant for both parties. Again, that they are then describable seems reasonable in many contexts.

Actions may be complicated objects, involving multiple dimensions like leisure or work, monetary payment, schooling etc. Those may translate directly into prizes. They might also generate a probability distribution over prizes, which is independent of $s \in S$. Prices may be thought of as costly or beneficial to the principal, the agent or both.

[41] The mapping from action pairs to lotteries must be surjective, such that every act can be identified with some contract.

[42] The arg max exists, because menus are compact. If it is not unique, ties can be broken in the agent's interest.

axioms ranks contracts according to

$$V(g) = E_{\phi}\left[\int_S l(s)\left(\alpha_s(g(i)) \cdot s\right) d\mu(s\,|i)\right],$$

where $\mu(s\,|i)$ are unique in the appropriate sense.

What about the agent? She must be able to assign an event dependent cost, $c(x,i)$, to every prize $x \in Z$. Let $c(i) \in \mathbb{R}^k$ be the vector of these costs.[43] Further she must assess probabilities of events according to objective probabilities, $\phi : I \to [0,1]$.[44] Lastly, assume that the agent believes that the representation above reveals the principal's assessment of the uncertainty over her own future tastes. Then, conditional on learning the principal's ranking of contracts, the agent ranks contracts according to

$$W(g) = -E_{\phi}\left[\int_S \left(\alpha_s(g(i)) \cdot c(i)\right) d\mu(s\,|i)\right].$$

Note that $W(g)$ depends on the conditional subjective probabilities, $\mu(s\,|i)$, as perceived by the principal but not on the taste dependent intensities, $l(s)$. In our axiomatic setup these two are distinct concepts.

The assumption that rankings of contracts are commonly known is usually required in contract theory and justified by some informal story of learning from past observations.[45] As this assumption is not our focus, we make it without doing the game theoretic complexity of the contracting problem justice. Instead we address the additional assumption required in the *incomplete* contracting literature: In order to allow both parties to rank all contracts, it has to be assumed that they believe in the same probability distribution over utility-payoffs, ex ante.[46] This ad hoc assumption is made for lack of a useful choice theoretic model of the bounded rationality involved. It is troubling in the context of unforeseen contingencies, where asymmetric information seems natural. Our domain is not only well suited to describe

[43] This cost might be a benefit. In fact, linear transformations of the cost schedule are irrelevant. It is labeled as a cost only because in many applications a benefit for the principal corresponds to a cost for the agent.

[44] In general she may asses them according to $\pi_{agent} : I \to [0,1]$. This assessment may or may not agree with the principal's assessment, $\pi_{principal}$. If it does not, the question arises, whether parties update their believes after learning the other party's assessment. Typically probabilities of describable events are assumed to be assessed according to objective probabilities, $\phi : I \to [0,1]$, by both parties.

[45] For example, contracts signed in a large homogenuous population might be observed.

[46] See, for example, Hart and Moore (1988). Maskin and Tirole (1999) show that, even for the case of unobservable actions and under restrictions that are often satisfied in the literature, this assumption allows to achieve all those payoffs with incomplete contracts that could be achieved by a hypothetical complete contract, which can condition on the indescribable contingencies.

the type of incomplete contracts laid out above, but for those contracts our axioms also give choice theoretic content to the assumption of common beliefs.

Forecasting behavior based on beliefs elicited from the principal's ranking of contracts is an inductive step. The underlying assumption that the class of representations we consider maps his actual decision making process is not directly falsifiable. However, it can be falsified indirectly on the basis of its predictions. The agent might, thus, be comfortable to make this assumption not only because it is intuitive from introspection, but also because past agents have found it to generate the right predictions.

More complicated situations should be addressed. For example the principal might not foresee the indescribable contingencies relevant for his future taste perfectly, himself. In that case the agent might want to distinguish those beliefs based on information, from those representing ambiguity.[47]

Even when ambiguity is not an issue, both parties might have exclusive information about indescribable contingencies relevant for the other party's future taste. Then each party may update their beliefs after learning the other party's preferences. Sequentiality and a model of learning would have to be incorporated to generate common beliefs ex ante. We leave this as an area for future research.

5. Conclusion

The notion of a taste space is attractive, because in principle it allows distinction of consequences and probabilities. In the context of preference for flexibility this distinction, in turn, reconciles choice with Bayesian decision making, which is at the heart of the notion of rationality. However, identifying the two conceptually distinct components through preferences has proven difficult. This paper proposes a richness requirement on an exogenous state space, interpreted as information, which allows their unique identification. The exogenous state is chosen by Nature. We conclude by suggesting a reinterpretation:

Consider information, which can be determined by an experimenter instead of Nature. In many contexts the experimenter cannot credibly offer alternative informations about states of nature. If he can, then information is typically interpreted as just another dimension of the consumption bundle. A context, where DM's ranking changes with "information" conveyed by the experimenter and where it is not interpreted as part of the consumption bundle, is framing. A frame is information, which seems to be irrelevant to the rational evaluation of

[47] A richer model would be required to accomodate this possibility. Dekel, Lipman and Rustichini (1998-b) point out that contingencies, which DM can not foresee, are difficult to reconcile with the notion of a state space. We know of no work that uniquely identifies the two types of beliefs, even in a setting without subjective states.

alternatives, but which may affect choice.

One possible interpretation of frames is suggested by Sher and McKenzie (2006). They propose that logically equivalent frames may not be informationally equivalent: they convey information about the sender's knowledge about relevant aspects of the choice situation. Alternatively, DM might be susceptible to frames for a multitude of other reasons, like hedonic forces, cognitive load or reference dependent preferences.

Consider pairs (A, f), where A is a menu and f is a frame.[48] DM has preferences over the domain $\mathcal{A} \times I$, where I is now interpreted as the collection of all frames, $f \in I$. The adaptation of our axioms to this new domain is straight forward. To paraphrase the identifying assumption, Richness, in this context: if there is preference for flexibility with respect to two indifferent menus, then those can be reframed, so as to break the indifference. Hence, frames do have an impact on the evaluation of menus.

The representation implied by the axioms suggests interpreting DM's susceptibility to frames as Bayesian decision making. The underlying model is not specified, but the uniqueness result of Theorem 2 lets us classify the information content of changing frame f to frame f' by comparing the probability distributions $\mu(s\,|f)$ and $\mu(s\,|f')$ they induce.

If DM truly was a Bayesian decision maker (in the sense specified by our model,[49]) then $\mu(s\,|f)$ should predict how often taste s governs her future choice. Whether and when it does, is an empirical question.

6. Appendix

6.1. Proof of Theorem 1

For any nontrivial event $i \in I$ (which exists according to Corollary 1), $V_i(A)$ and $v_i(A)$ agree on $\overline{\mathcal{A}}$ up to a positive affine transformation, as established by Corollary 2. Thus there is an event dependent, positive scaling factor $\pi'(i)$, such that $v_i(A) = \pi'(i)\, V_i(A)$ for all $A \in \overline{\mathcal{A}}$, where $\pi'(i) = 0$, if and only if i is trivial.

Let V' represent $\succ$ on G and $V' \equiv v$ on $\overline{G}$. Due to Continuity, there is a convex act $\overline{g} \in \overline{G}$ for all $g \in G$, such that $\overline{g}(i) \sim_i g(i)$. Then, according to Proposition 1, $V'(g) = V'(\overline{g}) = \sum_{i \in I} v_i(\overline{g}(i)) = \sum_{i \in I} \pi'(i)\, V_i(\overline{g}(i))$. According to Theorem DLRS, $V_i(\overline{g}(i)) = V_i(g(i))$.

Hence, $g \succ h$ implies $\sum_{i \in I} \pi'(i)\, V_i(g(i)) > \sum_{i \in I} \pi'(i)\, V_i(h(i))$. Therefore

$$V'(g) = \sum_{i \in I} \pi'(i) \left[\int_S l(s) \max_{\alpha \in g(i)} (\alpha \cdot s)\, d\mu_i(s) \right]$$

[48] This domain is suggested by Salant and Rubinstein (2007).

[49] This may be the case, even if she is not explicitly aware of the information content she assigns to frames. For example a reference point introduced by a frame might persist and influence future choice the way DM "expects" it to.

represents $\succ$. Since v is unique only up to positive affine transformations, $\pi'(i)$ can be normalized to be a probability measure, $\pi(i)$.

Interpreting $\mu(s|i) := \mu_i(s)$ as a conditional probability measure over the taste space S, define

$$V(g) := \sum_{i \in I} \pi(i) \left[\int_S l(s) \max_{\alpha \in g(i)} (\alpha \cdot s)\, d\mu(s|i) \right]$$

to establish the sufficiency statement in Theorem 1. That Axioms 1-5 are necessary for the existence of the representation is straight forward to verify. ■

6.2. Proof of Theorem 2

To establish that Axiom 6 is *sufficient* for the uniqueness, consider $\Delta(Z)$ as a $k-1$ dimensional simplex of lotteries. We now construct the menus K_ρ and K_ρ^ε that allow identification of individual tastes in a sequence of four claims.

Definitions 11:

• *For $X \subseteq Z$, let $S_X \subseteq S$ denote all tastes under which all $x \in X$ are the worst prizes, and all $x \in Z \backslash X$ are strictly better. For example, $S_{\{z\}}$ denotes all tastes under which $z \in Z$ is the unique worst prize. Let $\sharp(X)$ be the cardinality of X.*

•
$$K_X := \left\{ \alpha \in \Delta(Z) \,\middle|\, \sum_{y \in Z \backslash X} \alpha^2(y) = r^2 \text{ and } \alpha(x) = \alpha(y) \text{ for } x, y \in X \right\}.$$
[50]

•
$$L_X := \left\{ \alpha \in \Delta(Z) \,\middle|\, \sum_{y \in Z \backslash X} \alpha^2(y) = (r \cos \varepsilon)^2 \text{ and } \alpha(z) = 1 - \sum_{y \in Z \backslash X} \alpha(y) \text{ for some } z \in X \right\}.$$

• *Let α^X be the projection of lottery α into the hyperplane with $\alpha(x) = 0$ for all $x \in X$. $\measuredangle\left(\alpha^X, \beta^X\right)$ is the angle between lottery α and β, when projected into this plane.*

•
$$N_X^\varepsilon(\alpha) = \left\{ \beta \in K_X \,\middle|\, \measuredangle\left(\alpha^X, \beta^X\right) < \varepsilon \right\}.$$

Figure 1a shows $K_{\{z\}}$ for the case $Z = \{x, y, z\}$ and $r^2 = \frac{1}{2}$.[51]

[50] The projection of its elements α into the hyperplane with $\alpha(x) = 0$ for all $x \in X$ is the positive orthant of the $k - c$ dimensional sphere of radius $r < \frac{1}{\sqrt{k-1}}$ around $\langle 1, x \rangle$.

[51] In the case of only three prizes, $K_{\{z\}}$ is a subset of the Marshak triangle with prize z at the origin.

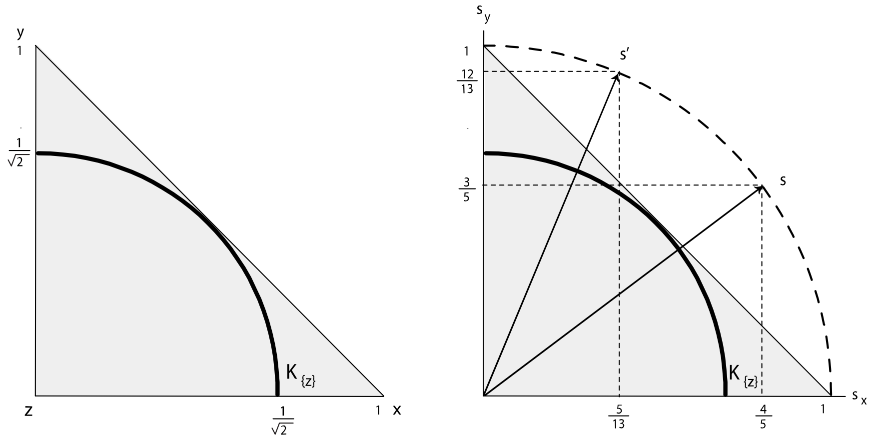

Figure 1a: $K_{\{z\}}$ in the Marshak triangle. b: Endogenous state space $S^* = \{s, s'\}$.

K_X allows to identify individual tastes in S_X. For $\sharp(X) = 1$, L_X is empty. According to Claim 2 below, L_X contains the most preferred lottery in $s' \in S_Y$ for $Y \subset X$. Adding this menu to K_X makes sure that the utility under those tastes is not reduced, when $N_X^\varepsilon(\alpha^X)$ is removed from K_X, if α is relevant under some taste $s \in S_X$. Note that for $s \in S_X$ the most preferred lottery in $K_X \cup L_X$ is in K_X. Lastly, $K_X \cap L_X = \emptyset$.

Claim 1: $\arg\max_{\alpha \in K_X} (\alpha \cdot s)$ *is a singleton and*

$$\arg\max_{\alpha \in K_X} (\alpha \cdot s) \neq \arg\max_{\alpha \in K_X} (\alpha \cdot s')$$

for all $s \neq s'$, $s \in S_X$, $s' \in S_Y$ *and* Y *not a strict subset of* X.

Proof: Y not being a strict subset of X implies that for all $y \in Y \backslash X$, $s'_y = 0$ and $s_y > 0$. Thus, $\max_{\alpha \in K_X} (\alpha \cdot s)$ is solved uniquely for α parallel to s, implying $\alpha(y) > 0$. $\max_{\alpha \in K_X} (\alpha \cdot s')$ is solved for $\alpha(y) = 0$. $\square$

For $s \in S_X$, the menu $\{\alpha \in K_X \,|\alpha \text{ is parallel to } s\}$ is then a singleton, as K_X is a subset of a sphere. Figure 1b illustrates for the case $Z = \{x, y, z\}$ and $S^* = \left\{s = \left(\frac{4}{5}, \frac{3}{5}, 0\right), s' = \left(\frac{5}{13}, \frac{12}{13}, 0\right)\right\}$.

Definition 12: *Let* $\alpha_s^X \in K_X$ *be the lottery that identifies* s. *Conversely, let* $s_\alpha^X \in S_X$ *be the taste parallel to lottery* $\alpha \in K_X$.

Claim 1 implies that $\measuredangle\left(\alpha_s^X, \alpha_{s'}^X\right) > 0$ for $s \neq s'$, $s \in S_X$, $s' \in S_Y$ and Y not a strict subset of X.

Claim 2: *For* $s' \in S_Y$ *with* $Y \subset X$, $\measuredangle\left(\alpha_s^Y, \alpha_{s'}^Y\right) > \delta$ *for all* $s \in S_X$, *and* $r < \frac{1}{\sqrt{k-1}}$, *there is* $\overline{\varepsilon}(\delta) > 0$, *such that for* $\varepsilon < \overline{\varepsilon}(\delta)$,

$$\arg\max_{\alpha \in K_X \cup L_X} (\alpha \cdot s') \subseteq L_X.$$

Proof: $\sharp(Y) < \sharp(X) \leq k - 1$. Suppose, for $s' \in S_Y$, α was the most preferred lottery in K_X. Split $\alpha \cdot s'$ into two parts: $\sum_{x \in Z \backslash X} \alpha(x) s'_x$ and $\sum_{x \in X} \alpha(x) s'_x$. By virtue of being in K_X, α will assign the same weight to all $x \in X$. For $x \in Y$, $s'_x = 0$. All prizes in $X \backslash Y$ might be valued equally:

$$\sum_{x \in X} \alpha(x) s'_x \leq \left(\max_{x \in X} s'_x\right) \left(1 - \sum_{x \in Z \backslash X} \alpha(x)\right) \frac{\sharp(X) - \sharp(Y)}{\sharp(X)}.$$

Now $\sharp(Y) \geq 1$ and $\sharp(X) \leq k-1$. Therefore,

$$\sum_{x \in X} \alpha(x) s'_x \leq \left(\max_{x \in X} s'_x\right) \left(1 - \sum_{x \in Z \backslash X} \alpha(x)\right) \frac{k-2}{k-1}.$$

By choosing $\beta \in L_X$ with $\measuredangle\left(\beta^X, \alpha^X\right) = 0$, DM can at least guarantee herself

$$\sum_{x \in Z \backslash X} \alpha(x) s'_x \cos\varepsilon + \left(\max_{x \in X} s'_x\right) \left(1 - \sum_{x \in Z \backslash X} \beta(x)\right),$$

where

$$1 - \sum_{x \in Z \backslash X} \beta(x) > 1 - \sum_{x \in Z \backslash X} \alpha(x)$$

for $\varepsilon > 0$. A sufficient condition for β to be preferred over α under taste s' is therefore:

$$(1 - \cos\varepsilon) \sum_{x \in Z \backslash X} \alpha(x) s'_x < \left(\max_{x \in X} s'_x\right) \left(1 - \sum_{x \in Z \backslash X} \alpha(x)\right) \frac{1}{k-1}.$$

Now $\measuredangle\left(\alpha_s^Y, \alpha_{s'}^Y\right) > \delta$ for $s \in S_X$ and $s' \in S_Y$ bounds $\max_{x \in X} s'_x$ away from zero and $r < \frac{1}{\sqrt{k-1}}$ bounds $\sum_{x \in Z \backslash X} \alpha(x)$ away from 1. Thus, there is $\overline{\varepsilon}(\delta)$, such that for $\varepsilon < \overline{\varepsilon}(\delta)$, β is preferred over α under taste s'. □

Definition 13: *Given a probability distribution over tastes, $\rho : S \to \mathbb{R}_+$, define a particular convex combination of menus:*

$$K_\rho^\varepsilon := \sum_{X \subseteq Z} \int_{s \in S_X} \left(\left(K_X \backslash N_X^\varepsilon\left(\alpha_s^X\right)\right) \cup L_X\right) \rho(s)\, ds.$$

Let K_ρ denote the case $\varepsilon = 0$.

Note that $K_\rho \succ_i K_\rho^\varepsilon$ for all $i \in I$ and $\varepsilon > 0$. Figure 2 illustrates $K_{\{z\}} \backslash N_{\{z\}}^\varepsilon\left(\alpha_s^{\{z\}}\right)$ for the case $Z = \{x, y, z\}$ and $S^* = \left\{s = \left(\frac{4}{5}, \frac{3}{5}, 0\right), s' = \left(\frac{5}{13}, \frac{12}{13}, 0\right)\right\}$.

Claim 3: *For the representation in Theorem 2, the cardinality of*

$$S^* := \bigcup_{i \in I} support\left(\mu\left(s \mid i\right)\right)$$

must be bounded above by the cardinality of I, T.

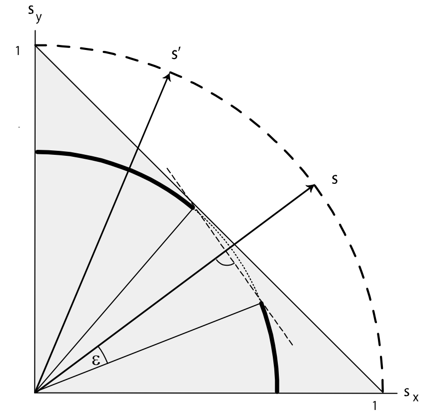

Figure 2: $K_{\{z\}}$ *versus* $K_{\{z\}} \backslash N^{\varepsilon}_{\{z\}} (\alpha_s)$.

Proof: Let $N \in [1, \infty]$ be the cardinality of S^*. We now show that $N \leq T$. Let ρ and ρ' have finite support. Then, for ε small enough, there is $s_1 \in S^*$, such that $\rho'(s_1) \neq \rho(s_1)$ implies $K^{\varepsilon}_{\rho} \nsim_i K^{\varepsilon}_{\rho} \cup K^{\varepsilon}_{\rho'}$ and $K^{\varepsilon}_{\rho'} \nsim_i K^{\varepsilon}_{\rho} \cup K^{\varepsilon}_{\rho'}$ for some $i \in I$. Suppose $T < N$. Then there are tastes $\{s_2, ..., s_{T+1}\} \in S^*$. Fix some

$$\rho(s) : \begin{cases} > 0 \text{ for } s \in \{s_1, ..., s_{T+1}\} \\ = 0 \text{ otherwise} \end{cases}.$$

According to the representation in Theorem 2 (taking into account that ρ has finite support here,)

$$V\left(g_i^{K_\rho}\right) - V\left(g_i^{K^{\varepsilon}_\rho}\right) = \phi(i) \sum_{s \in S} \left(V\left(g_i^{K_X \cup L_X}\right) - V\left(g_i^{\left(K_X \backslash N^{\varepsilon}_X\left(\alpha^X_s\right)\right) \cup L_X}\right)\right) \rho(s).$$

Thus, $V\left(g_i^{K_\rho}\right) - V\left(g_i^{K^{\varepsilon}_\rho}\right)$ is linear in $\rho(s)$. Look for

$$\rho'(s) : \begin{cases} > 0 \text{ for } s \in \{s_1, ..., s_{T+1}\} \\ = 0 \text{ otherwise} \end{cases}$$

which satisfies the T linear equations

$$V\left(g_j^{K_\rho}\right) - V\left(g_j^{K^{\varepsilon}_\rho}\right) = V\left(g_j^{K_{\rho'}}\right) - V\left(g_j^{K^{\varepsilon}_{\rho'}}\right)$$

for all $j \in I$. These conditions do not fully specify the $T+1$ components of $\rho'(s)$, and hence $\rho'(s_1) \neq \rho(s_1)$ must be possible, implying $K^{\varepsilon}_{\rho} \nsim_i K^{\varepsilon}_{\rho} \cup K^{\varepsilon}_{\rho'}$ for some $i \in I$. This contradicts Axiom 6. Hence, $N \leq T$ must indeed hold and, thus, the cardinality of S^* is bounded above by T. $\square$

According to the representation,

$$V\left(g_i^{K_\rho}\right) - V\left(g_i^{K^{\varepsilon}_\rho}\right) = \phi(i) \sum_{X \subseteq Z} \int_{s \in S_X} \left(V\left(g_i^{K_X \cup L_X}\right) - V\left(g_i^{\left(K_X \backslash N^{\varepsilon}_X\left(\alpha^X_s\right)\right) \cup L_X}\right)\right) \rho(s)\, ds.$$

For lottery $\alpha \in K_X$ and taste $s \in S_X$,

$$l(s)(\alpha \cdot s) = rl(s) \cos\left(\measuredangle\left(\alpha, \alpha^X_s\right)\right).$$

Claim 4: *For $\varepsilon > 0$ small enough,*

$$V\left(g_i^{K_\rho}\right) - V\left(g_i^{K_\rho^\varepsilon}\right) = \phi(i)\, r\,(1-\cos\varepsilon) \sum_{s\in S^*} l(s)\,\rho(s)\,\mu(s\,|i).$$

Proof: Consider how

$$l(s')\left(\max_{\alpha\in K_X\cup L_X}(\alpha\cdot s') - \max_{\alpha\in\left(K_X\setminus N_X^\varepsilon(\alpha_s^X)\right)\cup L_X}(\alpha\cdot s')\right)$$

is evaluated for any given taste s' and $\varepsilon > 0$ small enough. As S^* is finite, Claim 1 allows choosing δ, such that $\measuredangle\left(\alpha_s^X, \alpha_{s'}^X\right) > \delta$ for all $Y \nsubseteq X$ or $Y = X$, $s \in S_X$ and $s' \in S_Y$. Let $s \in S_X$ and $s' \in S_Y$ and consider the following three exhaustive cases:

i) $Y = X$: Then L_X is not relevant for the maximum of either menu. Therefore, consider instead

$$l(s')\left(\max_{\alpha\in K_X}(\alpha\cdot s') - \max_{\alpha\in K_X\setminus N_X^\varepsilon(\alpha_s^X)}(\alpha\cdot s')\right) = rl(s')\left(1-\cos\left(\max\left\{0,\left\{\varepsilon - \measuredangle\left(\alpha_s^X,\alpha_{s'}^X\right)\right\}\right\}\right)\right).$$ [52]

ii) $Y \nsubseteq X$: As $\measuredangle\left(\alpha_s^X, \alpha_{s'}^X\right) > \delta$, $\varepsilon < \delta$ ensures

$$\max_{\alpha\in K_X\cup L_X}\{\alpha\cdot s'\} - \max_{\alpha\in\left(K_X\setminus N_X^\varepsilon(\alpha_s^X)\right)\cup L_X}\{\alpha\cdot s'\} = 0.$$

iii) $Y \subset X$: Then Claim 2 states that for $\varepsilon < \bar{\varepsilon}(\delta)$, K_X is not relevant for the maximum. In that case,

$$\max_{\alpha\in K_X\cup L_X}\{\alpha\cdot s'\} - \max_{\alpha\in\left(K_X\setminus N_X^\varepsilon(\alpha_s^X)\right)\cup L_X}\{\alpha\cdot s'\} = 0.$$

So for ε small enough and for $s \in S_X$,

$$V\left(g_i^{K_X\cup L_X}\right) - V\left(g_i^{\left(K_X\setminus N_X^\varepsilon(\alpha_s^X)\right)\cup L_X}\right) = \phi(i)\int_{N_X^\varepsilon(\alpha_s^X)} rl\left(s_{\alpha'}^X\right)\left(1-\cos\left(\varepsilon - \measuredangle\left(\alpha_s^X,\alpha'\right)\right)\right)\mu\left(s_{\alpha'}^X\,|i\right)d\alpha'.$$

Consequently,

$$V\left(g_i^{K_\rho}\right) - V\left(g_i^{K_\rho^\varepsilon}\right) = \phi(i)\, r\sum_{X\subseteq Z}\int_{S_X}\rho(s)\int_{N_X^\varepsilon(\alpha_s^X)}\left(1-\cos\left(\varepsilon-\measuredangle\left(\alpha_s^X,\alpha'\right)\right)\right) l\left(s_{\alpha'}^X\right)\mu\left(s_{\alpha'}^X\,|i\right)d\alpha' ds.$$

Choose $\varepsilon > 0$ small enough, such that $\varepsilon < \delta$. Then

$$\int_{N_X^\varepsilon(\alpha_s^X)}\mu\left(s_{\alpha'}^X\,|i\right)d\alpha' = \mu(s\,|i).$$

Define $\rho(s) := 0$ for $s \notin S^*$. Then

$$V\left(g_i^{K_\rho}\right) - V\left(g_i^{K_\rho^\varepsilon}\right) = \phi(i)\, r\,(1-\cos\varepsilon)\sum_{s\in S^*} l(s)\,\rho(s)\,\mu(s\,|i).\ \square$$

[52] Compare Figure 2.

Now suppose there are two distinct representations, such that

$$V\left(g_i^A\right) = \phi\left(i\right)\sum_{S^*} l\left(s\right)\max_{\alpha\in A}\left(\alpha\cdot s\right)\mu\left(s\left|i\right.\right) + const(i)$$

and

$$\widehat{V}\left(g_i^A\right) = \phi\left(i\right)\sum_{\widehat{S^*}} \widehat{l}\left(s\right)\max_{\alpha\in A}\left(\alpha\cdot s\right)\widehat{\mu}\left(s\left|i\right.\right) + \widehat{const}(i).$$

First we show that $\widehat{l}\left(s\right) = l\left(s\right)$ up to a linear transformation and that $S^* = \widehat{S^*}$. Consider

$$\rho\left(s\right) := \frac{\frac{1}{l(s)}}{\sum_{S^*}\frac{1}{l(s)}} \text{ and } \widehat{\rho}\left(s\right) := \frac{\frac{1}{\widehat{l}(s)}}{\sum_{\widehat{S^*}}\frac{1}{\widehat{l}(s)}}.$$

Then, by Claim 4,

$$V\left(g_i^{K_\rho}\right) - V\left(g_i^{K_\rho^\varepsilon}\right) \propto \phi\left(i\right)\sum_{S^*}\mu\left(s\left|i\right.\right) = \phi\left(i\right)$$

for all $i \in I$, because $\mu\left(s\left|i\right.\right)$ is a probability measure. Thus, on the one hand,

$$V\left(g_i^{K_\rho}\right) - V\left(g_i^{K_\rho^\varepsilon}\right) = \left(V\left(g_j^{K_\rho}\right) - V\left(g_j^{K_\rho^\varepsilon}\right)\right)\frac{\phi\left(i\right)}{\phi\left(j\right)}$$

for all $j \in I$ and on the other hand

$$\widehat{V}\left(g_i^{K_{\widehat{\rho}}}\right) - \widehat{V}\left(g_i^{K_{\widehat{\rho}}^\xi}\right) = \left(\widehat{V}\left(g_j^{K_{\widehat{\rho}}}\right) - \widehat{V}\left(g_j^{K_{\widehat{\rho}}^\xi}\right)\right)\frac{\phi\left(i\right)}{\phi\left(j\right)}$$

for all $j \in I$, which implies

$$V\left(g_i^{K_{\widehat{\rho}}}\right) - V\left(g_i^{K_{\widehat{\rho}}^\xi}\right) = \left(V\left(g_j^{K_{\widehat{\rho}}}\right) - V\left(g_j^{K_{\widehat{\rho}}^\xi}\right)\right)\frac{\phi\left(i\right)}{\phi\left(j\right)},$$

because V and $\widehat{V}$ represent the same preference and are unique up to positive affine transformations. Choose ε and ξ, such that

$$V\left(g_i^{K_\rho}\right) - V\left(g_i^{K_\rho^\varepsilon}\right) = V\left(g_i^{K_{\widehat{\rho}}}\right) - V\left(g_i^{K_{\widehat{\rho}}^\xi}\right).$$

This is possible, because both sides are positive, continuous in ε and ξ and vanish in the limit of $\varepsilon \to 0$ and $\xi \to 0$, respectively. Then the observations above imply

$$V\left(g_j^{K_\rho}\right) - V\left(g_j^{K_\rho^\varepsilon}\right) = V\left(g_j^{K_{\widehat{\rho}}}\right) - V\left(g_j^{K_{\widehat{\rho}}^\xi}\right)$$

for all $j \in I$. In order to apply Axiom 6, translate this into a preference statement:

$$\frac{1}{2}K_\rho + \frac{1}{2}K_{\widehat{\rho}}^\xi \sim_j \frac{1}{2}K_\rho^\varepsilon + \frac{1}{2}K_{\widehat{\rho}}$$

for all $j \in I$.

Claim 4 implies that $c_{K_\rho, K^\varepsilon_\rho}(s) \propto \rho(s)$ for ε small enough. Unless $S^* = \widehat{S^*}$ and $l(s)$ and $\widehat{l}(s)$ disagree only by a linear transformation, $\rho(s) > \widehat{\rho}(s)$ on some $s \in S^*$. Because all tastes in S^* are relevant in some positive measure event $j \in I$, there must be $j \in I$, such that

$$\begin{aligned} V\left(g_j^{\left(\frac{1}{2}K_\rho+\frac{1}{2}K^\xi_{\widehat{\rho}}\right)\cup\left(\frac{1}{2}K^\varepsilon_\rho+\frac{1}{2}K_{\widehat{\rho}}\right)}\right) &= V\left(g_j^{\frac{1}{2}K_\rho+\frac{1}{2}K_{\widehat{\rho}}}\right) - \frac{1}{2}\phi(i)\sum_{s\in S^*}\min\left(c_{K_\rho,K^\varepsilon_\rho}(s), c_{K_{\widehat{\rho}},K^\xi_{\widehat{\rho}}}(s)\right)\mu(s|i) \\ &> V\left(g_j^{\frac{1}{2}K_\rho+\frac{1}{2}K_{\widehat{\rho}}}\right) - \frac{1}{2}\phi(i)\sum_{s\in S^*} c_{K_\rho,K^\varepsilon_\rho}(s)\mu(s|i) = V\left(g_j^{\frac{1}{2}K^\varepsilon_\rho+\frac{1}{2}K_{\widehat{\rho}}}\right) \end{aligned}$$

or

$$\left(\frac{1}{2}K_\rho + \frac{1}{2}K^\xi_{\widehat{\rho}}\right) \cup \left(\frac{1}{2}K^\varepsilon_\rho + \frac{1}{2}K_{\widehat{\rho}}\right) \succ_j \frac{1}{2}K^\varepsilon_\rho + \frac{1}{2}K_{\widehat{\rho}}.$$

Thus, by Axiom 6, there is $k \in I$, such that

$$\frac{1}{2}K_\rho + \frac{1}{2}K^\xi_{\widehat{\rho}} \not\sim_k \frac{1}{2}K^\varepsilon_\rho + \frac{1}{2}K_{\widehat{\rho}},$$

a contradiction. Hence $S^* = \widehat{S^*}$ and $\widehat{l}(s) = l(s)$ up to linear transformations on S^*.

Next we have to show that the measure $\mu(.|i)$ is unique for all $i \in I$ with $\phi(i) > 0$. This follows immediately from the result in DLR (their Theorem 1), that $\widehat{\mu}(s|i)\widehat{l}(s) \propto \mu(s|i)l(s)$ for the case of a finite taste space. Alternatively, it can be shown directly. To do so, fix i. Suppose $s \in S_X$. For a parameter $\chi \in (0,1]$ and $y \in Z\backslash X$, let $p^y_x := \frac{\chi}{s_y}$. For $y \in X$, let $p^y_x := 1$. Consider χ, such that for all $y \in Z\backslash X$, $1 - p^y_x \geq 0$. Then $p^y_x s_y + (1-p^y_x)s_x = \chi$ is constant on $Z\backslash X \times X$ for all χ. Let

$$P := \{\langle p^y_x, y; (1-p^y_x), x\rangle \,|x \in X, y \in Z\}.^{53}$$

For $y \in Z\backslash X$ and $x \in X$, the one-sided derivative $\left.\frac{\partial V\left(g_i^P\right)}{\partial p^y_x}\frac{dp^y_x}{d\chi}\right|_-$ is defined everywhere, as $V\left(g_i^P\right)$ is convex[54] and continuous in p^y_x. Write it as $\left.\frac{\partial V\left(g_i^P\right)(\{p^y_x\})}{d\chi}\right|_-$, implying that, for the purpose of taking the derivative, $V\left(g_i^P\right)$ is considered a function only of p^y_x. Analogously write $\left.\frac{\partial V\left(g_i^P\right)\left(\{p^y_x\}^{y\in\widetilde{Y}}_{x\in\widetilde{X}}\right)}{d\chi}\right|_-$, implying that for the derivative, $V\left(g_i^P\right)$ is considered a function of all p^y_x with $y \in \widetilde{Y} \subseteq Z$ and $x \in \widetilde{X} \subseteq X$.[55]

Suppose $y \in Z\backslash X$ and $x \in X$. Note that, for $x \in X$, $\langle 1, x\rangle \in P$. Therefore $\langle p^y_x, y; (1-p^y_x), x\rangle \in \arg\max_{\alpha\in P} U_{s'}(\alpha)$ implies $s'_y > s'_x$. Hence, under taste s', lowering χ makes all lotteries in $\left\{\langle p^y_x, y; (1-p^y_x), x\rangle \,\middle|\, x \in \widetilde{X}, y \in \widetilde{Y}\right\}$ less desirable. All lotteries in $P\backslash\left\{\langle p^y_x, y; (1-p^y_x), x\rangle \,\middle|\, x \in \widetilde{X}, y \in \widetilde{Y}\right\}$ remain unchanged. Under which tastes, then, will lowering of χ matter? Under all those tastes $s' \in S^*$, for which all most preferred lotteries become worse, or all $s' \in S^*$ with $\arg\max_{\alpha\in P}(\alpha \cdot s') \subseteq$

---

[53] Note that p^y_x and P are defined with respect to s. For notational simplicity we do not index them accordingly.

[54] $V\left(h_i^P\right)$ is linearly increasing in q^y_x in all the states, where $\langle p^y_x, y; (1-p^y_x), x\rangle$ is the most preferred lottery in P. $\langle 1, x\rangle$ is also available in P. Therefore, the subset of S where this is the case can only gain elements, when p^y_x is increased. This implies convexity.

[55] We need to use this awkward notation, because we are changing the relevant values p^y_x simultaneously.

$\left\{ \langle p_x^y, y; (1-p_x^y), x\rangle \middle| x \in \widetilde{X}, y \in \widetilde{Y} \right\}$. So

$$\left.\frac{\partial V\left(g_i^P\right)\left(\{p_x^y\}_{x\in\widetilde{X}}^{y\in\widetilde{Y}}\right)}{d\chi}\right|_{-} \propto \phi(i)\left(\sum_{\left\{s' \middle| \arg\max_{\alpha\in P} U_{s'}(\alpha) \subseteq \left\{ \langle p_x^y, y; (1-p_x^y), x\rangle \middle| x\in\widetilde{X}, y\in\widetilde{Y}\right\}\right\}} l(s')\,\mu(s'|i)\right).$$

We are interested in isolating the value $\phi(i)\, l(s)\, \mu(s|i)$ for the taste $s \in S_X$, which generated the menu P. Accounting yields that

$$\phi(i)\, l(s)\, \mu(s|i) \propto \left.\frac{\partial V\left(g_i^P\right)\left(\{p_x^y\}_{x\in X}^{y\in Z\setminus X}\right)}{d\chi}\right|_{-} - \sum_{\{z\}\subseteq Z} \left.\frac{\partial V\left(g_i^P\right)\left(\{p_x^y\}_{x\in X\setminus\{z\}}^{y\in(Z\setminus X)\setminus\{z\}}\right)}{d\chi}\right|_{-}$$

$$+\sum_{\{z,z'\}\subseteq Z} \left.\frac{\partial V\left(g_i^P\right)\left(\{p_x^y\}_{x\in X\setminus\{z,z'\}}^{y\in(Z\setminus X)\setminus\{z,z'\}}\right)}{d\chi}\right|_{-} + ... + (-1)^{Z-2} \sum_{x\in X, y\in Z\setminus X} \left.\frac{\partial V\left(g_i^P\right)\left(\{p_x^y\}\right)}{d\chi}\right|_{-}.$$

Given $i \in I$, the value of the expected utility representation across menus is unique up to positive affine transformations. $l(s)$ is unique up to linear transformations. The specific probability $\phi(i)$ is given. Then, the requirement that $\mu(s|i)$ must be a probability measure uniquely identifies it for $\phi(i) > 0$.

This establishes that Axiom 6 is sufficient for the uniqueness statement in Theorem 2.

It remains to establish that Axiom 6 is also *necessary*. Suppose to the contrary that the representation holds with the stated uniqueness, but Axiom 6 is violated. Then, there are two menus $A, B \in \mathcal{A}$, such that $A \sim_j B$ for all $j \in I$ and $A \cup B \succ_i B$ for some $i \in I$. $A \sim_j B$ for all $j \in I$ implies $\sum_{S^*} c_{A,B}(s)\,\mu(s|j) = 0$ for all $j \in I$. $A \cup B \succ_i B$ implies that $c_{A,B}(s)$ cannot be zero under all tastes, so it must be positive under some tastes and negative under others. For us it is important that it is not constant across tastes: define $\widehat{\mu}(s|i) := \frac{(1+\eta c_{A,B}(s))\mu(s|i)}{\sum_{S^*}(1+\eta c_{A,B}(s))\mu(s|i)}$, where η is small enough, such that $1+\eta c_{A,B}(s) > 0$ for all $s \in S^*$. Accordingly define $\widehat{l}(s) := \frac{l(s)}{1+\eta c_{A,B}(s)}$. Clearly the function

$$\widehat{V}(g) := \sum_I \phi(i) \sum_{S^*} \widehat{l}(s) \max_{\alpha\in g(i)} (\alpha \cdot s)\, \widehat{\mu}(s|i)$$

is a representation of $\succ_i$, when evaluated in acts g_i^A. As such, it is unique up to positive affine transformations. To verify that it represents $\succ$, it is, therefore, sufficient to find two menus, $A \succ_j B$ for all $j \in I$, for which the relative cost of getting g_j^B instead of g_j^A across I is the same according to $\widehat{V}(g)$ as according to $V(g)$. Consider again

$$\rho(s) := \frac{\frac{1}{l(s)}}{\sum_{S^*} \frac{1}{l(s)}}$$

Then, for $\varepsilon > 0$ small enough, it is established above that $V\left(g_i^{K_\rho}\right) - V\left(g_i^{K_\rho^\varepsilon}\right) \propto \phi(i)$. But, according to

Claim 4,

$$\widehat{V}\left(g_i^{K_\rho}\right)-\widehat{V}\left(g_i^{K_\rho^\varepsilon}\right)=\phi(i)\,r\,(1-\cos\varepsilon)\sum_{s\in S^*}\widehat{l}(s)\,\rho(s)\,\widehat{\mu}(s\,|i)\propto\frac{\phi(i)}{1+\eta\sum\limits_{S^*}c_{A,B}(s)\,\mu(s\,|i)}=\phi(i)\,.$$

This contradicts the uniqueness statement in Theorem 2. Thus, Axiom 6 must hold. ■

6.3. Proof of Proposition 2

That i) implies ii) is demonstrated above. The reverse follows from Theorem 2.

It remains to be shown that if there is another pair of menus, $A', B' \in A$, such that $A' \sim_j B'$ for all $j \in I$ and $A' \cup B' \succ_i B'$ for some $i \in I$, then they add another set of possible representations, if and only if

$$\frac{c_{A',B'}(s)}{c_{A',B'}(s')}\neq\frac{c_{A,B}(s)}{c_{A,B}(s')}$$

for some $s, s' \in S$. That this condition is sufficient for the existence of additional representations is obvious. To see that it is necessary, suppose there was a representation $\widehat{V}(g)$ with $\widehat{\mu}(s\,|i) \neq \frac{(1+\eta c_{A,B}(s))\mu(s|i)}{\sum_{S^*}(1+\eta c_{A,B}(s))\mu(s|i)}$ for all η. There must be some non-constant function $c : S \to \mathbb{R}$, such that $\widehat{\mu}(s\,|i) \equiv \frac{(1+\eta c(s))\mu(s|i)}{\sum_{S^*}(1+\eta c(s))\mu(s|i)}$ for some $\eta > 0$ and $c(s) \neq c_{A,B}(s)$. $\succ_i$ mandates that $\widehat{l}(s) \propto \frac{l(s)}{1+\eta c(s)}$. Because $\widehat{V}(g)$ represents the same preference as $V(g)$, $\sum_{S^*} c(s)\,\mu(s\,|i)$ must be constant. Hence, there is some non-constant function $\widetilde{c} : S \to \mathbb{R}$, with $\sum_{S^*} \widetilde{c}(s)\,\mu(s\,|i) = 0$ for all $i \in I$. Let $\widetilde{c}^+(s)$ and $\widetilde{c}^-(s)$ be the positive and negative part of $\widetilde{c}(s)$, respectively. Choose $A' := \frac{1}{2}K_{\rho^+} + \frac{1}{2}K_{\rho^-}^\xi$ and $B' := \frac{1}{2}K_{\rho^+}^\varepsilon + \frac{1}{2}K_{\rho^-}$, where $\rho^+(s) \propto \widetilde{c}^+(s)$ and $\rho^-(s) \propto \widetilde{c}^-(s)$. Then, $A' \sim_i B'$ for all $i \in I$, but $A' \cup B' \succ_j B'$ for some $j \in I$, because $c_{A',B'}(s)$ is not constant. Thus A' and B' violate Axiom 6. They satisfy $\frac{c_{A',B'}(s)}{c_{A',B'}(s')} \neq \frac{c_{A,B}(s)}{c_{A,B}(s')}$ by construction. ■

6.4. Proof of Proposition 1':

That $v(g) = \sum_{t=1}^{T} v_{D_t}(g(D_t))$ for $g \in \overline{G} \cap G_{\{D_t\}}$ represents $\succ$ confined to $\overline{G} \cap G_{\{D_t\}}$, is implied by Proposition 1.

If the simple act g is constant on each element of $\{D_t\}_{t=1}^{T}$, then it is also constant on each element of a finer partition $\{D'_t\}_{t=1}^{T'}$. For $\tau \subseteq \{1, ..., T'\}$, such that $D_t = \bigcup_{t\in\tau} D'_t$, the usual induction argument yields

$$\begin{aligned}&\frac{1}{\sharp\tau}\left(g^*(D_1),...,g^*(D_{t-1}),A,g^*(D_{t+1}),...,g^*(D_T)\right)+\frac{\sharp\tau-1}{\sharp\tau}g^*\\ =\;&\sum_{t\in\tau}\frac{1}{\sharp\tau}\left(g^*(D'_1),...,g^*\left(D'_{t-1}\right),A,g^*\left(D'_{t+1}\right),...,g^*(D'_{T'})\right),\end{aligned}$$

and thus $v_{D_t}(A) = \sum_{t\in\tau} v_{D'_t}(A)$. Therefore, $v(g) = \sum_{t=1}^{T} v_{D_t}(g(D_t))$ for $g \in \overline{G} \cap G_{\{D_t\}}$ represents $\succ$ constrained to all simple acts, $g \in \overline{G} \cap G^*$.

The uniqueness statement follows immediately from the uniqueness in Proposition 1. That the representation implies continuity and linearity of v and, thus, the axioms is obvious. ■

6.5. Proof of Theorem 1'

As suggested in the text, first establish the result for simple acts and then show that those are dense in the space of all acts. Once this is established, verify that Definition 7 can be employed. Corollary 2 still holds, where i is replaced with D.

Claim 5: *If $\succ$ satisfies Axioms 1-5, then there are a set of bounded positive numbers $\{l(s)\}_{s\in S}$, a collection of probability measures $\{\mu_D(s)\}_{D\in\mathcal{F}}$ and a countably additive probability measure π on $\mathcal{F}$, such that, for $g \in G_{\{D_t\}}$,*

$$V(g) = \sum_{t=1}^{T} \pi(D_t) \int_S l(s) \max_{\alpha \in g(D_t)} (\alpha \cdot s)\, d\mu_{D_t}(s)$$

represents $\succ$ on G^*. Furthermore, there is a function $v : G \to \mathbb{R}$ as in Proposition 1 that agrees with V on G^*.

Proof: Just as in the proof of Theorem 1, establish that there is an event dependent, positive scaling factor $\pi'(D)$, such that

$$v(g) = \sum_{t=1}^{T} \pi'(D_t) \int_S l(s) \max_{\alpha \in g(D_t)} (\alpha \cdot s)\, d\mu_{D_t}(s)$$

for $g \in G_{\{D_t\}}$, where v represents $\succ$. $\pi'(D) = 0$, if and only if D is trivial. $\succ_D$ is then represented by $\int_S l(s) \max_{\alpha\in A} (\alpha \cdot s)\, d\mu_D(s)$. Holding utilities fixed, it is a straight forward variation of AA's classical result, that $\succ_D$ identifies $\mu_D(s)$ uniquely. Thus, it obviously identifies $\pi'(D)\,\mu_D(s)$ up to the value $\pi'(D)$. Now consider a partition $\{D_t\}_{t=1}^{T}$ with $D \cup D' \in \{D_t\}_{t=1}^{T}$ and a finer partition $\{D_t'\}_{t=1}^{T'}$ with $D, D' \in \{D_t'\}_{t=1}^{T'}$. According to the proof of Proposition 1', $v_{D\cup D'}(A) = v_D(A) + v_{D'}(A)$. As $l(s)$ does not depend on D, the representation for the finer partition must then assign the same relative weight to any taste s, as the representation for the coarser partition:

$$\mu_{D\cup D'}(s) \propto \pi'(D)\,\mu_D(s) + \pi'(D')\,\mu_{D'}(s)$$

for all $s \in S$ and $D, D' \in \mathcal{F}$. Thus, for $\mu_{D\cup D'}(s)$ to be a probability measure, it must hold that $\pi'(D \cup D') = \pi'(D) + \pi'(D')$. Inductively establish that

$$\pi'\left(\bigcup D_t\right) = \sum \pi'(D_t)$$

for $\bigcup D_t \in \mathcal{F}$. $\mathcal{F}$ is a σ-algebra, so it includes all countable unions of its elements. Since v is unique only up to positive affine transformations, $\pi(D) \propto \pi'(D)$ can be normalized, such that $\pi(D)$ is a countably additive probability measure.

For $g \in G^*$, define

$$V(g) := \sum_{t=1}^{T} \pi(D_t) \int_S l(s) \max_{\alpha \in g(D_t)} (\alpha \cdot s)\, d\mu_{D_t}(s)$$

to establish the Claim 5. $\square$

Claim 6: *The simple acts G^* are dense in G in the topology defined on G.*

Proof: We have to argue that every neighborhood of an act $g \in G$ in the product topology contains a simple act. Let $p_i : G \to G_i$ be the natural projection from G to $G_i = \mathcal{A}$ and let $B_\varepsilon(A) \subseteq \mathcal{A}$ be an open ball of radius $\varepsilon > 0$ around $A \in \mathcal{A}$,

$$B_\varepsilon(A) := \{B \in \mathcal{A} \,|\, d_h(A, B) < \varepsilon\}.$$

It suffices to show that, for every act $g \in G$, there is a simple act in every finite intersection of sets of the form $p_i^{-1}(B_\varepsilon(g(i))) \subseteq G$.[56] Let a finite set $I' \subseteq I$ index the relevant dimensions for this intersection. Then we need to establish that there is always a simple act h with

$$\max_{i \in I'} d_h(g(i), h(i)) < \varepsilon.$$

Let $L \subset \Delta(Z)$ be a finite set of lotteries over Z, such that for all $\alpha \in \Delta(Z)$ there is $\alpha' \in L$ with $d_p(\alpha, \alpha') < \varepsilon$. This set exists, because $\Delta(Z)$ is compact. Let $\mathcal{A}'$ be the set of all subsets of L. Then $\mathcal{A}' \subset \mathcal{A}$, and for all $A \in \mathcal{A}$ there is $A' \in \mathcal{A}'$ with $d_h(A, A') < \varepsilon$ by the definition of $d_h(A, B)$. Thus, there is an act in $\bigcap_{I'} p_i^{-1}(B_\varepsilon(g(i)))$ with support only in $\mathcal{A}'$. Because I' is finite and $\mathcal{F}$ the Borel σ-algebra, there is finite partition $\{D_t\}$ of I, such that $i, j \in I'$ and $i \in D_t$ imply $j \notin D_t$. Thus, for every $g \in G$ and for all $\varepsilon > 0$, there is a simple act in $\bigcap_{I'} p_i^{-1}(B_\varepsilon(g(i)))$. $\square$

Claim 6 implies that

$$v(g) \equiv \sum_{t=1}^{T} \pi(D_t) \int_S l(s) \max_{\alpha \in g(D_t)} (\alpha \cdot s)\, d\mu_{D_t}(s)$$

on G^*, which can be guaranteed according to Claim 5, uniquely determines the continuous function $v(g)$ on all of G.

We want to use Definition 7. Hold $l(s)$ fixed. It is bounded by construction. For a simple act, $g_n \in G_{\{D_t\}}$, consider the function $\varphi_n : I \to \mathbb{R}$, defined as

$$\varphi_n(i) := \int_S l(s) \max_{\alpha \in g_n(D)} (\alpha \cdot s)\, d\mu_D(s)$$

for $i \in D \in \{D_t\}$. Then, the task is to find a sequence of simple acts, $\langle g_n \rangle \subseteq G^*$, such that φ_n converges from below to the bounded function

$$\varphi(i) := \int_S l(s) \max_{\alpha \in g(i)} (\alpha \cdot s)\, d\mu_i(s)$$

for a given act $g \in G$ and some measure $\mu_i(s)$:

[56] Open sets in the product topology are the product of open sets in the topology d_h on $\mathcal{A}$, which coincide with $\mathcal{A}$ for cofinitely many $i \in I$.

First, for $g_n \in G_{\{D_t\}}$, let $D^n(i)$ be such that $i \in D^n(i) \in \{D_t\}$. Because $g_n \in G_{\{D_t\}}$ can always be expressed by using a finer partition and because $\mathcal{F}$ is the Borel σ-algebra, it is without loss of generality to assume $\lim_{n\to\infty} D^n(i) = \{i\}$. Given $\mu_D(s)$, $l(s)$ is unique. Axiom 7 then implies that $\mu_i(s) := \lim_{g_n \to g} \mu_{D^n(i)}(s)$ is well defined. $(\alpha \cdot s)$ is continuous; thus, $g_n(i) \to g(i)$ for $g_n \to g$ holds by construction.

Second, compactness of $\Delta(Z)$ and Continuity (Axiom 2) imply that the set of acts with only singletons in their support has a worst element, $\underline{g}$. Axiom 5 then implies that $g \succcurlyeq \underline{g}$ for all $g \in G$. For a singleton $\{\alpha\}$,

$$\int_S l(s) \max_{\alpha \in \{\alpha\}} (\alpha \cdot s)\, d\mu_i(s) = \sum_{x \in Z} \left(\alpha(x) \int_S l(s)\, s_x d\mu_i(s) \right).$$

For $z = \arg\min_{x \in X} \left(\int_S l(s)\, s_x d\mu_i(s) \right)$, this expression is minimized in $\alpha = \langle 1, z \rangle$. Thus, $\underline{g}$ has support in $\{\{\langle 1, z\rangle\} \,|z \in Z\}$, which is a finite set. Hence $\underline{g}$ is simple.

With a simple act as a worst act, there must then be a sequence of simple acts, such that $g_n(i) \to g(i)$ from below. Continuity of v and Definition 7 give

$$E_\pi \left[\int_S l(s) \max_{\alpha \in g(i)} (\alpha \cdot s)\, d\mu_i(s) \right] = v(g).$$

Interpreting $\mu(s\,|i) := \mu_i(s)$ as a probability measure over the taste space S, conditional on the information $i \in I$, yields the representation in Theorem 1':

$$V(g) = E_\pi \left[\int_S l(s) \max_{\alpha \in g(i)} (\alpha \cdot s)\, d\mu(s\,|i) \right].$$

This completes the proof of the sufficiency statement in Theorem 1'. That the axioms are also necessary for the existence of the representation is straight forward to verify. ■

6.6. Proof of Theorem 2'

As in the proof of Theorem 2, we want to relate the ranking of acts to something that is proportional to $\int_S l(s)\, \rho(s)\, d\mu(s\,|D)$ for a given probability measure $\rho : S \to \mathbb{R}_+$. In the case of Theorem 2, S^* turns out to be finite. It is, therefore, possible to construct two menus, whose payoff differ only under one taste. This is more complicated now, because making a menu preferred less by a finite amount under one taste will invariably make it worse under neighboring tastes, too. Given a probability measure ρ, the proof below defines a specific sequence of probability measures, $\widetilde{\rho}(n) : S \to \mathbb{R}_+$, $\widetilde{\rho}(n) \to \rho$, and a sequence of sufficiently small numbers, $\varepsilon(n) \underset{n\to\infty}{\to} 0$. Write $\widetilde{\rho}$ and ε for ease of notation, when there is no risk of confusion. $K_{\widetilde{\rho}}^{\varepsilon}$ is the convex combination of menus as defined in Definitions 11, constructed with respect to $\widetilde{\rho}(n)$ and $\varepsilon(n)$. Claim 7 states that, for the specifications of $\widetilde{\rho}$ and ε given in the proof of the claim, the cost of having to choose from $K_{\widetilde{\rho}}^{\varepsilon}$ instead of $K_{\widetilde{\rho}}$ in event D, $\int_{S'} l(s) \left(\max_{\alpha \in K_{\widetilde{\rho}}} \{\alpha \cdot s\} - \max_{\alpha \in K_{\widetilde{\rho}}^{\varepsilon}} \{\alpha \cdot s\} \right) d\mu(s\,|D)$, is proportional to $\int_{S'} l(s)\, \rho(s)\, d\mu(s\,|D)$ in the limit of $n \to \infty$. This is what we want. To state the claim, one more definition is needed. Suppose $D \in \mathcal{F}$ induces measure $\mu(s\,|D)$ on the taste space S.

Definition 14: *For $S' \subseteq S$, let*

$$\overline{\mu}(S') := E_\phi \left[\int_{S'} d\mu(s|i) \right].$$

Claim 7: *For S' a Borel set with $\overline{\mu}(S') > 0$ and for $D \in \mathcal{F}$ with $\phi(D) > 0$,*

$$\frac{\int_{S'} l(s) \left(\max_{\alpha \in K_{\widetilde{\rho}}} \{\alpha \cdot s\} - \max_{\alpha \in K^{\varepsilon}_{\widetilde{\rho}}} \{\alpha \cdot s\} \right) d\mu(s|D)}{r \int_{N^{\varepsilon}_X(\widehat{\alpha}^X)} \left(1 - \cos\left(\varepsilon - \measuredangle\left(\widehat{\alpha}^X, \beta^X\right)\right)\right) d\beta \int_{S'} l(s)\rho(s) d\mu(s|D)} \to 1$$

for $n \to \infty$, for $s \in S_X$, for $\widehat{\alpha} = \arg\max_{\alpha} \int_{N^{\varepsilon}_X(\alpha^X)} \left(1 - \cos\left(\varepsilon - \measuredangle\left(\alpha^X, \beta^X\right)\right)\right) d\beta$ and for $\widetilde{\rho}$ and ε as constructed in the proof of this claim.

Proofing this claim is the main work. The following definition, corollary and claim are also for $\widetilde{\rho}$ and ε as constructed in the proof of Claim 7.

Definition 15 *Given a representation $V(g)$ as in Theorem 1 and for $\overline{\rho}(s)$, such that $\overline{\rho}(s)\, l(s)$ is equal to a constant function $\mu(s|D)$-almost everywhere, let*

$$\lambda(D, D') := \lim_{n\to\infty} \frac{p_{K_{\widetilde{\overline{\rho}}}, K^{\varepsilon}_{\widetilde{\overline{\rho}}}}(D)}{p_{K_{\widetilde{\overline{\rho}}}, K^{\varepsilon}_{\widetilde{\overline{\rho}}}}(D')},$$

for all $s \in S$.

If there are multiple representations as in Theorem 1, $\lambda(D, D')$ must be independent of the representation chosen, as it is defined by the relative ranking of two particular menus, $\{\langle 1, z\rangle | z \in Z\}$ and $\{\langle 1, \underline{z}\rangle\}$, as can be seen from the definition of $p_{A,B}(D)$, Definition 8.

With Definition 15 and Claim 7, a corollary and another claim can be established for the limit of $n \to \infty$:

Corollary 3:

$$\frac{p_{K_{\widetilde{\rho}}, K^{\varepsilon}_{\widetilde{\rho}}}(D)}{p_{K_{\widetilde{\rho}}, K^{\varepsilon}_{\widetilde{\rho}}}(D')} \to \frac{\int_S l(s)\rho(s) d\mu(s|D)}{\int_S l(s)\rho(s) d\mu(s|D')} \lambda(D, D')$$

for $D, D' \in \mathcal{F}$ with $\phi(D) > 0$ and $\phi(D') > 0$.

Proof:

$$\frac{p_{K_{\widetilde{\rho}}(r),K^{\varepsilon}_{\widetilde{\rho}}(r)}(D)}{p_{K_{\widetilde{\rho}}(r),K^{\varepsilon}_{\widetilde{\rho}}(r)}(D')} \rightarrow \frac{\int_S l(s)\left(\max_{\alpha\in K_{\widetilde{\rho}}(r)}\{\alpha\cdot s\} - \max_{\alpha\in K^{\varepsilon}_{\widetilde{\rho}}(r)}\{\alpha\cdot s\}\right)d\mu(s|D)}{\int_S l(s)\left(\max_{\alpha\in K_{\widetilde{\rho}}(r)}\{\alpha\cdot s\} - \max_{\alpha\in K^{\varepsilon}_{\widetilde{\rho}}(r)}\{\alpha\cdot s\}\right)d\mu(s|D')}\lambda(D,D')$$

$$\rightarrow \frac{\int_S l(s)\rho(s)\,d\mu(s|D)}{\int_S l(s)\rho(s)\,d\mu(s|D')}\lambda(D,D'),$$

where the last limit is implied by Claim 7. □

For the case, where $\rho(s)\,l(s)$ is constant, the corollary just restates Definition 15.

Claim 8: *If for any $D\in\mathcal{F}$ with $\phi(D)>0$, $\rho(s)\,l(s)$ is not equal to a constant function $\mu(s|D)$-almost everywhere, then there is $D'\in\mathcal{F}$ with $\phi(D')>0$, such that*

$$\frac{p_{K_{\widetilde{\rho}},K^{\varepsilon}_{\widetilde{\rho}}}(D)}{p_{K_{\widetilde{\rho}},K^{\varepsilon}_{\widetilde{\rho}}}(D')} \nrightarrow \lambda(D,D').$$

With Corollary 3 and Claim 8 at hand, the proof is analogous to the finite case: suppose to the contrary that there was a second representation,

$$\widehat{V}(g) = E_{\phi}\left[\int_S \widehat{l}(s)\max_{\alpha\in g(i)}(\alpha\cdot s)\,d\widehat{\mu}(s|i)\right],$$

such that $l(s)$ and $\widehat{l}(s)$ disagree by more than a linear transformation on a set of tastes S' with positive measure $\overline{\mu}(S')$.[57] Consider the probability measure

$$\widehat{\rho}(s) := \frac{\frac{1}{\widehat{l}(s)}}{\int_S \frac{1}{\widehat{l}(s')}ds'}$$

and $\widetilde{\widehat{\rho}}(s)$ constructed with respect to $\widehat{\rho}(s)$ as $\widetilde{\rho}(s)$ with respect to $\rho(s)$. According to $\widehat{V}(g)$, Definition 15 implies that

$$\frac{p_{K_{\widetilde{\widehat{\rho}}},K^{\varepsilon}_{\widetilde{\widehat{\rho}}}}(D)}{p_{K_{\widetilde{\widehat{\rho}}},K^{\varepsilon}_{\widetilde{\widehat{\rho}}}}(D')} \rightarrow \lambda(D,D')$$

for all $D,D'\in\mathcal{F}$. But $\widehat{\rho}(s)\,l(s)$ is not equal to a constant function $\mu(s|D)$-almost everywhere for all $D\in\mathcal{F}$. Claim 8 then implies that there is $D'\in\mathcal{F}$ with $\phi(D')>0$, such that, according to $V(g)$,

$$\frac{p_{K_{\widetilde{\widehat{\rho}}},K^{\varepsilon}_{\widetilde{\widehat{\rho}}}}(D)}{p_{K_{\widetilde{\widehat{\rho}}},K^{\varepsilon}_{\widetilde{\widehat{\rho}}}}(D')} \nrightarrow \lambda(D,D').$$

[57] They might agree on every set $S'\subseteq S$ with $\overline{\mu}(S')>0$, but disagree on a set with $\overline{\widehat{\mu}}(S')>0$. To rule out this case, just switch the labels in this argument.

Because$\frac{p_{K_{\widetilde{\widehat{\rho}}},K^{\varepsilon}_{\widetilde{\widehat{\rho}}}}(D)}{p_{K_{\widetilde{\widehat{\rho}}},K^{\varepsilon}_{\widetilde{\widehat{\rho}}}}(D')}$ and $\lambda(D, D')$ are determined by $\succ$, this is a contradiction. Hence, no such representation $\widehat{V}(g)$ can exist, and $l(s)$ is unique for all $s \in S$ up to linear transformations and up to changes on a set S' with $\overline{\mu}(S') = 0$.

We also have to show that the Borel measure $\mu(.\,|D)$ is unique up to measure zero changes to establish the sufficiency of our axioms. Again, this follows directly from DLR (their Theorem 1). Alternatively, from the proof for the case of finite $\mathcal{F}$, copy the part that establishes

$$\phi(D)\, l(s)\, \mu(s\,|D) \propto \left.\frac{\partial V\left(g_D^P\right)\left(\{p_x^y\}_{x\in X}^{y\in Z\setminus X}\right)}{d\chi}\right|_{-} - \sum_{\{z\}\subseteq Z} \left.\frac{\partial V\left(g_D^P\right)\left(\{p_x^y\}_{x\in X\setminus\{z\}}^{y\in(Z\setminus X)\setminus\{z\}}\right)}{d\chi}\right|_{-}$$

$$+ \sum_{\{z,z'\}\subseteq Z} \left.\frac{\partial V\left(g_D^P\right)\left(\{p_x^y\}_{x\in X_s\setminus\{z,z'\}}^{y\in(Z\setminus X)\setminus\{z,z'\}}\right)}{d\chi}\right|_{-} + ... + (-1)^{Z-2} \sum_{x\in X, y\in Z\setminus X} \left.\frac{\partial V\left(g_D^P\right)(\{p_x^y\})}{d\chi}\right|_{-}.$$

For the case where $\mathcal{F}$ is infinite, S is potentially infinite, so $\mu(s\,|D)$ may be a density. Given $D \in \mathcal{F}$, the value of the expected utility representation across menus is unique up to affine transformations. $l(s)$ is unique in the sense established above. The specific probability $\phi(D)$ is given. Then the requirement that $\mu(s\,|D)$ must be a probability measure uniquely identifies $\mu(S'\,|D)$ for all $S' \subseteq S$. With this the sufficiency of Axiom 6' for the uniqueness is established, once Claims 7 and 8 are proofed.

Proof of Claim 7: First, note that $S \subset [0,1]^k$. Partition $[0,1]^k$ the following way: index prizes in Z by $\kappa \in \{1, ..., k\}$. For every prize κ, define

$$\widetilde{S}_\kappa(n) := \left\{\frac{r}{n}\,|r \in \{0, 1, ..., n\}\right\}$$

with typical element $\widetilde{s}_\kappa^r$.

Definition 16: *For*

$$S_\kappa(n) := \{s_\kappa^r\,|r \in \{0, 1, ..., n-1\}\},$$

where $s_\kappa^r = \widetilde{s}_\kappa^r + \xi(n)$, *define the half open intervals* $Q_n(s) = Q_n(s_1, ..., s_k)$

$$Q_n(s) := \{(x_1, ..., x_k)\,|s_\kappa^r < x_\kappa \Leftrightarrow s_\kappa^r < s_\kappa \text{ for all } \kappa \in \{1, ..., k\} \text{ and } r \in \{0, ..., n-1\}\}.$$

and H_{s_κ} *as the hyperplane* $\{s_\kappa\} \times [0,1]^{k-1}$, *which is orthogonal to* κ.

The intervals have the hyperplanes as boundaries.

Definition 17: *Given* $\delta(n) > 0$, *define*

$$\Sigma(n) := \bigcup_{X\subseteq Z} \{s \in S_X\,|\exists s' \notin Q_n(s) \cap S_X \text{ with } \measuredangle(\alpha_s^x, \alpha_{s'}^x) \leq \delta(n) \text{ for some } x \in X\}$$

and

$$S_X^\delta(n) := \left\{ s \in S_X \,\middle|\, \text{For } X \subset X' \; \nexists s' \in S_{X'} \text{ with } \measuredangle\left(\alpha_s^X, \alpha_{s'}^X\right) < \delta(n) \right\}.$$

Corollary 4: *For $S' \subseteq S$, the closure of $S' \setminus \bigcup_{X \subseteq Z} S_X^\delta(n)$ is contained in $\Sigma(n)$.*

Proof: Immediate.

Lemma 2: *For any $b > 0$ and $n > 0$, it is possible to choose $\xi(n)$ and $\delta(n)$, such that $\overline{\mu}(\Sigma(n)) < b$.*

Definition 18: *Let $B_\xi(s)$ be the open ξ neighborhood of s in $[0,1]^k$. Let*

$$B_\xi(S) := \bigcup_{s \in S} B_\xi(s).$$

Proof of Lemma 2: Given a sequence $\langle \xi(n) \rangle_{n=1,\dots,\infty}$, there is a sequence $\langle \delta(n) \rangle_{n=1,\dots,\infty}$, such that, for $s \in S_X$,

$$\{ s' \in S_X \,|\, \measuredangle(\alpha_s^x, \alpha_{s'}^x) < \delta(n) \text{ for some } x \in X \} \subseteq B_{\xi(n)}(s).$$

Consider the sets $\{ s \in S \,|\, s_\kappa \in (\widetilde{s}_\kappa^r, \widetilde{s}_\kappa^r + 2\xi) \}$. For any $\vartheta > 0$ choose ξ small enough, such that

$$\overline{\mu}(\{ s \in S \,|\, s_\kappa \in (\widetilde{s}_\kappa^r, \widetilde{s}_\kappa^r + 2\xi) \}) < \vartheta.$$

To see this, decompose $\overline{\mu}(S')$ into a continuous and a discrete component,

$$\overline{\mu}(S') = \overline{\mu}^c(S') + \overline{\mu}^d(S'),$$

where μ^c and μ^d each sum to at most one. Continuity of $\overline{\mu}^c$ allows to find ξ small enough to guarantee

$$\overline{\mu}^c(\{ s \in S \,|\, s_\kappa \in (\widetilde{s}_\kappa^r, \widetilde{s}_\kappa^r + 2\xi) \}) < \frac{\vartheta}{2}.$$

$\overline{\mu}^d$ has (at most) countable support $S^d = \{ s_1^d, s_2^d, \dots \}$. The cdf of $\overline{\mu}^d$ is a convergent sequence with limit $\sum_{s \in S^d} \overline{\mu}^d(\{s\})$. Thus, there is a finite number T, such that

$$\sum_{s \in S^d} \overline{\mu}^d(\{s\}) - \sum_{s \in \{ s_1^d, \dots, s_t^d \}} \overline{\mu}^d(\{s\}) < \frac{\vartheta}{2}$$

for all $t \geq T$. Choose ξ small enough, such that $\{ s \in S \,|\, s_\kappa \in (\widetilde{s}_\kappa^r, \widetilde{s}_\kappa^r + 2\xi) \} \cap \{ s_1^d, \dots, s_T^d \} = \emptyset$. Taking the two together, $\overline{\mu}(\{ s \in S \,|\, s_\kappa \in (\widetilde{s}_\kappa^r, \widetilde{s}_\kappa^r + 2\xi) \}) < \vartheta$. Now note that

$$\Sigma(n) \subseteq \bigcup_{\kappa \in \{1,\dots,k\}} \bigcup_{r \in \{0,1,\dots,n-1\}} \{ s \in S \,|\, s_\kappa \in (\widetilde{s}_\kappa^r, \widetilde{s}_\kappa^r + 2\xi) \}.$$

Choose ϑ, such that $kn\vartheta < b$ to establish the lemma.‖

According to Lemma 2, $\langle \xi(n) \rangle_{n=1,\dots,\infty}$ and $\langle \delta(n) \rangle_{n=1,\dots,\infty}$ can be chosen, such that $\overline{\mu}(\Sigma(n)) \to 0$.

Define K_ρ^ε exactly as in the case where I is finite, except that S is now potentially not finite:

Definition 13': *Given a probability distribution over tastes, $\rho : S \to \mathbb{R}_+$, define a particular convex combination of menus:*

$$K_\rho^\varepsilon := \sum_{X \subseteq Z} \int_{s \in S_X} \left(\left(K_X \backslash N_X^\varepsilon \left(\alpha_s^X \right) \right) \cup L_X \right) \rho(s)\, ds.$$

Let K_ρ denote the case $\varepsilon = 0$.

Lastly define $\widetilde{\rho}(s)$ given $\{Q_n\}$:

Definition 19: *Define $\widetilde{\rho}(s)$ to be the probability measure, such that*

- *for all $s \in S_X^\delta(n)$ with $\overline{\mu}(Q_n(s)) > 0$,*

$$\widetilde{\rho}(s) \propto \frac{\int_{Q_n(s)} \rho(s')\, d\overline{\mu}(s')}{\overline{\mu}(Q_n(s))}$$

- *otherwise, $\widetilde{\rho}(s) = 0$.*

The menu $K_{\widetilde{\rho}}^\varepsilon(n)$ is now specified. To evaluate

$$\int_{S'} l(s) \left(\max_{\alpha \in K_{\widetilde{\rho}}} \{\alpha \cdot s\} - \max_{\alpha \in K_{\widetilde{\rho}}^\varepsilon} \{\alpha \cdot s\} \right) d\mu(s \,|D),$$

first evaluate

$$l(s') \left(\max_{\alpha \in K_X \cup L_X} \{\alpha \cdot s'\} - \max_{\alpha \in \left(K_X \backslash N_X^\varepsilon(\alpha_s^X)\right) \cup L_X} \{\alpha \cdot s'\} \right)$$

for any given taste s' and $\varepsilon > 0$. For tastes $s \in S_X \backslash S_X^\delta(n)$, $\widetilde{\rho}(s) = 0$. Therefore, consider only $s \in S_X^\delta(n)$, $s' \in S_Y(n)$. Distinguish four exhaustive cases:
i) $Y = X$: Then L_X is not relevant for the maximum of either menu and $N_X^\varepsilon\left(\alpha_s^X\right) \subset K_X$. Hence,

$$l(s') \left(\max_{\alpha \in K_X \cup L_X} \{\alpha \cdot s'\} - \max_{\alpha \in \left(K_X \backslash N_X^\varepsilon(\alpha_s^X)\right) \cup L_X} \{\alpha \cdot s'\} \right) =$$

$$rl(s') \left(1 - \cos\left(\max\left\{0, \varepsilon - \measuredangle\left(\alpha_s^X, \alpha_{s'}^X\right)\right\}\right)\right),$$

as in the finite case.
ii) $Y \nsubseteq X$: For $s \in S_X^\delta$ and $s'' \in S_{X \cup Y}$, $\measuredangle\left(\alpha_s^X, \alpha_{s''}^X\right) > \delta$ by the construction of S_X^δ. For $s' \in S_Y$, there must be $s'' \in S_{X \cup Y}$, such that $\measuredangle\left(\alpha_s^X, \alpha_{s'}^X\right) \geq \measuredangle\left(\alpha_s^X, \alpha_{s''}^X\right)$. Then $\varepsilon < \delta$ ensures

$$\max_{\alpha \in K_X \cup L_X} \{\alpha \cdot s'\} - \max_{\alpha \in \left(K_X \backslash N_X^\varepsilon(\alpha_s^X)\right) \cup L_X} \{\alpha \cdot s'\} = 0.$$

iii) $Y \subset X$ and $s' \in S_Y^\delta$: Then $\measuredangle\left(\alpha_s^Y, \alpha_{s'}^Y\right) > \delta$. Hence, Claim 2 states that for $\varepsilon < \overline{\varepsilon}(\delta)$, K_X is not relevant for the maximum. In that case,

$$\max_{\alpha\in K_X\cup L_X}\{\alpha\cdot s'\}-\max_{\alpha\in\left(K_X\setminus N_X^{\varepsilon}(\alpha_s^X)\right)\cup L_X}\{\alpha\cdot s'\}=0.$$

iv) $Y\subset X$ and $s'\in S_Y\setminus S_Y^{\delta}$:

$$c\left(\varepsilon,n,s,s'\right):=\max_{\alpha\in K_X\cup L_X}\{\alpha\cdot s'\}-\max_{\alpha\in\left(K_X\setminus N_X^{\varepsilon}(\alpha_s^X)\right)\cup L_X}\{\alpha\cdot s'\}.$$

In s' also the probability α assigns to prizes in $X\setminus Y$ may generate utility. Dropping r from the expression in case i clearly yields an upper bound on $c\left(\varepsilon,n,s,s'\right)$:

$$c\left(\varepsilon,n,s,s'\right)<l\left(s'\right)\left(1-\cos\left(\max\left\{0,\varepsilon-\measuredangle\left(\alpha_s^X,\alpha_{s'}^X\right)\right\}\right)\right).$$

Let (numerator) and (denominator) refer to the expression in Claim 7. Then the four cases imply

$$\text{(numerator)}:=\frac{1}{r}\int\limits_{S'}l\left(s\right)\left(\max_{\alpha\in K_{\widetilde{\rho}}(n)}\{\alpha\cdot s\}-\max_{\alpha\in K_{\widetilde{\rho}}^{\varepsilon}(n)}\{\alpha\cdot s\}\right)d\mu\left(s\left|D\right.\right)=$$

$$\sum_{X\subseteq Z}\int\limits_{S'\cap S_X}l\left(s\right)\left(\int\limits_{N_X^{\varepsilon}(\alpha_s^X)}\left(1-\cos\left(\varepsilon-\measuredangle\left(\alpha_s^X,\beta^X\right)\right)\right)\widetilde{\rho}\left(s_\beta^X\right)d\beta\right)d\mu\left(s\left|D\right.\right)+C(\varepsilon,n)\underset{\varepsilon\to0}{\longrightarrow}0,$$

where the limit is independent of $\widetilde{\rho}$ and

$$C(\varepsilon,n):=\frac{1}{r}\sum_{X\subseteq Z}\sum_{Y\subset X}\int\limits_{S'\cap\left(S_Y\setminus S_Y^{\delta}(n)\right)}\left(\int\limits_{N_X^{\varepsilon}(\alpha_s^X)}c\left(\varepsilon,n,s_\beta^X,s\right)\widetilde{\rho}\left(s_\beta^X\right)d\beta\right)d\mu\left(s\left|D\right.\right).$$

$$\text{(denominator)}:=\int_{N_X^{\varepsilon}(\widehat{\alpha}_s^X)}\left(1-\cos\left(\varepsilon-\measuredangle\left(\widehat{\alpha}_s^X,\beta^X\right)\right)\right)d\beta\int\limits_{S'}l\left(s\right)\rho\left(s\right)d\mu\left(s\left|D\right.\right)\underset{\varepsilon\to0}{\longrightarrow}0.$$

To establish the claim, we first show that $\lim\limits_{n\to\infty}\frac{\text{(numerator)}-C(\varepsilon,n)}{\text{(denominator)}}=1$ and second that $\lim\limits_{n\to\infty}\frac{C(\varepsilon,n)}{\text{(denominator)}}=0$.

Subclaim 7.1: $\lim\limits_{n\to\infty}\frac{\text{(numerator)}-C(\varepsilon,n)}{\text{(denominator)}}=\lim\limits_{n\to\infty}\frac{\sum\limits_{X\subseteq Z}\int\limits_{S'\cap S_X}l(s)\int\limits_{N_X^{\varepsilon}(\alpha_s^X)}\widetilde{\rho}\left(s_\beta^X\right)d\beta d\mu(s|D)}{\int\limits_{N_X^{\varepsilon}(\widehat{\alpha}^X)}d\beta\int\limits_{S'}l(s)\rho(s)d\mu(s|D)}$ for any given $\widetilde{\rho}$.

Proof: Apply l'Hospital's rule twice with respect to ε. When taking the necessary derivatives, ε appears only in linear combinations of trigonometric functions, which vanish at the integration boundary:

$$\frac{\partial}{\partial\varepsilon}\left(\text{(numerator)}-C(\varepsilon,n)\right)=\sum_{X\subseteq Z}\int\limits_{S'\cap S_X}l\left(s\right)\left(\int\limits_{N_X^{\varepsilon}(\alpha_s^X)}\sin\left(\varepsilon-\measuredangle\left(\alpha_s^X,\beta^X\right)\right)\widetilde{\rho}\left(s_\beta^X\right)d\beta\right)d\mu\left(s\left|D\right.\right)\underset{\varepsilon\to0}{\longrightarrow}0$$

$$\frac{\partial^2}{\partial\varepsilon^2}\left((\text{numerator})-C(\varepsilon,n)\right)=\sum_{X\subseteq Z}\int_{S'\cap S_X} l\left(s\right)\left(\int_{N_X^{\varepsilon}\left(\alpha_s^X\right)}\cos\left(\varepsilon-\measuredangle\left(\alpha_s^X,\beta^X\right)\right)\widetilde{\rho}\left(s_\beta^X\right)d\beta\right)d\mu\left(s\,|D\,\right)$$

$$\frac{\partial}{\partial\varepsilon}\left(\text{denominator}\right)=\int_{N_X^{\varepsilon}\left(\widehat{\alpha}^X\right)}\sin\left(\varepsilon-\measuredangle\left(\widehat{\alpha}^X,\beta^X\right)\right)d\beta\int_{S'} l\left(s\right)\rho\left(s\right)d\mu\left(s\,|D\,\right)\underset{\varepsilon\to 0}{\longrightarrow}0$$

$$\frac{\partial^2}{\partial\varepsilon^2}\left(\text{denominator}\right)=\int_{N_X^{\varepsilon}\left(\widehat{\alpha}^X\right)}\cos\left(\varepsilon-\measuredangle\left(\widehat{\alpha}^X,\beta^X\right)\right)d\beta\int_{S'} l\left(s\right)\rho\left(s\right)d\mu\left(s\,|D\,\right)$$

Note that $\frac{\partial}{\partial\varepsilon}\left(\text{denominator}\right)\big|_{\varepsilon>0}\neq 0$ and $\frac{\partial^2}{\partial\varepsilon^2}\left(\text{denominator}\right)\Big|_{\varepsilon>0}\neq 0$, hence, l'Hospital's rule can be applied.

For $\varepsilon\left(n\right)$ as specified above,

$$\lim_{n\to\infty}\frac{\sum\limits_{X\subseteq Z}\int\limits_{S'\cap S_X} l\left(s\right)\left(\int\limits_{N_X^{\varepsilon}\left(\alpha_s^X\right)}\cos\left(\varepsilon-\measuredangle\left(\alpha_s^X,\beta^X\right)\right)\widetilde{\rho}\left(s_\beta^X\right)d\beta\right)d\mu\left(s\,|D\,\right)}{\int\limits_{N_X^{\varepsilon}\left(\widehat{\alpha}^X\right)}\cos\left(\varepsilon-\measuredangle\left(\widehat{\alpha}^X,\beta^X\right)\right)d\beta\int\limits_{S'} l\left(s\right)\rho\left(s\right)d\mu\left(s\,|D\,\right)}$$

$$=\quad\lim_{n\to\infty}\frac{\sum\limits_{X\subseteq Z}\int\limits_{S'\cap S_X} l\left(s\right)\int\limits_{N_X^{\varepsilon}\left(\alpha_s^X\right)}\widetilde{\rho}\left(s_\beta^X\right)d\beta d\mu\left(s\,|D\,\right)}{\int\limits_{N_X^{\varepsilon}\left(\widehat{\alpha}^X\right)}d\beta\int\limits_{S'} l\left(s\right)\rho\left(s\right)d\mu\left(s\,|D\,\right)}.\ \|$$

Thus, for $\frac{\frac{\partial^2}{\partial\varepsilon^2}\left((\text{numerator})-C(\varepsilon,n)\right)}{\frac{\partial^2}{\partial\varepsilon^2}\left(\text{denominator}\right)}\underset{n\to\infty}{\to}1$, we have to show

$$\frac{\sum\limits_{X\subseteq Z}\int\limits_{S'\cap S_X} l\left(s\right)\int\limits_{N_X^{\varepsilon}\left(\alpha_s^X\right)}\widetilde{\rho}\left(s_\beta^X\right)d\beta d\mu\left(s\,|D\,\right)}{\int\limits_{N_X^{\varepsilon}\left(\widehat{\alpha}^X\right)}d\beta\int\limits_{S'} l\left(s\right)\rho\left(s\right)d\mu\left(s\,|D\,\right)}\underset{n\to\infty}{\to}1.$$

This must hold for any Borel set S' with $\mu\left(S'\,|D\,\right)>0$. Note that

$$\widetilde{\rho}\left(s_\beta^X\right)\propto\frac{\int\limits_{Q_n\left(s_\beta^X\right)}\rho\left(s'\right)d\overline{\mu}\left(s'\right)}{\overline{\mu}\left(Q_n\left(s_\beta^X\right)\right)}$$

is constant for $\beta\in N_X^{\varepsilon}\left(\alpha_s^X\right)$ for all $s\in S'\cap S_X$ and $s\notin\Sigma\left(n\right)$. According to Lemma 2, $\xi\left(n\right)$ and $\delta\left(n\right)$ can be chosen, such that $\overline{\mu}\left(\Sigma\left(n\right)\right)\to 0$. Hence, it is sufficient to show

$$\int_{S'} l\left(s\right)\widetilde{\rho}\left(s\right)d\mu\left(s\,|D\,\right)\to\int_{S'} l\left(s\right)\rho\left(s\right)d\mu\left(s\,|D\,\right).$$

for any Borel set S'.

Because $\widetilde{\rho}\left(s\right)$ and $\rho\left(s\right)$ are both probability measures,

$$\widetilde{\rho}\left(s\right)\propto\frac{\int\limits_{Q_n\left(s\right)}\rho\left(s'\right)d\overline{\mu}\left(s'\right)}{\overline{\mu}\left(Q_n\left(s\right)\right)}\to\rho\left(s\right)$$

implies $\widetilde{\rho}(s) \to \rho(s)$.

Because $l(s)$ in the representation of Theorem 1 is integrable over S, $\mu\left(\widehat{S}\,|D\right) = 0$ implies $\int_{\widehat{S}} l(s)\, d\mu(s\,|D) = 0$. Hence, considering $l(.)\,\mu(.\,|D)$ as a measure on S, we want to establish this convergence almost everywhere according to $l(.)\,\mu(.\,|D)$. It is sufficient to show that this holds almost everywhere according to $\overline{\mu}(.)$, because $\int_{\widehat{S}} l(s)\, d\mu(s\,|D) > 0$ implies $\mu\left(\widehat{S}\,|D\right) > 0$, which in turn implies $\overline{\mu}\left(\widehat{S}\right) > 0$, because by assumption, $\phi(D) > 0$.

Subclaim 7.2: For every $\widehat{S} \subseteq S$:

$$\int_{\widehat{S}} \frac{\int_{Q_n(s)} \rho(s')\, d\overline{\mu}(s')}{\overline{\mu}(Q_n(s))} d\overline{\mu}(s) \to \int_{\widehat{S}} \rho(s)\, d\overline{\mu}(s).$$

Lemma 3:

$$\int_{\bigcup_{\widehat{S}} Q_n(s)} f(s)\, d\overline{\mu}(s) \to \int_{\widehat{S}} f(s)\, d\overline{\mu}(s)$$

for every bounded function $f : S \to \mathbb{R}$.

Proof: It is sufficient to show the lemma for all Borel sets, or, because S is compact and metrizable by the standard metric on $\mathbb{R}^k$, for all compact sets. Let $\widehat{S}$ be a compact set. Then for $s \notin \widehat{S}$ there is N, such that $s \notin \bigcup_{\widehat{S}} Q_n(s)$ for all $n > N$. To see this, recall that $Q_n(s)$ was defined to have length $\frac{1}{n}$ in every dimension. Let $\frac{1}{N} < \min_{s' \in \widehat{S}} (d(s, s'))$, where $d(s, s')$ is the standard Euclidian distance in $\mathbb{R}^k$. This implies that $\bigcup_{\widehat{S}} Q_n(s) \searrow \widehat{S}$ and, hence, $\overline{\mu}\left(\bigcup_{\widehat{S}} Q_n(s)\right) \to \overline{\mu}\left(\widehat{S}\right)$.[58] Thus, for $f : S \to \mathbb{R}$ bounded,

$$\int_{\bigcup_{\widehat{S}} Q_n(s)} f(s)\, d\overline{\mu}(s) \to \int_{\widehat{S}} f(s)\, d\overline{\mu}(s). \;\|$$

Proof of Subclaim 7.2: Because $\rho(s)$ is a probability measure, it is clearly bounded. Thus, on the one hand, Lemma 3 implies

$$\int_{\bigcup_{\widehat{S}} Q_n(s)} \frac{\int_{Q_n(s)} \rho(s')\, d\overline{\mu}(s')}{\overline{\mu}(Q_n(s))} d\overline{\mu}(s) \to \int_{\widehat{S}} \frac{\int_{Q_n(s)} \rho(s')\, d\overline{\mu}(s')}{\overline{\mu}(Q_n(s))} d\overline{\mu}(s).$$

On the other hand, $Q_n(s)$ are disjoint, and hence

$$\begin{aligned} \int_{\bigcup_{\widehat{S}} Q_n(s)} \frac{\int_{Q_n(s)} \rho(s')\, d\overline{\mu}(s')}{\overline{\mu}(Q_n(s))} d\overline{\mu}(s) &= \sum_{Q \in \left\{Q_n(s) \middle| s \in \widehat{S}\right\}} \int_Q \frac{\int_Q \rho(s')\, d\overline{\mu}(s')}{\overline{\mu}(Q)} d\overline{\mu}(s) = \\ \sum_{Q \in \left\{Q_n(s) \middle| s \in \widehat{S}\right\}} \int_Q \rho(s')\, d\overline{\mu}(s') &= \int_{\bigcup_{\widehat{S}} Q_n(s)} \rho(s')\, d\overline{\mu}(s') \to \int_{\widehat{S}} \rho(s')\, d\overline{\mu}(s'), \end{aligned}$$

[58] See, for example, Theorem 2.1 in Billingsley (1995).

again by the lemma. ‖

So indeed,

$$\lim_{n\to\infty}\frac{(\text{numerator})-C(\varepsilon,n)}{(\text{denominator})}=1.$$

Now we only have to check that $\lim_{n\to\infty}\frac{C(\varepsilon,n)}{(\text{denominator})}=0$. According to case iv above, $C(\varepsilon,n)<\overline{C}(\varepsilon,n)$, where

$$\overline{C}(\varepsilon,n):=\frac{1}{r}\sum_{X\subseteq Z}\sum_{Y\subset X}\int\limits_{S'\cap\left(S_Y\setminus S_Y^{\delta}(n)\right)} l(s)\left(\int\limits_{N_X^{\varepsilon}(\alpha_s^X)}\left(1-\cos\left(\max\left\{0,\varepsilon-\measuredangle\left(\alpha_s^X,\beta^X\right)\right\}\right)\right)\widetilde{\rho}\left(s_\beta^X\right)d\beta\right)d\mu\left(s\,|D\right).$$

As before, according to l'Hospital's rule,

$$\lim_{n\to\infty}\frac{\overline{C}(\varepsilon,n)}{(\text{denominator})}=\lim_{n\to\infty}\frac{\frac{\partial^2}{\partial\varepsilon^2}\overline{C}(\varepsilon,n)}{\frac{\partial^2}{\partial\varepsilon^2}(\text{denominator})}.$$

$$=\lim_{n\to\infty}\frac{\sum\limits_{X\subseteq Z}\sum\limits_{Y\subset X}\int\limits_{S'\setminus S_Y^{\delta}(n)} l(s)\left(\int\limits_{N_X^{\varepsilon}(\alpha_s^X)}\cos\left(\varepsilon-\measuredangle\left(\alpha_s^X,\beta^X\right)\right)\widetilde{\rho}\left(s_\beta^X\right)d\beta\right)d\mu\left(s\,|D\right)}{\int\limits_{N_X^{\varepsilon}(\widehat{\alpha}^X)}\cos\left(\varepsilon-\measuredangle\left(\widehat{\alpha}^X,\beta^X\right)\right)d\beta\int\limits_{S'} l(s)\,\rho(s)\,d\mu\left(s\,|D\right)}.$$

By Corollary 4, $S'\setminus S_Y^{\delta}(n)\subseteq\Sigma(n)$. Thus, $\overline{\mu}(\Sigma(n))\to 0$ implies $\lim_{n\to\infty}\frac{\overline{C}(\varepsilon,n)}{(\text{denominator})}=0$.

This completes the proof of the Claim 7. □

Proof of Claim 8: Let

$$\overline{\rho}(s):=\frac{\frac{1}{l(s)}}{\int_S\frac{1}{l(s')}ds'}.$$

Then $\overline{\rho}(s)\,l(s)$ is equal to a constant function $\mu(s\,|D)$-almost everywhere. Construct $\widetilde{\overline{\rho}}(s)$ with respect to $\overline{\rho}(s)$, like $\widetilde{\rho}(s)$ is constructed with respect to $\rho(s)$. Lastly let $C_n=\frac{1}{2}K_{\widetilde{\rho}}+\frac{1}{2}K_{\widetilde{\overline{\rho}}}$ for each n. Then

$$\frac{p_{C_n,\frac{1}{2}K_{\widetilde{\rho}}^{\varepsilon}+\frac{1}{2}K_{\widetilde{\overline{\rho}}}}(D)}{p_{C_n,\frac{1}{2}K_{\widetilde{\rho}}^{\varepsilon}+\frac{1}{2}K_{\widetilde{\overline{\rho}}}}(D')}=\frac{p_{K_{\widetilde{\rho}},K_{\widetilde{\rho}}^{\varepsilon}}(D)}{p_{K_{\widetilde{\rho}},K_{\widetilde{\rho}}^{\varepsilon}}(D')}$$

holds by construction. According to Definition 15,

$$\frac{p_{K_{\widetilde{\overline{\rho}}},K_{\widetilde{\overline{\rho}}}^{\xi}}(D)}{p_{K_{\widetilde{\overline{\rho}}},K_{\widetilde{\overline{\rho}}}^{\xi}}(D')}\to\lambda(D,D')$$

for all $D'\in\mathcal{F}$. For $\varepsilon(n)$ small enough, Continuity (Axiom 2) allows to find $\xi(n)$, such that

$$\frac{p_{K_{\widetilde{\rho}}(r),K_{\widetilde{\rho}}^{\varepsilon}(r)}(D)}{p_{K_{\widetilde{\overline{\rho}}}(\overline{r}),K_{\widetilde{\overline{\rho}}}^{\xi}(\overline{r})}(D)}=1,$$

which implies

$$\frac{p_{C_n,\frac{1}{2}K^{\varepsilon}_{\widetilde{\rho}}+\frac{1}{2}K_{\widetilde{\widetilde{\rho}}}}(D)}{p_{C_n,\frac{1}{2}K_{\widetilde{\rho}}+\frac{1}{2}K^{\xi}_{\widetilde{\widetilde{\rho}}}}(D)}=1$$

for all n.

Suppose $l(s)\rho(s)$ is not constant $\mu(s|D)$-almost everywhere. This implies that there is $S'\subset S$ with

$$\int_{S'} l(s)\rho(s)\,d\mu(s|D)<\int_{S'} l(s)\overline{\rho}(s)\,d\mu(s|D)$$

and $\mu(S'|D)>0$.

According to Claim 7,

$$\frac{\int_S l(s)\left(\max_{\alpha\in C_n}\{\alpha\cdot s\}-\max_{\alpha\in\frac{1}{2}K^{\varepsilon}_{\widetilde{\rho}}+\frac{1}{2}K_{\widetilde{\widetilde{\rho}}}}\{\alpha\cdot s\}\right)d\mu(s|D)}{\int_S l(s)\left(\max_{\alpha\in C_n}\{\alpha\cdot s\}-\max_{\alpha\in\frac{1}{2}K_{\widetilde{\rho}}+\frac{1}{2}K^{\xi}_{\widetilde{\widetilde{\rho}}}}\{\alpha\cdot s\}\right)d\mu(s|D)}\to\frac{\int_{S'} l(s)\rho(s)\,d\mu(s|D)}{\int_{S'} l(s)\overline{\rho}(s)\,d\mu(s|D)}$$

and hence,

$$\frac{\int_S l(s)\left(\max_{\alpha\in C_n}\{\alpha\cdot s\}-\max_{\alpha\in\frac{1}{2}K_{\widetilde{\rho}}+\frac{1}{2}K^{\xi}_{\widetilde{\widetilde{\rho}}}}\{\alpha\cdot s\}\right)d\mu(s|D)}{\int_S l(s)\left(\max_{\alpha\in C_n}\{\alpha\cdot s\}-\max_{\alpha\in\left(\frac{1}{2}K_{\widetilde{\rho}}+\frac{1}{2}K^{\xi}_{\widetilde{\widetilde{\rho}}}\right)\cup\left(\frac{1}{2}K^{\varepsilon}_{\widetilde{\rho}}+\frac{1}{2}K_{\widetilde{\widetilde{\rho}}}\right)}\{\alpha\cdot s\}\right)d\mu(s|D)}<1$$

for all n. This is equivalent to

$$\frac{p_{C_n,\frac{1}{2}K_{\widetilde{\rho}}+\frac{1}{2}K^{\xi}_{\widetilde{\widetilde{\rho}}}}(D)}{p_{C_n,\left(\frac{1}{2}K_{\widetilde{\rho}}+\frac{1}{2}K^{\xi}_{\widetilde{\widetilde{\rho}}}\right)\cup\left(\frac{1}{2}K^{\varepsilon}_{\widetilde{\rho}}+\frac{1}{2}K_{\widetilde{\widetilde{\rho}}}\right)}(D)}\nrightarrow 1.$$

Axiom 6' then requires that there be $D'\in\mathcal{F}$, such that

$$\frac{p_{C_n,\frac{1}{2}K_{\widetilde{\rho}}+\frac{1}{2}K^{\xi}_{\widetilde{\widetilde{\rho}}}}(D')}{p_{C_n,\frac{1}{2}K^{\varepsilon}_{\widetilde{\rho}}+\frac{1}{2}K_{\widetilde{\widetilde{\rho}}}}(D')}\nrightarrow 1.$$

As noted before, this requires $\phi(D')>0$. The observations above then imply

$$\frac{p_{C_n,\frac{1}{2}K^{\varepsilon}_{\widetilde{\rho}}+\frac{1}{2}K_{\widetilde{\widetilde{\rho}}}}(D)}{p_{C_n,\frac{1}{2}K^{\varepsilon}_{\widetilde{\rho}}+\frac{1}{2}K_{\widetilde{\widetilde{\rho}}}}(D')}=\frac{p_{K_{\widetilde{\rho}},K^{\varepsilon}_{\widetilde{\rho}}}(D)}{p_{K_{\widetilde{\rho}},K^{\varepsilon}_{\widetilde{\rho}}}(D')}\nrightarrow\lambda(D,D').$$

This establishes Claim 8. □

It remains to show that Axiom 6' is also necessary. The argument requires only slight changes compared to the finite case: suppose to the contrary that the representation holds with the stated uniqueness, but Axiom 6' is violated. Then, there are sequences $\langle A_n\rangle,\langle B_n\rangle,\langle C_n\rangle\subseteq\mathcal{A}$, which converge in the Hausdorff topology, with $\frac{p_{C_n,A_n\cup B_n}(D)}{p_{C_n,B_n}(D)}\nrightarrow 1$ for some $D\in\mathcal{F}$ and $\frac{p_{C_n,A_n}(D')}{p_{C_n,B_n}(D')}\to 1$ for all $D'\in\mathcal{F}$. $\frac{p_{C_n,A_n}(D')}{p_{C_n,B_n}(D')}\to 1$ for all $D' \in \mathcal{F}$ implies that

$$\frac{\int\limits_S c_{A_n,B_n}(s)\,\mu(s\,|D')}{\int\limits_S c_{C_n,B_n}(s)\,\mu(s\,|D')} \to 0$$

for all $D' \in \mathcal{F}$. $\frac{p_{C_n,A_n\cup B_n}(D)}{p_{C_n,B_n}(D)} \nrightarrow 1$ implies

$$\frac{c_{A_n,B_n}(s)}{\int\limits_S c_{C_n,B_n}(s)\,\mu(s\,|D)} \nrightarrow 0$$

$\mu(s\,|D)$-almost everywhere. In complete analogy to the finite case, define

$$\widehat{\mu}(s\,|D) := \left(1 + \eta \frac{c_{A_n,B_n}(s)}{\int\limits_S c_{C_n,B_n}(s)\,\mu(s\,|D)}\right)\mu(s\,|D),$$

where η is small enough, such that $1 + \eta\frac{c_{A_n,B_n}(s)}{\int_S c_{C_n,B_n}(s)\mu(s|D)} > 0$ for all $s \in S$. From here the argument is identical to the one in the finite case. Thus, Axiom 6' must hold. ■

6.7. Proof of Proposition 3

Proof of Lemma 1: For any given $i \in I$, $V(g)$ and $\widehat{V}(g)$ represent the same preference, $\succ_i$. Define

$$\mu'(s\,|i) := \frac{l(s)}{\widehat{l}(s)}\mu(s\,|i).$$

Then $\widehat{\mu}(s\,|i)$ must be a probability measure with $\widehat{\mu}(s\,|i) \propto \mu'(s\,|i)$. Hence

$$\widehat{\mu}(s\,|i) = \frac{\mu'(s\,|i)}{\int\limits_S d\widehat{\mu}(s\,|i)},$$

and consequently

$$\widehat{l}(s)\,\widehat{\mu}(s\,|i) = \frac{l(s)\,\mu(s\,|i)}{\int\limits_S d\mu'(s\,|i)}.$$

At the same time $V(g)$ and $\widehat{V}(g)$ have to represent the same preference across I. Then $\phi(i)$ must be a probability measure with

$$\phi(i) \propto \pi(i)\frac{\int\limits_S l(s)\max\limits_{\alpha\in A}(s\cdot\alpha)\,d\mu(s\,|i)}{\int\limits_S \widehat{l}(s)\max\limits_{\alpha\in A}(s\cdot\alpha)\,d\widehat{\mu}(s\,|i)} = \pi(i)\int\limits_S d\mu'(s\,|i) = \pi(i)\int\limits_S \frac{l(s)}{\widehat{l}(s)}d\mu(s\,|i),$$

which establishes Lemma 1. ‖

I is assumed to have finite cardinality T. According to Lemma 1, $\widehat{l}(s)$ has to solve the system of equations

$$\phi(i) \propto \pi(i)\sum_{S^*}\frac{l(s)}{\widehat{l}(s)}\mu(s\,|i).$$

We want to establish that there is a neighborhood of π, such that all ϕ in this neighborhood allow an alternative representation, $\widehat{V}(g)$. Interpret π and ϕ as vectors in $\mathbb{R}_+^T$. Denote with $\mu(s) \in \mathbb{R}_+^T$ the vector with i-th component $\mu(s|i)$ and with $\pi \odot \mu(s) \in \mathbb{R}_+^T$ the component wise product of those vectors. The system of equations has a solution with $\widehat{l}(s) > 0$, if and only if ϕ is in the interior of the positive linear span of $\{\pi \odot \mu(s)\}_{s \in S^*}$.

Lemma 4: Under the conditions of Proposition 3, $\{\mu(s)\}_{s \in S^*}$ are linearly independent.

Proof: Suppose not. Let $n \in \{1, ..., T\}$ index the tastes in S^*. Then there must be parameters c_n for $n \in \{1, ..., T-1\}$, such that $\mu(s_T) = \sum_{n \in \{1, ..., T-1\}} c_n \mu(s_n)$. Then for some $\tau \in (0, \infty) \setminus \{1\}$, we can define $\mu'(s|i)$ to be probability measures, such that

$$\mu'(s_T) \propto \tau \mu(s_T) \text{ and } \frac{\mu'(s_n|i)}{\mu'(s_m|i)} = \frac{\mu(s_n|i)}{\mu(s_m|i)}$$

for all $n, m \in \{1, ..., T-1\}$ and all $i \in I$. Then

$$l'(s_n) := l(s_n) \frac{\mu(s_n|i)}{\mu'(s_n|i)}$$

is well defined for all $n \in \{1, T\}$, and

$$V(g) = E_\pi \left[\sum_{S^*} l'(s) \max_{\alpha \in g(i)} (\alpha \cdot s) \mu'(s|i) \right].$$

This contradicts Theorem 2. $\|$

$\pi \in \mathbb{R}_+^T$. Thus, $\{\pi \odot \mu(s)\}_{s \in S^*}$ must also be linearly independent. S^* is also assumed to have cardinality T. Therefore $\{\pi \odot \mu(s)\}_{s \in S^*}$ spans $\mathbb{R}^T$, and the positive linear span of $\{\pi \odot \mu(s)\}_{s \in S^*}$ is open in $\mathbb{R}_+^T$. π can be expressed as a linear combination, which assigns unit weight to T linearly independent vectors: $\pi = \sum_{S^*} \pi \odot \mu(s)$. Hence, π is in the interior of the positive linear span of $\{\pi \odot \mu(s)\}_{s \in S^*}$. This establishes the first part of Proposition 3: under the conditions of the proposition, there is a neighborhood of π in $\mathbb{R}^T$, such that all ϕ in this neighborhood allow an alternative representation, $\widehat{V}(g)$. Since the solution of a linear system of equations is continuous in all parameters, it is continuous in π. Thus $\frac{l(s)}{\widehat{l}(s)} \approx 1$, which implies $\frac{\mu(s|i)}{\widehat{\mu}(s|i)} \approx 1$. This establishes Proposition 3. ■

References

[1] Anscombe, Francis J. and Robert J. Aumann (1963), "A Definition of Subjective Probability." *Annals of Mathematical Statistics*, 34, 199-205.

[2] Billingsley, Patrick (1995), Probability and Measure. John Wiley and Sons, New York, 3rd edition.

[3] Dekel, Eddie, Barton L. Lipman and Aldo Rustichini (1998-a), "Recent Developments in Modeling Unforeseen Contingencies." *European Economic Review*, Vol. 42, 523-542.

[4] —— (1998-b), "Standard State-Space Models Preclude Unawareness." *Econometrica*, Vol 66, No. 1, 159-173.

[5] —— (2001), "Representing Preferences with a Unique Subjective State Space." *Econometrica*, Vol 69, No. 4, 891-934.

[6] Dekel, Eddie, Barton L. Lipman, Aldo Rustichini and Todd Sarver (2007), "Representing Preferences with a Unique Subjective State Space: Corrigendum." *Econometrica*, Vol. 75, No.2, 591-600.

[7] Epstein, Larry G. (2006), "An axiomatic model of non-Bayesian updating." *Review of Economic Studies*, Vol. 73, 413-436.

[8] Epstein, Larry G. and Kyoungwon Seo (2007), "Subjective States: a more Robust Model." mimeo.

[9] Fishburn, Peter C. (1970), *Utility Theory for Decision Making.* John Wiley and Sons, New York. Reprinted by Krieger Press, Huntington, New York, 1979.

[10] Grant, Simon and Edi Karni (2005), "Why does it Matter that Beliefs and Valuations be Correctly Represented?" *International Economic Review*, Vol. 46, No. 3, 917-934.

[11] Gul, Faruk and Wolfgang Pesendorfer (2001), "Temptation and Self-Control." *Econometrica*, Vol.69, No. 6, 1403-1435.

[12] Hart, Oliver and John Moore (1999), "Foundations of Incomplete Contracts." *Review of Economic Studies*, Vol. 66, No.1, 115-138.

[13] Karni, Edi and Philippe Mongin (2000), "On the Determination of Subjective Probability by Choice." *Management Science*, 46, 233-248.

[14] Karni, Edi (2006), "Bayesian Decision Theory and the Representation of Beliefs." mimeo.

[15] Kopylov, Igor (2007), "Temptation, Guilt, and Finite Subjective State Spaces." mimeo.

[16] Kreps, David M. (1979), "A Representation Theorem for "Preference for Flexibility."" *Econometrica*, Vol. 47, No.3, 565-577.

[17] —— (1988), Notes on the Theory of Choice. Westviev Press, Boulder, CO.

[18] —— (1992), "Static Choice and Unforeseen Contingencies," in Partha Dasgupta, Douglas Gale, Oliver Hart and Eric Maskin, editors, *Economic Analysis of Markets and Games: Essays in Honor of Frank Hahn.* MIT Press, 259-281.

[19] Maskin, Eric and Jean Tirole (1999), "Unforeseen Contingencies and Incomplete Contracts." *Review of Economic Studies*, 66, 83-114.

[20] Nehring, Klaus (1996), "Preference for Flexibility and Freedom of Choice in a Savage Framework." UC-Davis working paper.

[21] —— (1999), "Preference for Flexibility in a Savage Framework." *Econometrica*, Vol. 67, No.1, 101-119.

[22] Ozdenoren, Emre (2002), "Completing the State Space with Subjective States." *Journal of Economic Theory*, Vol. 105, No. 2, 531-539.

[23] Salant, Yuval and Ariel Rubinstein (2007), "(A, f), Choice with Frames." mimeo.

[24] Savage, Leonard J. (1954), *The Foundations of Statistics.* John Wiley and Sons, New York. Revised edition, Dover, New York, 1972.

[25] Sher, Shlomi and Craig R.M. McKenzie (2006), "Information leakage from logically equivalent frames." *Cognition, 101, 467-494.*

[26] Tirole, Jean (1999), "Incomplete Contracts: Where do we Stand?" *Econometrica*, Vol. 67, No.4, 741-781.

[27] Von Neumann, John and Oscar Morgenstern (1944), Theory of Games and Economic Behavior. Princeton University Press, Princeton, NJ.