Conditional Preference for Flexibility: Eliciting Beliefs from Behavior*

Philipp Sadowski†

November 29, 2009

Abstract

Following Kreps (1979), I consider a decision maker with uncertain beliefs about her future tastes. 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. I allow choice among menus to depend on exogenous states, interpreted as information. My axioms yield a representation that uniquely identifies beliefs, provided information is sufficiently relevant for choice. 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

The expected utility model of von Neumann and Morgenstern (1944, henceforth vNM,) explains choice under risk by considering probabilities and tastes separately. In the context of choice under subjective uncertainty, the corresponding separation of beliefs and tastes is a central concern. For the one extreme case, where all subjective uncertainty can be captured by observable states of the world, the works of Savage (1954) and Anscombe and Aumann (1963, henceforth AA,) achieve this separation. In the other extreme case, where none of the subjective uncertainty can be captured by observable states of the world, uncertainty can be

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

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

modeled with a subjective state space. Kreps (1979, henceforth Kreps) and Dekel, Lipman and Rustichini (2001, henceforth DLR,)[1] find that the separation is not possible in this case.

In the general case some subjective uncertainty can be described by observable states of the world, but potentially not all. This paper analyzes a model of choice under such general subjective uncertainty, which features the AA and DLR models as special cases.[2] The model separately identifies beliefs over subjective states and tastes, provided that observable states are "relevant enough". A tight behavioral characterization of relevant enough is given.[3]

I call observable states of the world information. An act assigns a menu of alternatives for future choice to every information.[4] Timing of choice is the following. In period 1, the decision maker (DM) chooses an act. Between periods 1 and 2 information arrives. In period 2 the act is evaluated and DM chooses from the resulting menu. Only period 1 choice is observed. If information does not account for all subjective uncertainty that resolves between periods 1 and 2, then even commitment to a contingent plan of period 2 choice is costly and one should observe conditional preference for retaining flexibility: All else being equal, DM prefers the act that assigns a larger menu to any particular information.

This paper provides a representation of Conditional Preference for Flexibility (CPF.) As in DLR, subjective uncertainty is modeled by a subjective state space, which collects all possible tastes that might govern DM's choice in period 2. I call it the taste space. DM conditions her beliefs about her future tastes on information. Holding information fixed, choice over menus is represented by a subjective expected utility as in DLR. The central new axiom, Relevance of Information, is equivalent to the unique identification of utilities and conditional beliefs in this representation.

My model allows any finite or topological information space, I, and any finite prize space. An act, g, assigns every information, i, a contingent menu, $g(i)$, of lotteries over prizes. Accordingly, the taste space, S, collects all possible vNM rankings of lotteries over prizes. In the case of finite I, choice has a CPF representation , if it can be represented by

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

[1]Throughout the paper I refer to the version of their model that represents preference for flexibility. Dekel, Lipman, Rustichini and Sarver (2007), henceforth DLRS, is a relevant corrigendum.

[2]In Savage and Kreps' models there is no objective uncertainty (or risk), while AA, DLR, and the present paper consider a combination of subjective and objective uncertainty.

[3]Ozdenoren (2002) also accomodates this general case. In the terminology of the present paper he assumes that the observable state is irrelevant.

[4]This is in analogy to the terminology in Savage. The notion of "contingent menus" appears in Epstein (2006). Nehring (1999) calls acts with menus as outcomes "opportunity acts".

where $\mu(s|i)$ is a subjective probability measure on S. $\mu(s|i)$ is interpreted as the belief that taste s occurs, conditional on information i. U_s is a realized vNM utility function that represents taste s. ϕ is a probability distribution over information.

Theorem 1 takes the CPF representation and the distribution ϕ as given.[5] It establishes that conditional beliefs $\mu(s|i)$ are unique and utilities U_s are unique in the appropriate sense, if and only if choice between acts satisfies the *Relevance of Information* axiom. The axiom is formulated in terms of DM's induced ranking of menus conditional on information, which is derived from her choice over acts. Say that two menus are *the same for DM*, if for any information she considers the union of those menus to be as good as either of the menus individually. Relevance of Information requires that if two menus are not the same for DM, then there must be information under which she prefers one over the other.

Theorem 2 states that choice over acts has a CPF representation, if and only if it satisfies the immediate extensions of the AA and DLR axioms. These axioms have no implications for the effect of information on the conditional ranking of menus. They are necessary axioms even for a more general representation, where not only beliefs but also utilities depend on the observable state of the world, i. However, the interpretation of observable states as information (and with it the separation of beliefs and objectives) relies on the fact that only beliefs are updated when the state changes. Theorem 2, therefore, implies that the interpretation of observable states as information is always possible, as it does not constrain period 1 choice.

The usual choice theoretic approach is to take the 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 1 relates beliefs, which are parameters of the representation, to period 1 choice behavior.

I propose to go further. If DM has private knowledge about the contingencies underlying the formation of her taste,[6] then the natural inductive step is to employ her beliefs about future tastes to forecast period 2 choice behavior. Doing so implies that the CPF representation 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,[7] on the other hand the model can be refuted, if its forecasts do not agree with observation.

[5] ϕ could be objective. If ϕ is subjective as suggested above, it must also be elicited from choice. This is the content of Theorem 3.

[6] Kreps (1992) points out that a subjective taste space naturally accounts for contingencies that are not just indescribable, but unforeseen, at least by the observer.

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

Being able to forecast behavior can be important in strategic situations. As an illustrative example, consider a retailer, who writes a contract with a supplier today. The demand, s, facing the retailer tomorrow will be either high (h) or low (l). Today s is unknown to both parties, tomorrow it will become the private knowledge of the retailer. The only relevant public information that becomes available tomorrow is consumer confidence, i, a general market indicator, which will also be either high (H) or low (L). Thus, a contract, g, can only condition on consumer confidence, not on demand. Clearly the most efficient contract might give the retailer some choice of supply quantities, q, contingent on consumer confidence. From the perspective of the retailer, such a contract is an act in the terminology of this paper. Routinely one might write down the following objective function for the retailer's choice of contract:

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

First, take consumer confidence, $i \in \{H, L\}$, as given. The retailer can then order any quantity in $g(i)$. If tomorrow she faces demand $s \in \{h, l\}$, she will choose the quantity q that maximizes her profits, $U_s(q)$. Today she does not know tomorrow's demand, but she can assign probabilities conditional on consumer confidence, $\mu(s\,|i)$. She values the menu $g(i)$ at its expected value, $\sum_{s \in \{h,l\}} \mu(s\,|i) \max_{q \in g(i)} (U_s(q))$. Second, she takes an expectation over different levels of consumer confidence according to a probability distribution ϕ. This is an example of a CPF representation.[8] The supplier's profit from supplying the quantity q does not directly depend on the demand the retailer faces, s. The supplier's expected profit depends on the probability of s only because, within the constraints of the contract, the retailer orders the quantity that maximizes her profits, given s. If the supplier does not understand the contingencies underlying s, then he can not assign a probability to s. Because of this asymmetry in information, the supplier's ability to rank contracts depends on learning the retailer's beliefs.

Contracting models usually have to assume that, first, parties know each other's ranking of contracts and that, second, they share common beliefs about future utility-payoffs, when writing the contract. The first assumption raises the complex game theoretic question of how parties learn each other's ranking of contracts. This question is rarely addressed and

[8] A general contract may give the retailer some choice between non-degenerate lotteries, α, over different quantities as alternatives for future choice, contingent on consumer confidence. For example the contract might commit the supplier to a certain action, like the quantity of units to pack and the type of packing material to use. This action might have probabilistic implications for the quantity of intact units received by the retailer.

is not my focus here. Instead, I am concerned with the second assumption. If two parties write a contract in the face of unforeseen or indescribable contingencies, which are relevant for one party's future utility-payoffs, then there can be efficiency gains from giving some control rights to this party, as in the example.[9] Since 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."* Beliefs that are elicited from a party's ranking of contracts give choice theoretic substance to the assumption of common beliefs.[10]

Section 2 demonstrates the generic identification of beliefs in the example above. Section 3 lays out the model and establishes Theorems 1 and 2, first for a finite information space and then for a general topological information space. Section 4 contains Theorem 3, which combines the two results. Section 5 comments in more detail on possible implications for contracting. Section 6 concludes.

2. Illustration of Identification of Beliefs

Consider the following three specifications of the CPF representation of the retailer's choice between contracts in the example above, where final outcomes are lotteries, α, over quantities.

• Irrelevant information: Suppose that the retailer's beliefs are independent of information about consumer confidence, $\mu(h\,|H) = \mu(h\,|L) = \mu(h)$. In this case her induced ranking of menus is independent of information and it is without loss of generality to consider only contracts with $g(H) = g(L)$. If g is such an unconditional contract, then

$$V(g) = \sum_{s\in\{h,l\}} \mu(s) \max_{\alpha\in g(H)} (U_s(\alpha)).$$

[9] Only states of the world are indescribable here. Actions (or lotteries over prizes) can be described ex ante. Whether such a contract is considered incomplete is a definitional question. See section 5 in Hart and Moore (1999) for a discussion.

[10] 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."*

This is an example of DLR's representation. To see that beliefs are not identified, consider a different probability distribution $\widehat{\mu}(s)$ on $S = \{h, l\}$ and rescaled utilities

$$\widehat{U}_s(x) = U_s(x) \frac{\mu(s)}{\widehat{\mu}(s)}.$$

Then

$$\sum_{s \in \{h,l\}} \mu(s) \left(\max_{\alpha \in g(H)} U_s(\alpha) \right) \equiv \sum_{s \in \{h,l\}} \widehat{\mu}(s) \left(\max_{\alpha \in g(H)} \widehat{U}_s(\alpha) \right).$$

This is the fundamental indeterminacy in the Kreps and DLR models and variations of those.

• No preference for flexibility: Suppose that $\mu(h|H) = 1$ and $\mu(h|L) = 0$. Now subjective uncertainty is perfectly captured by the observable information states and it is without loss of generality to identify h with H and l with L. There is then no preference for flexibility and one can confine attention to contracts with lotteries instead of menus as outcomes. If $g(i) = \alpha_i$ is such a fully specified contract, then

$$V(g) = \sum_{i \in \{H,L\}} \phi(i) U_i(\alpha_i).$$

This is an example of AA's representation.

• Preference for Flexibility and Relevant Information: Lastly, suppose the retailer believes that she faces high demand with higher probability, if consumer confidence is also high, $1 > \mu(h|H) > \mu(h|L) > 0$. Further suppose that there was another representation of the same ranking of contracts with beliefs $\widehat{\mu}(s|i)$ and tastes $\widehat{U}_l$ and $\widehat{U}_h$:

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

V and $\widehat{V}$ have to generate the same ranking of contracts.

Consider two quantities (or degenerate lotteries) q_h and q_l such that the retailer prefers to receive q_h if demand is high and q_l if demand is low, that is, $U_h(q_h) - U_h(q_l) > 0$ and $U_l(q_h) - U_l(q_l) < 0$. Slightly abusing notation, I denote a lottery that gives q_h with probability α and q_l with probability $1 - \alpha$ by α. I denote by $\{\alpha, \beta\}$ the menu that contains lotteries α and β.

Suppose for $\beta < \alpha$ and $\delta, \varepsilon \in (0, 1 - \alpha)$ the retailer is indifferent between the two contracts

$$\begin{aligned} g &= \begin{pmatrix} \{\alpha+\delta,\beta\} \text{ if } i=H \\ \{\alpha,\beta\} \text{ if } i=L \end{pmatrix} \\ g' &= \begin{pmatrix} \{\alpha,\beta\} \text{ if } i=H \\ \{\alpha+\varepsilon,\beta\} \text{ if } i=L \end{pmatrix}. \end{aligned}$$

$\beta < \alpha$ implies that α is relevant for the value of these contracts only under taste h. Hence, $g \sim g'$ implies that

$$\phi(H)\,\mu(h\,|H)\,\delta\,(U_h(q_h) - U_h(q_l)) = \phi(L)\,\mu(h\,|L)\,\varepsilon\,(U_h(q_h) - U_h(q_l))\,.$$

An analogous equality must hold for the parameters of $\widehat{U}$. Therefore,

$$\frac{\mu(h\,|H)}{\mu(h\,|L)} = \frac{\varepsilon\phi(L)}{\delta\phi(H)} = \frac{\widehat{\mu}(h\,|H)}{\widehat{\mu}(h\,|L)}.$$

Similarly,

$$\frac{\mu(l\,|H)}{\mu(l\,|L)} = \frac{\widehat{\mu}(l\,|H)}{\widehat{\mu}(l\,|L)}.$$

Since μ and $\widehat{\mu}$ are both probability measures, it follows immediately that $\mu \equiv \widehat{\mu}$. That the scaling of the expected utility functions U_h and U_l is unique up to a common linear transformation is implied by standard arguments applied to the comparison of contracts which disagree only under information i.

This reasoning can be generalized to any finite information space, I: whenever a representation features at least as many linearly independent probability measures over the taste space, indexed by $i \in I$, as there are relevant tastes, then beliefs are uniquely identified and the scaling of utilities is uniquely identified up to a common linear transformation. For the proof of Theorem 1, however, no particular representation is given. The theorem implies that the CPF representation of any ranking that satisfies Relevance of Information must have this feature.

3. The Model

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. Write explicitly $\alpha = \langle \alpha(x), x; \alpha(y), y; ... \rangle$, where $\alpha(x)$ is the probability α assigns to $x \in Z$ etc. When there is no risk of confusion, x also denotes the degenerate lottery $\langle 1, x \rangle$. Let $\mathcal{A}$ be the collection of all compact subsets of $\Delta(Z)$ with menus A, B, C as elements.[11]

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 topological space, then $\mathcal{F}$ is the Borel σ-algebra.

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 alternative 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 $\mathcal{A}$.

The following concepts are important throughout the paper.

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

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

To define DM's induced ranking of menus A and B conditional on an event $D \in \mathcal{F}$, consider acts that give menu A or B, respectively, in event D and some arbitrary but fixed

[11] Compactness is not essential. If menus were not compact, maximum and minimum would have to be replaced by supremum and infimum, respectively, in all that follows.

default menu, A^*, in the event not D. Comparing those acts induces a ranking $\succ_D$ over menus. In the context of the model, $\succ_D$ turns out to be independent of A^*.

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

$$g_D^A(i) := \begin{cases} A \text{ for } i \in D \\ A^* \text{ otherwise} \end{cases}.$$

Let $\succ_D$ be the induced binary relation on $\mathcal{A} \times \mathcal{A}$, $A \succ_D B$, if and only if $g_D^A \succ g_D^B$. $\not\succ_D$ 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 $A \succ_D B$.

In period 2 objects of choice are lotteries over the prize space. The taste space (the collection of all conceivable period 2 tastes) is the collection of all vNM rankings of lotteries. The following definition is due to DLRS.

Definition 3:

$$S = \left\{ s \in \mathbb{R}^k \,\middle|\, \sum_t s_t = 0 \text{ and } \sum_t s_t^2 = 1 \right\}$$

is the *taste space.*[12]

S collects all possible realized vNM utilities, twice normalized. Every taste in S is a vector with k components where each entry can be thought of as specifying the relative utility associated with the corresponding prize.[13]

Definition 4: For any topological information space I and preference relation $\succ$, call (ϕ, μ, U) a representation of *Conditional Preference for Flexibility (CPF)*, if ϕ is a probability measure on I, $\mu = \{\mu(.\,|i)\}_{i \in I}$ is a family of probability measures on S, $U = \{U_s(.)\}_{s \in S}$ is a family of vNM utilities where U_s represents taste s and the objective function

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

represents $\succ$.

[12] DLRS refer to S as the universal state space.

[13] 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.

E_ϕ denotes an appropriately defined expectation. The following two subsections consider I to be a finite and a generic topological space, respectively.

3.1. Information as a Finite Space

Assume that I is finite and let $i, j \in \mathcal{F}$ denote the elementary events of the sigma algebra $\mathcal{F}$. Then the CPF representation (ϕ, μ, U) corresponds to the objective function

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

If U_s is a vNM representation of taste s, then it must have the form $U_s(\alpha) = l(s)(s \cdot \alpha) + c(s)$, where $s \cdot \alpha$ is the vector product of state s and lottery α, $l(s)$ is the "intensity" of taste s and $c(s)$ is some constant. The relative intensity of utilities together with beliefs determines how DM trades of gains across tastes. The constants $c(s)$ have no behavioral content. In addition any changes on measure zero subsets of S have no behavioral content. This motivates the next definition.

Definition 5: For the CPF representation (ϕ, μ, U)
i) The *space of relevant information,* $I^* \subseteq I$, is the minimal set with $\phi(I^*) = 1$.
$\mu = \{\mu(.\,|i)\}_{i \in I}$ is unique, if the measure $\mu(s\,|i)$ is unique for all $i \in I^*$.
ii) The *space of relevant tastes,* $S^* \subseteq S$, is the minimal set with $\mu(S^*\,|i) = 1$ for all $i \in I^*$.[14]
$U = \{U_s(.)\}_{s \in S}$ is *essentially unique*, if U_s are unique up to a common linear transformation, the addition of constants $c(s)$ and up to changes on $S \backslash S^*$.

An axiomatization of the CPF representation is given in Theorem 2. The distribution ϕ is identified from behavior in Theorem 3. The main concern, however, is to separately identify beliefs μ and objectives U, *provided* that DM's choice over acts has a CPF representation for a given distribution ϕ.

Axiom 1 *(Relevance 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 paraphrase Axiom 1: whenever two menus are not the same for DM, then there is some information for which they are not equally good. If A and B were the same for DM, then she should be willing to choose from $A \cup B$ by simply ignoring A. This can not be the

[14] S^* can be thought of as the tastes DM considers possible.

case, if $A \cup B \nsim_i B$ for some $i \in I$. Throughout the interpretation is that, ultimately, only the chosen item matters for the value of a menu. If $A \nsim_i B$, then Axiom 1 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 1 requires that changing information can make either one or the other case more relevant, namely that there is $j \in I$ with $A \nsim_j B$. Axiom 1 is not a strong assumption in the sense that it is local; it only requires breaking indifference. Intuitively, the axiom specifies situations under which information has to be relevant at all. It does not require information to be very informative. This should make it easier for the experimenter to find a rich enough information space.

Theorem 1: *If, given a probability distribution ϕ on I, $\succ$ has the CPF representation (ϕ, μ, U), then statements i and ii below are equivalent and imply iii:*
i) $\succ$ satisfies Axiom 1,
ii) μ is unique and U is essentially unique,
iii) the cardinality of S^* is bounded above by the cardinality of I^*.

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 1, 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.

Another difficulty in the application and interpretation of models of preference for flexibility is the generically infinite subjective state space. Theorem 1 conveniently constrains the space of relevant tastes, S^*, to be finite. Axiom 1 implies this finiteness, because I 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.[15] Section 3.2 considers I to be a general topological space, lifting the constraint on the cardinality of S^*.

If the CPF representation is viewed as purely descriptive of period 1 choice, then the identification of beliefs relates parameters of the representation to period 1 choice behavior only. However, if DM has private information about the formation of her future taste, then

[15] Kopylov (2009) 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.

the natural inductive step is to employ the beliefs about future tastes elicited in Theorem 1 to forecast period 2 choice. Doing so implies that the CPF representation is interpreted as a map of the decision making process and that period 1 choice which satisfies Axiom 1 constrains period 2 choice frequencies. On the one hand this inductive step must be justified for a particular application. On the other hand its validity can be refuted, if observed behavior is not in line with the model's predictions. The assumption that allows the identification of beliefs and objectives in this work is similar to the assumption underlying the uniqueness results in AA and Savage. Here, changing information only leads to updated beliefs. There, a state-independent ordinal ranking implies a state-independent cardinal ranking of prizes.[16,17]

The ability to forecast period 2 choice frequencies is relevant in strategic situations, for example in the context of contracts. Section 5 elaborates.

Both types of exogenous uncertainty in my domain are essential for the uniqueness result: on the one hand, DLR find that menus over lotteries alone do not allow to distinguish objectives and beliefs $\mu(s)$. There has to be some channel through which to vary one, but not the other. In the CPF representation, 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 objectives and beliefs in the axiomatic setup developed by Savage (1954).[18] To establish the uniqueness result, the payoff a menu generates must be varied independently for different tastes. This is possible only because DM can be offered lotteries over prizes.

That the collection of tastes which DM considers possible, S^*, must indeed be finite is argued in the discussion of the theorem above. To see how Relevance of Information implies unique beliefs and utilities, suppose there were two CPF representations of the same preference relation, (ϕ, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$. Suppose further that for information $i \in I$ one could construct menus $K \sim_i \widehat{K}$, such that K generates constant payoff across tastes according to (ϕ, μ, U) and $\widehat{K}$ according to $\left(\phi, \widehat{\mu}, \widehat{U}\right)$. Changing information changes only DM's beliefs about her future tastes. If a menu generates the same payoff for every taste,

[16] AA's representation can be viewed as a special case of the CPF representation, where there is only one taste. Their identifying assumption is the corresponding special case: the scaling of the vNM utility indexed by this one taste is independent of information.

[17] Karni and co-authors, for example Grant and Karni [2005] and Karni [2009a and 2009b,] elaborate the point that interpreting Savage's or AA's unique subjective probabilities over objective states as DM's true beliefs may be misleading. The CPF model is not immune to this critique. However, beliefs in the former models have no direct implications for observable behavior while beliefs over tastes in the CPF model have implications for period 2 choice frequencies. Hence, the current model is more readily measured against the quality of its predictions.

[18] Following Nehring (1996), a companion paper to the one cited above, Epstein and Seo (2009) 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.

then the conditional value of the menu is independent of information. Hence, $K \sim_j \widehat{K}$ for all $j \in I$ would have to hold. At the same time, if (ϕ, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ were distinct, $\widehat{K}$ would not generate constant payoff across tastes according to (ϕ, μ, U). Therefore $K \cup \widehat{K} \succ_{j'} K$ for some $j' \in I$. Relevance of Information would then imply that there is $j \in I$ with $K \nsim_j \widehat{K}$, a contradiction. This rough intuition does not quite work, because the construction of menus that generate the same payoff for every taste is not possible in general. However, because $S^* \subset S$ is finite, one can construct pairs of menus (A, B for (ϕ, μ, U) and $\widehat{A}, \widehat{B}$ for $\left(\phi, \widehat{\mu}, \widehat{U}\right)$) for which the difference in payoffs is constant across tastes. Let K be the convex combination of menus $\frac{1}{2}A + \frac{1}{2}\widehat{B}$ and let $\widehat{K} = \frac{1}{2}\widehat{A} + \frac{1}{2}B$. Then $K \sim_j \widehat{K}$ for all $j \in I$ and by the type of argument laid out above $K \cup \widehat{K} \succ_{j'} K$ for some $j' \in I$. This is enough to contradict the Relevance of Information assumption.

If Axiom 1 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 1? Suppose there is a CPF representation of $\succ$. 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:

Definition 6:

$$c_{A,B}(s) := \max_{\alpha \in A} U_s(\alpha) - \max_{\beta \in B} U_s(\beta)$$

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

$$\sum_{S^*} c_{A,B}(s)\, \mu(s|j) = 0$$

for all $j \in I$. This suggests the following Proposition.

Proposition 1: *Suppose (ϕ, μ, U) is a CPF representation of $\succ$. 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 family of representations $\left\{\left(\phi, \widehat{\mu}, \widehat{U}\right)\right\}_\eta$ based on $\widehat{\mu}(s|i) = \frac{\left(1+\eta c_{A,B}(s)\right)\mu(s|i)}{\sum_{S^*}\left(1+\eta c_{A,B}(s)\right)\mu(s|i)}$ and $\widehat{U}_s = \frac{U_s}{1+\eta c_{A,B}(s)}$, indexed by $\eta > -\frac{1}{c_{A,B}(s)}$.*

If there is another pair of menus $A', B' \in A$ satisfying i), then there is another such family 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.

It is now time to axiomatize the CPF representation. When I use the general notation $D \in \mathcal{F}$ this indicates that a statement is also relevant for a general topological information space and the induced sigma-algebra, as discussed in Section 3.2. As mentioned above, the axioms are direct extensions of standard assumptions:

Axiom 2 *(Preference): $\succ$ is asymmetric and negatively transitive.*

Axiom 3 *(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 4 *(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 $h \in G$.

If a convex combination of menus were defined as a lottery over menus, then the motivation of Independence in my setup would be the same as in more familiar contexts. Uncertainty would resolve before DM consumes an item from one of the menus. However, following DLR and Gul and Pesendorfer (2001), I define the convex combination 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 this setup *"indifference as to when uncertainty is resolved."*[19]

Axiom 5 *(Nontriviality): There are* $g,h \in G$*, such that* $g \succ h$.

The next axiom considers DM's induced ranking of menus, $\succ_D$. As long as some subjective uncertainty is not resolved with information, $\succ_D$ should exhibit preference for flexibility. This is captured by the central axiom in Kreps, which states that larger menus are weakly better than smaller menus:

Axiom 6 *(Monotonicity):* $A \cup B \succcurlyeq_D A$ *for all* $A, B \in \mathcal{A}$ *and* $D \in \mathcal{F}$.

Corollary 1: *If* $\succ$ *satisfies Axioms 2-6, 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*[20] *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 realized vNM utility functions, the subjective state space S_D can be replaced by the taste space S for all $D \in \mathcal{F}$. Note that the taste space does not include the taste where DM is indifferent between all prizes, implicitly assuming nontriviality of the

[19] Both DLR and Gul and Pesendorfer elaborate this argument.
[20] See footnote 3 in DLRS.

ex-post preferences over prizes.[21,22]

Theorem 2: $\succ$ *satisfies Axioms 2-6, if and only if it has a CPF representation.*

Proof: See Appendix.

The proof first establishes an additively separable representation of $\succ$ confined to convex subsets of $\Delta(Z)$ via the Mixture Space Theorem. Because of order denseness, this representation pins down an additively separable representation of $\succ$ on all of $\Delta(Z)$, $V(g) = \sum_{i \in I} v_i(g(i))$. Now suppose V_i represents $\succ_i$. Because of the uniqueness implied by the Mixture Space Theorem, V_i must agree up to scaling with v_i. The scaling is absorbed by $\phi(i)$, which is then normalized to be a probability distribution. Thus, an act is evaluated by

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

Note that this is AA's representation, where my acts have menus as outcomes, while AA acts have lotteries as outcomes.[23] Indeed, Axioms 2-4 imply AA's axioms. Furthermore, Axioms 2-6 imply DLRS' axioms, according to Corollary 1. According to Theorem DLRS, $\succ_i$ can then be represented by

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

where $U_{s,i}$ are vNM utility functions.[24] Pick any $j \in I$ and define $U_s := U_{s,j}$. Rescaling $\mu_i(s)$ allows representing $\succ_i$ by

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

[21] 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 $D \in \mathcal{F}$.

[22] DLR further establish that, for the smallest taste space S_D, which allows a representation of $\succ_D$, $closure(S_D)$ is unique. This closure is independent of information, if and only if 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 $A \cup B \not\sim_{D'} A$.

[23] In terms of the contracting example, Anscombe-Aumann acts correspond to completely specified contracts.

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

for all $i \in I$. Combining the two yields the CPF representation (ϕ, μ, U) where

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

represents $\succ$. The intensity of each taste is endogenous, but it is fixed across information.

Clearly Axioms 2-6 are also necessary for the generic combination of the AA and DLRS representations,

$$\widehat{V}(g) = \sum_{i \in I} \phi(i) \widehat{V}_i(g(i)) = \sum_{i \in I} \phi(i) \left[\int_S \max_{\alpha \in g(i)} (U_{s,i}(\alpha)) \, d\mu_i(s) \right]$$

where exogenous states impact not only probabilities $\mu(s\,|i)$, but also the intensities of tastes. Theorem 2 implies that there is a CPF representation of $\succ$ whenever the more general representation $\widehat{V}$ exists. The assumption that information impacts only beliefs does, therefore, not constrain period 1 choice.

Remark: Let $\widetilde{S} = S \times \mathbb{R}_+$ collect all pairs of vNM rankings *and* intensities. Suppose

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

represents $\succ$. This representation is even more general than the representation $\widehat{V}$ above. Theorem 2 implies that there is a family of probability measures $\widetilde{\mu} = \{\widetilde{\mu}(\widetilde{s}\,|i)\}_{i \in I}$ on $\widetilde{S}$, that allows to represent $\succ$ and for which every taste, $s \in S$, corresponds to at most one state in its support. It is straight forward to verify that this $\widetilde{\mu}$ has the smallest possible support $S^* \subset \widetilde{S}$ among all measures that allow a representation of $\succ$. Thus, restricting attention to CPF representations is equivalent to considering those representations based on the subjective state space $\widetilde{S}$, which utilize only a minimal amount of subjective states in the sense of DLR. According to Theorem 1, $\widetilde{\mu}$ is unique.

3.2. Information as a Topological Space

If the information space I is finite, Axiom 1 limits the cardinality of the taste space, S. This is no longer the case when I is infinite. This sub-section generalizes the previous one and considers I to be a generic topological space. The reader may choose to proceed directly to section 4 without a loss in the continuity of ideas. Here and in the proofs, definitions and results that generalize those in the previous sub-section are distinguished by a prime on their label.

Recall that $\mathcal{F}$ is the Borel σ-algebra on I. The expectation under probability measures on $\mathcal{F}$ can only be calculated directly for simple functions.[25] For general functions it is defined as an appropriate limit:

Definition 7 *(Based on Definition 10.12 in Fishburn (1970)):* For a countably additive probability measure π on $\mathcal{F}$ and a bounded measurable function $\varphi : I \to \mathbb{R}$, let $\langle \varphi_n \rangle$ be a sequence of simple functions, $\varphi_n : I \to \mathbb{R}$, 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.

Definition 5': For the CPF representation (ϕ, μ, U)
i) $\mu = \{\mu (. \,|i)\}_{i \in I}$ is unique, if 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.
ii) $U = \{U_s (.)\}_{s \in S}$ is essentially unique, if U_s are unique up to a common linear transformation, the addition of constants $c(s)$ and changes on a set $S' \subset S$ with $E_\phi \left[\int_{S'} d\mu (s \,|i) \right] = 0$.

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 entire prize space, Z, is preferred over the worst prize.

Definition 8: Given $D \in \mathcal{F}$, let $\underline{z}$ be the worst prize, $A \succcurlyeq_D \{\underline{z}\}$ for all $A \in \mathcal{A}$.[26] For $A, B \in A$, define $p_{A,B}(D) \in (-1, 1)$, such that
i) for $A \succcurlyeq_D B$, $p = p_{A,B}(D)$ solves

$$\frac{1}{1+p} A + \frac{p}{1+p} \{\underline{z}\} \sim_D \frac{1}{1+p} B + \frac{p}{1+p} Z,$$

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

[26] If $\succ$ can be represented by a CPF representation, then this prize must exist because Z is finite and because $\succ_D$ must obviously satisfy *Monotonicity.*

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

Note that $p_{A,B}(D) \neq 0$ implies that D is nontrivial.

If two sequences of menus, $\langle A_n \rangle$ and $\langle B_n \rangle$, approach each other, then the cost of getting 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 1' *(Relevance 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 1' implies Axiom 1, where i is substituted by D. To see this, note that Axiom 1 holds trivially unless there is $D \in \mathcal{F}$, such that $A \cup B \nsim_D B$ and $A \sim_D B$. This implies $p_{C,B}(D) = p_{C,A}(D)$ and $p_{C,A \cup B}(D) \neq p_{C,B}(D)$. Define the constant sequences $A_n := A$ and $B_n := B$ and let $C_n := C \succ_D A$. Then $\frac{p_{C_n, A_n \cup B_n}(D)}{p_{C_n, B_n}(D)} \nrightarrow 1$. Thus, according to Axiom 1', 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 1 is satisfied.[27]

Theorem 1': *If, given $\phi : I \to \mathbb{R}_+$, $\succ$ has the CPF representation (ϕ, μ, U), then $\succ$ satisfies Axiom 1', if and only if μ is unique and U is essentially unique.*

Proof: See Appendix.

The discussion of Theorem 1 applies here, including the implications of a partial failure of Axiom 1 and Axiom 1', respectively.

The intuition for the proof of Theorem 1 involves identifying taste $s \in S^*$ via two menus, where one is preferred over the other under taste s, but they generate the same payoff under every other relevant taste. If S is continuous, then the complication is that making a menu preferred less by a finite amount under one taste will invariably make it worse under similar tastes,[28] too. Therefore, individual tastes can only be identified in the limit where the less preferred and the more preferred menu approach each other. In this limit the cost of having

[27] If $p_{C_n, B_n}(D) \nrightarrow 0$, then Axiom 1 trivially implies Axiom 1'. Thus, Axiom 1' is only stronger than Axiom 1 for $p_{C_n, B_n}(D) \to 0$.

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

to choose from the less preferred menu instead of the more preferred menu tends to zero. Axiom 1' allows statements about the limit of the ratio of these costs for two different pairs of menus. The main idea of the proof is the same as for Theorem 1. To show that a similar construction is possible here, menus are best described in terms of their support functions.[29]

In addition to Axioms 2-6, an axiomatization of the CPF representation requires that $\succ_D$ does not change too much for small changes in D.

Axiom 7 *(Continuity in Information): For any $B \in \mathcal{A}$, the set $\{A \,|A \succ_D B\}$ is continuous in D.*

Theorem 2': *$\succ$ satisfies Axioms 2-7, if and only if $\succ$ has a CPF representation.*

Proof: See Appendix.

Straight forward changes to the proof of Theorem 2 establish the result for $\succ$ constrained to all simple acts.[30] The simple acts are shown to be dense in G under the topology defined on G. Ensuring that Definition 7 applies completes the proof.

4. Subjective versus Objective Probabilities of Information

Theorems 1 and 1' take the distribution ϕ on I and the CPF representation (ϕ, μ, U) as given and establish that μ and U are unique in the appropriate sense, if and only if information is relevant enough. ϕ might be objective in the sense that it corresponds to observed frequencies of information, or it might be subjective.

Consider first the case where ϕ is subjective and must also be elicited from behavior. Determining ϕ uniquely is analogous to the classical problem addressed by AA. Their unique identification of probabilities of exogenous states is based on the assumption of *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. In my setup, the combination of *information independence* and Axiom 1 would rule out any preference for flexibility. Thus, the assumption of information independence has to be constrained to a proper subset $\Psi \subset \mathcal{A}$ to be useful here. Having assumed state independent rankings, AA move on to consider only 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

[29] The introduction of support functions to the analysis of choice over menus is a major contributuion of DLR.

[30] The outcome of a simple act depends only on the event D in some finite partition $\{D_t \,|t \in \{1, .., T\}\}$.

result.[31] For the CPF representation it would amount to requiring that $\int_S \max_{\alpha \in A} U_s(\alpha)\, d\mu(s\,|i)$ 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.[32] Thus, the requirement must be limited to a *particular* collection of menus.

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 $\Omega = \{\$1, \$0\}$ to consist of the prizes "1 Dollar" and "nothing". The first part of Axiom 8 then requires all menus that consist only of lotteries that might pay out \$1 to be ranked independently of information. To motivate the requirement it is sufficient to assume that the value of \$1 (versus nothing) is independent of information.

Once AA restrict attention to representations with state independent utilities, there is no arbitrariness in their model. In contrast, preference for flexibility implies $\Omega \subset Z$. Hence, there could be $\Omega' \subset Z$, for which $\succ$ also satisfies the first part of Axiom 8, 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.

Theorem 3: *$\succ$ satisfies Axioms 1-6 and Axiom 8 (Axiom 1' and Axioms 2-8 if I is a general topological space,) if and only if it has a CPF representation, (ϕ, μ, U), where the evaluation of menus in $\Psi(\Delta(\Omega^*))$ is independent of information. For this representation ϕ is unique, μ is unique and U is essentially unique and constant across S for all $x \in \Omega^*$.*

[31] Compare to the discussion of Theorem 1.

[32] 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 $\int_S \max_{\alpha \in A} U_s(\alpha)\, d\mu(s\,|i)$ is linear, it can not be independent of $i \in I$.

Proof: In the class of representations, where $\int_S \max_{\alpha \in A} U_s(\alpha)\, d\mu(s|i)$ does not depend on $i \in I$ for all $A \in \Psi(\Delta(\Omega^*))$, the uniqueness of ϕ follows in complete analogy to the corresponding result in AA. Given the unique ϕ, Theorems 1 and 1' imply uniqueness of μ and essential uniqueness of U. Because a representation where $U_s(x)$ is constant across S for all $x \in \Omega^*$ clearly exists, the unique representation must have this feature. ■

Now consider the alternative case, where frequencies of information are observable. an observer who observes frequencies ϕ might be willing to assume that DM bases her evaluation of acts on ϕ, as long as a CPF representation based on ϕ *exists*.[33,34]

Proposition 2: *Suppose $\succ$ satisfies Axiom 1 and can be represented by (π, μ, U), where π has minimal support in the sense that I^* has the cardinality of S^*, T.*[35] *Then there is a neighborhood of π in $\mathbb{R}^T$, such that for any probability measure ϕ on I^* in this neighborhood there is a representation $\left(\phi, \widehat{\mu}, \widehat{U}\right)$, where $\widehat{U}$ and $\widehat{\mu}$ are continuous in ϕ.*

Proof: See Appendix.

This result about the robustness to small misspecifications of ϕ can be relevant in applications where beliefs are used to forecast period 2 choice: if the observer and the decision maker disagree slightly in their perception of the "objective" probabilities, then Theorem 1 can be applied[36] and the unique subjective probabilities of future tastes provided by Theorem 1 are at least a good approximation of DM's true beliefs.

[33] This is not always the case. For example if (π, μ, U) represents $\succ$ an there is an event $D \in \mathcal{F}$ that is trivial according to π but not according to ϕ, then there is no CPF representation based on ϕ.

[34] Alternatively, it is easy to strengthen Axiom 8, such that the unique CPF representation in Theorem 3 is based on those frequencies: *(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 $V\left(g_D^A\right) - V\left(g_D^B\right) = \left(V\left(g_{D'}^A\right) - V\left(g_{D'}^B\right)\right) \frac{\phi(D)}{\phi(D')}$ for $A, B \in \Psi(\Delta(\Omega))$.

[35] Such a representation can always be found by pooling all information states that are not needed to identify beliefs.

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

5. Asymmetric Information and Contracts

As illustrated by the example in the introduction, my domain has a natural interpretation in terms of contracts. At the time two parties write a contract, the event space I is describable. In addition there are indescribable contingencies that are more relevant for one party than for the other. It seems natural that information about those contingencies is also asymmetric. In order to focus on this asymmetry, I assume that each party foresees those and only those contingencies that are directly relevant to its own payoffs. Contingencies that are foreseen by both parties are describable.

Consider a principal and an agent who want to write a contract. Only the principal's valuations depend on indescribable contingencies, which are unforeseen only by the agent. Let S denote the principal's taste space. Actions are observable, so there is no risk of moral hazard. An action pair specifies actions to be taken by the principal and the agent, respectively. Each action pair induces a probability distribution over outcomes, which potentially dependents on the event $i \in I$, but not on indescribable contingencies.[37]

The contract can fully condition on uncertainty about the agent's payoff, because this uncertainty resolves entirely between writing the contract and taking action. Therefore, an efficient contract assigns some control rights to the principal: it specifies 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 a contract, $g : I \to \mathcal{A}$, specifies a menu of lotteries over outcomes for every event $i \in I$. The principal chooses from $g(i)$, after information i realizes and after the uncertainty about the contingencies that determine her taste over outcomes, $s \in S$, has been resolved. From the principal's point of view, the contract is an act in the terminology of the previous sections.

To agree on an efficient contract, both parties must be able to rank all contracts. The principal's ranking of contracts has a CPF representation and satisfies Axiom 1. For $\alpha_s(A) := \arg\max_{\alpha \in A} (\alpha \cdot s)$[38] the CPF representation can be written as

$$V(g) = E_\phi \left[\int_S U_s \left(\alpha_s^* \left(g(i) \right) \right) d\mu \left(s \,| i \right) \right],$$

where μ is unique. The agent assigns 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. Further he also assess probabilities of events

[37] Contingencies that impact the effect of actions on the probabilities of outcomes are directly relevant for both parties.

[38] The arg max exists, because menus are compact. If it is not unique, ties can be broken in favor of the agent.

according to the probability distribution $\phi : I \to [0, 1]$. Lastly, the agent believes that the CPF representation reveals the principal's assessment of the uncertainty about her own future tastes. Then, conditional on learning the principal's ranking of contracts, the agent ranks contracts according to

$$W(g) = E_\phi \left[\int_S \left(\alpha_s^* \left(g(i) \right) \cdot c(i) \right) d\mu \left(s \mid i \right) \right].$$

Note that $W(g)$ depends on the conditional subjective probabilities, μ, as perceived by the principal but not on the intensities of her tastes, U. In my 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.[39] As this assumption is not my focus, I make it without doing the game theoretic complexity of the contracting problem justice. Instead I 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.[40] 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. My domain is not only well suited to describe the type of (incomplete) contracts laid out above, but for those contracts my axioms also give choice theoretic substance 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 CPF representation maps her 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.

6. 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 ratio-

[39] Alternatively contracts signed in a large homogenuous population might be observed.
[40] Section 3 in Maskin and Tirole (1999) elaborates this point.

nality. However, identifying the two conceptually distinct components through preferences has proven difficult. This paper proposes to consider an exogenous state space, interpreted as information, which is relevant enough to allow their unique identification. The interpretation is that information is chosen by Nature. I conclude by suggesting a reinterpretation.

Consider information, which can be determined by the experimenter instead of Nature. In many contexts the experimenter cannot credibly offer alternative information about states of nature. If he can, then information is typically interpreted as just another dimension of the consumption bundle. In contrast, a frame is information which seems to be irrelevant to the rational evaluation of alternatives, but which may affect choice. The experimenter can change the frame (at least in a laboratory) and the frame is not interpreted as part of the consumption bundle.

One possible interpretation of frames is suggested by Sher and McKenzie (2006). They propose that logically equivalent frames may not be informationally equivalent, but convey information about the sender's knowledge about relevant but not explicitly specified aspects of the choice situation.[41]

Let $\{\succ_f\}_{f\in I}$ be a subset of $\mathcal{A} \times \mathcal{A} \times I$. Each binary relation $\succ_f$ is a subset of $\mathcal{A} \times \mathcal{A}$ and captures choice between menus in $\mathcal{A}$ under frame $f \in I$. The adaptation of my axioms to this new domain is straight forward. A representation of Preference for Flexibility with Frames is a pair (μ, U) where μ and U are as specified in Definition 4 and

$$V_f(A) = \int_S \max_{\alpha\in A} U_s(\alpha)\, d\mu(s\,|f)$$

represents $\succ_f$. To paraphrase the identifying assumption, "Relevance of Frame", in this context: if there is preference for flexibility with respect to two menus that are indifferent under one frame, then the choice can be reframed so as to break the indifference. Frames are relevant, if and only if the parameters of the representation are unique in the sense of Theorems 1.[42]

The representation suggests interpreting DM's susceptibility to frames as Bayesian decision making. The underlying model is not specified, but the uniqueness result allows classifying 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 the model,[43]) then

[41] Of course, DM might be susceptible to frames for a multitude of other reasons, like hedonic forces, cognitive load or reference dependent preferences.

[42] On this domain there is no probability measure over frames that corresponds to ϕ in the CPF representation. This simplifies the uniqueness statement and Theorem 3 becomes irrelevant.

[43] 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

$\mu(s\,|f\,)$ should predict how often taste s governs her future choice. Whether and when it does, is an empirical question.

7. Appendix

7.1. Proof of Theorem 1

The following definition is according to DLR.

Definition 11: Call $\sigma_A : S \to \mathbb{R}$ with $\sigma_A := \max_{p\in A}(p \cdot s)$ the *support function* of A.

Support functions have the following useful properties:[44]
(i)$A \subseteq B$ if and only if $\sigma_A \leqslant \sigma_B$
(ii) $\sigma_{\lambda A+(1-\lambda)B} = \lambda\sigma_A + (1-\lambda)\,\sigma_B$ whenever $0 \leq \lambda \leq 1$
(iii) $\sigma_{A\cap B} = \sigma_A \wedge \sigma_B$ and $\sigma_{(A\cup B)} = \sigma_A \vee \sigma_B$.
Denote by A_σ the maximal menu supported by σ, $A_\sigma = \bigcup_{\sigma_A=\sigma} A$. Let $\mathcal{A}_\sigma$ be the collection of all menus that are maximal with respect to some support function. Note that $A \in \mathcal{A}_\sigma$ iff A is convex. Let $\succ_i$ simultaneously denote the induced ranking of support functions, $\sigma \succ_i \xi$ if and only if $A_\sigma \succ_i A_\xi$.

Lemma 1: For ε small enough, $\sigma_\varepsilon := \varepsilon$ is a support function.

Proof: The $k-1$ dimensional hyperplane in $\mathbb{R}^k$ that contains S is $H_S = \left\{x \in \mathbb{R}^k\,|x \cdot \mathbf{1} = 0\right\}$. The hyperplane that contains the $k-1$ dimensional simplex of lotteries, $\Delta(Z)$, is $H_{\Delta(Z)} = \left\{x \in \mathbb{R}^k\,|x \cdot \mathbf{1} = 1\right\}$. These two hyperplanes are parallel. Choose ε small enough such that the $k-1$ dimensional ball $B_\varepsilon \subset H_{\Delta(Z)}$ with radius ε around the center of the simplex is itself inside the simplex, $B_\varepsilon \subset \Delta(Z)$. Then $\sigma_{B_\varepsilon} \equiv \varepsilon$. $\square$

Proof of Theorem 1, i)⇒iii): Suppose to the contrary that S^* is infinite or finite with $\#S^* > \#I$. The definition of S^* implies that one can find $\#I+1$ Borel Sets with non-empty interior, $\{S_t\}_{t=1}^{\#I+1}$, such that for all $t \leq \#I+1$ there exists $i \in I$ with $\mu(\text{int}(S_t)\,|i\,) > 0$. Since μ can have at most countably many atoms, one can further guarantee $\mu(Cl(S_t) \cap Cl(S_{t'})\,|i\,) = 0$ for all $t, t' \leq \#I+1$ and all $i \in I$.

"expects" it to.

[44]See, for example, Rockafellar (1972).

Claim 1: Given S_t, there is ε small enough and a support function ξ_t, such that $\xi_t = \varepsilon$ on $S \backslash S_t$, $\xi_t \geq \varepsilon$ on S_t and $x_t(i) := \int\limits_S [\xi_t(s) - \varepsilon] \, d\mu(s \,| i) > 0$ for some $i \in I^*$.

Proof of Claim 1: Remember that σ_ε supports a ball, B_ε, with radius ε around the center of the simplex. The maximal menu B with $\sigma_B \leq \sigma_\varepsilon$ on $S \backslash S_t$ includes all lotteries with $p \cdot s \leq \varepsilon$ for all $s \in S \backslash S_t$. This implies $\max\limits_{p \in B} (p \cdot s) > \varepsilon$ for all s in the non-empty interior of S_t. Hence, $\sigma_B > \sigma_\varepsilon$ must hold on int(S_t). Let $\xi_t := \sigma_B$. $\|$

We can solve the following system of $\#I + 1$ independent linear equations with variables $\{\alpha_t\}_{t \in \{1, \#I+1\}}$ for any $n > 0$ and some given t' :

$$\sum_{t=1}^{\#I+1} x_t(i) \, \alpha_t = 0 \text{ for all } i \in I \text{ and } \alpha_{t'} = \vartheta,$$

where x_t is as defined in Claim 1. Choose ϑ such that $\sum |\alpha_t| = 1$.The convex combination of finitely many menus is well defined, and by property (ii) above, the convex combination of finitely many support functions is, too. Thus one can define two support functions as

$$\begin{aligned} \xi &:= \sum_{t=1}^{\#I+1} |\alpha_t| \left(\mathbf{1}_{\alpha_t > 0} \xi_t + \mathbf{1}_{\alpha_t < 0} \varepsilon \right) \\ \sigma &:= \sum_{t=1}^{\#I+1} |\alpha_t| \left(\mathbf{1}_{\alpha_t > 0} \varepsilon + \mathbf{1}_{\alpha_t < 0} \xi_t \right) \end{aligned}$$

On the one hand, $\sum\limits_{t=1}^{\#I+1} x_t(i) \, \alpha_t = 0$ for all $i \in I$ immediately implies that $A_\xi \sim_i A_\sigma$. On the other hand, $\alpha_{t'} \neq 0$ implies that $A_\xi \cup A_\sigma \succ_i A_\xi$ for some $i \in I^*$, which contradicts Axiom 1.

Proof of Theorem 1, i)$\Rightarrow$ii):

Claim 2: For any positive function f on S^* there is $\overline{\alpha} > 0$ small enough, such that for any $0 < \alpha < \overline{\alpha}$ there are support functions ξ and σ with $\xi - \sigma|_{S^*} = \alpha f$.

Proof of Claim 2: List the elements of $S^* = \{s_1, s_2, ...\}$. Consider $\{S_t\}_{t=1}^{\#S^*}$ with $s_t \in S_t$ and $s_r \notin S_t$ for $r \neq t$. Construct $\xi_t(s)$ as in Claim 1. Let $x_t := \xi_t(s_t) - \varepsilon$. Choose $\{\alpha_t\}_{t=1}^{\#S^*}$ such that $\alpha_t x_t \propto f(s_t)$ and $\sum \alpha_t = 1$. Define $\overline{\xi} := \sum \alpha_t \xi_t$ and $\sigma := \varepsilon$. Then $\overline{\xi} - \sigma\big|_{S^*} \equiv \overline{\alpha} f$ for some $\overline{\alpha} > 0$. For $\alpha < \overline{\alpha}$ let $\xi := \alpha \overline{\xi} + (1 - \alpha) \varepsilon$. Then $\xi - \sigma|_{S^*} \equiv \alpha f$. $\|$

Suppose (ϕ, μ, U) and $\left(\phi, \widetilde{\mu}, \widetilde{U}\right)$ are two CPF representations of $\succ$ with S^* and $\widetilde{S^*}$ as the corresponding relevant taste spaces. Write the vNM expected utility $U_s(p)$ as $U_s(p) = l(s)(s \cdot p)$. As in the text, $l(s)$ captures the "intensity" of taste s. Let $f(s) \propto \frac{1}{l(s)}$ on S^*. Analogously let $\widetilde{f}(s) \propto \frac{1}{\widetilde{l}(s)}$ on $\widetilde{S^*}$. Find α and $\widetilde{\alpha}$ small enough, such that there are ξ and $\widetilde{\xi}$ with $\xi - \sigma|_{S^*} = \alpha f$ and $\widetilde{\xi} - \sigma\Big|_{\widetilde{S^*}} = \widetilde{\alpha}\widetilde{f}$ and $\widetilde{\xi} \sim_i \xi$. Because $f(s) \propto \frac{1}{l(s)}$ it must be true that $\max\limits_{p \in A_\xi} U_s(p) - \max\limits_{p \in A_\sigma} U_s(p)$ is constant across S^*. Consequently, $\sum\limits_{S^*} \left(\max\limits_{p \in A_\xi} U_s(p) - \max\limits_{p \in A_\sigma} U_s(p)\right) \mu(s|j)$ must be independent of j. This independence is meaningful in terms of $\succ$. It is easy to verify that it holds, if and only if

$$\frac{\phi(i)}{\phi(i)+\phi(j)} g_i^{A_\xi} + \frac{\phi(j)}{\phi(i)+\phi(j)} g_j^{A_\sigma} \sim \frac{\phi(i)}{\phi(i)+\phi(j)} g_i^{A_\sigma} + \frac{\phi(j)}{\phi(i)+\phi(j)} g_j^{A_\xi}$$

for all $i, j \in I^*$. The same argument, based on the representation $\left(\phi, \widetilde{\mu}, \widetilde{U}\right)$, implies that $\sum\limits_{\widetilde{S^*}} \left(\max\limits_{p \in A_{\widetilde{\xi}}} \widetilde{U}_s(p) - \max\limits_{p \in A_\sigma} \widetilde{U}_s(p)\right) \widetilde{\mu}(s|j)$ is independent of j. (ϕ, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ both represent $\succ$, and therefore $\sum\limits_{S^*} \left(\max\limits_{p \in A_{\widetilde{\xi}}} U_s(p) - \max\limits_{p \in A_\sigma} U_s(p)\right) \mu(s|j)$ must be independent of j. Hence, $\widetilde{\xi} \sim_j \xi$ for all $j \in I$. At the same time, $\max\limits_{p \in A_{\widetilde{\xi}}} U_s(p) - \max\limits_{p \in A_\sigma} U_s(p)$ is not constant across S^*, because (ϕ, μ, U) and $\left(\phi, \widetilde{\mu}, \widetilde{U}\right)$ are distinct, which implies $\widetilde{\alpha}\widetilde{f}(s)$ is not identical to $\alpha f(s)$ on S^* or on $\widetilde{S^*}$. W.l.o.g. suppose they disagree on S^*. Because $\widetilde{\xi} \sim_j \xi$, there must be $s', s'' \in S^*$ with $\widetilde{\alpha}\widetilde{f}(s') > \alpha f(s')$ and $\widetilde{\alpha}\widetilde{f}(s'') < \alpha f(s'')$. Hence, $A_{\widetilde{\xi}} \cup A_\xi \succ_j A_\xi$ for all $j \in I$ with $\mu(s'|j) > 0$. This contradicts Axiom 1. Hence, $S^* = \widetilde{S^*}$ and $l(s) \propto \widetilde{l}(s)$ on S^*. This establishes the essential uniqueness of U.

That the measure $\mu(.|i)$ is unique for all $i \in I$ with $\phi(i) > 0$ then 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.

Proof of Theorem 1, ii)⇒i): It remains to establish that Axiom 1 is also necessary. Suppose to the contrary that the representation exists with the stated uniqueness, but Axiom 1 is violated. Then, there are two 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 $i \in I$. $A \sim_j B$ for all $j \in I$ implies $\sum\limits_{S^*} c_{A,B}(s)\, \mu(s|j) = 0$ 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 the proof it is important that it is not constant across tastes: define $\widehat{\mu}(s|i) := \frac{\left(1+\eta c_{A,B}(s)\right)\mu(s|i)}{\sum_{S^*}\left(1+\eta c_{A,B}(s)\right)\mu(s|i)}$, 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 $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ 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 A_ξ and A_σ from the proof of claim 2. Their construction immediately implies that $V\left(g_i^{A_\xi}\right) - V\left(g_i^{A_\sigma}\right) \propto \phi(i)$ and

$$\widehat{V}\left(g_i^{A_\xi}\right) - \widehat{V}\left(g_i^{A_\sigma}\right) \propto \frac{\phi(i)}{1+\eta\sum_{S^*} c_{A,B}(s)\,\mu(s|i)} = \phi(i).$$

This contradicts the uniqueness statement in Theorem 1 i). Thus, Axiom 1 is necessary for this uniqueness statement. ■

7.2. Proof of Proposition 1

That i) implies ii) is demonstrated in the proof of Theorem 1. The reverse follows from Theorem 1.

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 $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ with $\widehat{\mu}(s|i) \neq \frac{\left(1+\eta c_{A,B}(s)\right)\mu(s|i)}{\sum_{S^*}\left(1+\eta c_{A,B}(s)\right)\mu(s|i)}$ 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 $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ represents the same preference as (ϕ, μ, U), $\sum_{S^*} c(s)\,\mu(s|i)$ 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. Following the proof of Claim 1 above, choose ξ^+ such that $\xi^+ - \sigma\big|_{S^*} = \alpha\widetilde{c}^+$ and ξ^- such that $\xi^- - \sigma\big|_{S^*} = \alpha\widetilde{c}^-$. Let $A' := A_{\xi^+}$ and $B' := A_{\xi^-}$.Then $A_{\xi^+} \sim_i A_{\xi^-}$ for all $i \in I$, but $A_{\xi^+} \cup A_{\xi^-} \succ_j A_{\xi^-}$ for some $j \in I$, because $c_{A',B'}(s)$ is not constant. Thus A' and B' violate Axiom 1. They satisfy $\frac{c_{A',B'}(s)}{c_{A',B'}(s')} \neq \frac{c_{A,B}(s)}{c_{A,B}(s')}$ by construction. ■

7.3. Proof of Theorem 2

Definition 12: 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.

Lemma 2: *$\succ$ constrained to $\overline{G}$ satisfies Axioms 2-4, 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 $v(g) = \sum\limits_{i \in I} v_i(g(i))$ 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 $v'(g) = \sum\limits_{i \in I} v_i'(g(i))$ 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. Lemma 2 is an application of the Mixture Space Theorem (Theorem 5.11 in Kreps (1988)),[45] where additive separability across I follows from the usual induction argument. $\square$

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 $\mathcal{A}$. $\square$

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 $\overline{G}$. Continuity implies that 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 Lemma 2, $V'(g) = V'(\overline{g}) = \sum\limits_{i \in I} v_i(\overline{g}(i)) = \sum\limits_{i \in I} \pi'(i) V_i(\overline{g}(i))$. According to Theorem DLRS, $V_i(\overline{g}(i)) = V_i(g(i))$. Hence, $g \succ h$ implies $\sum\limits_{i \in I} \pi'(i) V_i(g(i)) >$

[45] 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 $pg + (1-p) g'' \succ g' \succ qg + (1-q) g''$.

$\sum_{i \in I} \pi'(i) V_i(h(i))$. Therefore

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

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 2. That Axioms 2-6 are necessary for the existence of the representation is straight forward to verify. ■

7.4. Proof of Theorem 1'

The proof idea is the same as for Theorem 1. To show that Axiom 1 is sufficient for the uniqueness statement, first establish the analogous claim to Claim 2. The definition of support functions (definition 11) and all related notations remain relevant here.

Recall that $U_s(p)$ can be written as $U_s(p) = l(s)(s \cdot p)$. The function $l : S \to \mathbb{R}^+$ is strictly positive.[46] Consider the uninformative event $I \in \mathcal{F}$. Note that $\int_{S'} l d\mu(s|I)$ exists for any measurable $S' \subset S$, because the value of the menu supported by σ_ε in Lemma 1 is $\int_S \sigma_\varepsilon l d\mu(s|I) = \varepsilon \int_S l d\mu(s|I)$.

Lemma 3: There are support functions ξ and σ and a number $\alpha > 0$, such that $\mu(S'|I) - \int_{S'} \alpha(\xi - \sigma) l d\mu(s|I) < \varepsilon$. For $\alpha' > \alpha$ there are also support functions ξ' and σ', such that $\mu(S'|I) - \int_{S'} \alpha'(\xi' - \sigma') l d\mu(s) < \varepsilon$.

Proof:

Claim 3: If f is positive and integrable, then for any $\varepsilon > 0$, there is a continuous, positive function $g : S \to \mathbb{R}$ with bounded support, such that $\int_{S'} |f - g| d\mu(s) < \varepsilon$ for every measurable set $S' \subseteq S$.

[46] $l(s) = 0$ corresonds to the trivial state, which is not part of a CPF representation.

Proof: As f and μ are both weakly positive, $\int_S |f\mu(s)|\, ds$ exists. Thus, for every $\varepsilon > 0$, there exists a continuous function $g : S \to \mathbb{R}$ such that $\int_{S'} |g - f|\, d\mu(s) < \varepsilon$. See, for example, Billingsley, Theorem 17.10. Since f is positive, g can be chosen positive.$\|$

Given $\varepsilon > 0$, Claim 3 establishes that there is a continuous, positive function g , such that $\int_{S'} |l - g|\, d\mu(s) < \varepsilon$ for every measurable set $S' \subseteq S$. The function $\frac{1}{g} : S \to \mathbb{R}^+$ is then positive, bounded and continuous. Thus, for any $\varepsilon > 0$, g can be chosen such that

$$\int_{S'} |g - l| \frac{1}{g} d\mu(s) \leq \left\| \frac{1}{g} \right\|_\infty \int_{S'} |g - l|\, d\mu(s) < \frac{\varepsilon}{2}.$$

Claim 4 (Lemma 11 in DLR): The functions that are the difference of two support functions span a cone that is dense in $C(S)$, the space of continuous functions on S, the unit sphere in $\mathbb{R}^k$.

As l is positive, $\nu(S') := \int_{S'} l d\mu(s\,|I)$ is itself a measure.[47] Claim 4 then implies that for every $\varepsilon > 0$ there are two support functions ξ and σ and a number $\alpha > 0$, such that

$$\int_{S'} \left| \frac{1}{g} - \alpha(\xi - \sigma) \right| l d\mu(s) < \frac{\varepsilon}{2}$$

for every measurable set $S' \subseteq S$.

Hence,

$$\mu(S'\,|I) - \int_{S'} \alpha(\xi - \sigma)\, l d\mu(s\,|I) \leq \int_{S'} |1 - \alpha(\xi - \sigma)\, l|\, d\mu(s\,|I)$$

$$\leq \int_{S'} |g - l| \frac{1}{g} d\mu(s\,|I) + \int_{S'} \left| \frac{1}{g} - \alpha(\xi - \sigma) \right| l d\mu(s\,|I) < \varepsilon.$$

This establishes the first part of the lemma. To show the second part, consider $\alpha' = c\alpha$ with $c > 1$, then let $\sigma' = \sigma$ and $\xi' = \frac{1}{c}\xi + \left(1 - \frac{1}{c}\right)\sigma$. ξ' is a convex combination of support functions and therefore a support function and $\alpha'(\xi' - \sigma') \equiv \alpha(\xi - \sigma)$. This concludes the proof of Lemma 3. $\square$

[47] If information is ignored, in the sense that DM only gets to choose only between acts that do not condition on information, then preferences can be represented as in DLR. The measure ν corresponds to the measure featured in this representation. It is dominated by the measure $\mu(s\,|I)$ and the Radon-Nikodym derivative of v with respect to $\mu(.\,|I)$ evaluated in s is $l(s)$, the intensity of taste s.

Suppose (ϕ, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ are two CPF representations of $\succ$. Following Lemma 3, one can define a sequence of support functions $\langle \xi_n \rangle$ and $\langle \sigma_n \rangle$ and a sequence of numbers $\langle \alpha_n \rangle$, such that

$$\mu(S' \,|I) - \int_{S'} \alpha_n (\xi_n - \sigma_n)\, l d\mu(s\,|I) < \frac{1}{n}$$

for every measurable set $S' \subseteq S$ and for all $n > 0$. Analogously define $\left\langle \widehat{\xi}_n \right\rangle$ and $\langle \widehat{\sigma}_n \rangle$ and a sequence of numbers $\langle \widehat{\alpha}_n \rangle$ based on $\left(\phi, \widehat{\mu}, \widehat{U}\right)$. According to the second part of Lemma 3 it is possible to choose $\langle \alpha_n \rangle$ and $\langle \widehat{\alpha}_n \rangle$ such that

$$\int_S (\xi_n - \sigma_n)\, l d\mu(s\,|I) = \int_S \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) l d\mu(s\,|I)$$

and hence $\frac{1}{2}\xi_n + \frac{1}{2}\widehat{\sigma}_n \sim_I \frac{1}{2}\widehat{\xi}_n + \frac{1}{2}\sigma_n$ according to (ϕ, μ, U) for all $n > 0$.

Rewriting $p_{A,B}(D)$ as defined in definition 8 in terms of support functions yields $p_{A,B}(D) = \int_S (\sigma_A - \sigma_B)\, l d\mu(s\,|D)$. For the remainder of the proof, let A_n, B_n and C_n be defined, such that $\sigma_{A_n} = \frac{1}{2}\xi_n + \frac{1}{2}\widehat{\sigma}_n$, $\sigma_{B_n} = \frac{1}{2}\widehat{\xi}_n + \frac{1}{2}\sigma_n$ and $\sigma_{C_n} = \frac{1}{2}\sigma_n + \frac{1}{2}\widehat{\sigma}_n$.

Claim 5: $\frac{p_{C_n,A_n}(D)}{p_{C_n,B_n}(D)} \to 1$ for all $D \in \mathcal{F}$.

Proof: First note that

$$\begin{aligned}
\frac{p_{C_n,A_n}(D)}{p_{C_n,B_n}(D)} &= \frac{\int_S \frac{1}{2}(\sigma_n + \widehat{\sigma}_n - \xi_n - \widehat{\sigma}_n)\, l d\mu(s\,|D)}{\int_S \frac{1}{2}\left(\sigma_n + \widehat{\sigma}_n - \widehat{\xi}_n - \sigma_n\right) l d\mu(s\,|D)} \\
&= \frac{\int_S (\xi_n - \sigma_n)\, l d\mu(s\,|D)}{\int_S \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) l d\mu(s\,|D)}
\end{aligned}$$

By definition $\mu(S'\,|I) - \alpha_n \int_{S'} (\xi_n - \sigma_n)\, l d\mu(s\,|I) < \frac{1}{n}$ for every measurable set $S' \subseteq S$ and for all $n > 0$ implies that (i) $\lim\limits_{n\to\infty} \left[\alpha_n \int_S (\xi_n - \sigma_n)\, l d\mu(s\,|I)\right] = 1$, because μ is a probability measure and (ii) $\alpha_n (\xi_n - \sigma_n)\, l \to 1$ almost everywhere according to $\mu(s\,|I)$. The same observations can be made for $\left\langle \widehat{\xi}_n \right\rangle$, $\langle \widehat{\sigma}_n \rangle$, $\langle \widehat{\alpha}_n \rangle$ and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$.

For every $D \in \mathcal{F}$ the range of $\mu(.\,|D)$ is a subset of the range of μ and $S' \subseteq S$ is $\mu(.\,|D)$ measurable, if and only if it is $\mu(s\,|I)$ measurable. Hence, $\lim\limits_{n\to\infty} \left[\alpha_n \int_S (\xi_n - \sigma_n)\, l d\mu(s\,|D)\right] = 1$ for all $D \in \mathcal{F}$. Analogously $\lim\limits_{n\to\infty} \left[\widehat{\alpha}_n \int_S \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) \widehat{l} d\widehat{\mu}(s\,|D)\right] = 1$ for all $D \in \mathcal{F}$. As in the case of finite I, it is easy to verify that this independence is meaningful in terms of $\succ$. Hence, there is a sequence of numbers $\langle\beta_n\rangle$, such that $\lim_{n\to\infty}\left[\beta_n\int_S\left(\widehat{\xi}_n-\widehat{\sigma}_n\right)ld\mu\left(s\,|D\right)\right]=1$ for all $D\in\mathcal{F}$. Since $\frac{1}{2}\xi_n+\frac{1}{2}\widehat{\sigma}_n\sim_I\frac{1}{2}\widehat{\xi}_n+\frac{1}{2}\sigma_n$ for all $n>0$, it must be that $\frac{\alpha_n}{\beta_n}\to 1$. Together with observations (ii) above this implies that $\frac{\int_S(\xi_n-\sigma_n)ld\mu(s|D)}{\int_S(\widehat{\xi}_n-\widehat{\sigma}_n)ld\mu(s|D)}\to 1$ for all $D\in\mathcal{F}$. $\|$

Claim 6: If (ϕ,μ,U) and $\left(\phi,\widehat{\mu},\widehat{U}\right)$ are two CPF representations of $\succ$ that are distinct beyond the changes permitted in the uniqueness statement of Theorem 1', then $\frac{p_{C_n,A_n\cup B_n}(D)}{p_{C_n,B_n}(D)}\nrightarrow 1$.

Proof: First note that

$$\begin{aligned}\frac{p_{C_n,A_n\cup B_n}(I)}{p_{C_n,B_n}(I)} &= \frac{\int_S\frac{1}{2}\left(\sigma_n+\widehat{\sigma}_n-\max\left\{\xi_n+\widehat{\sigma}_n,\widehat{\xi}_n+\sigma_n\right\}\right)ld\mu\left(s\,|I\right)}{\int_S\frac{1}{2}\left(\sigma_n+\widehat{\sigma}_n-\widehat{\xi}_n-\sigma_n\right)ld\mu\left(s\,|I\right)}\\ &= \frac{\int_S\max\left\{\xi_n-\sigma_n,\widehat{\xi}_n-\widehat{\sigma}_n\right\}ld\mu\left(s\,|I\right)}{\int_S\left(\widehat{\xi}_n-\widehat{\sigma}_n\right)ld\mu\left(s\,|I\right)}\end{aligned}$$

It follows immediately from the uniqueness statements in Theorems 3 and 4 in DLR, that $\mu\left(s\,|D\right)$ and $\widehat{\mu}\left(s\,|D\right)$ share the same support in the sense of Definition 5 and that $l\left(s\right)\mu\left(s\,|D\right)$ is unique up to rescaling for any $D\in\mathcal{F}$. Thus, if (ϕ,μ,U) and $\left(\phi,\widehat{\mu},\widehat{U}\right)$ are distinct in the sense of the claim, then the corresponding functions l and $\widehat{l}$ have to be distinct. Consequently, there is $S'\subset S$, such that $\int_{S'}\frac{\widehat{l}}{l}d\mu\left(s\,|I\right)\neq\mu\left(S'\,|I\right)$. Thus, for n large enough, $\lim_{n\to\infty}\left[\alpha_n\int_{S'}\left(\widehat{\xi}_n-\widehat{\sigma}_n\right)ld\mu\left(s\,|D\right)\right]=c\neq\mu\left(S'\,|I\right)$. W.l.o.g. suppose that $c>1$. Then $\lim_{n\to\infty}\left[\alpha_n\int_S\max\left\{\xi_n-\sigma_n,\widehat{\xi}_n-\widehat{\sigma}_n\right\}ld\mu\left(s\,|D\right)\right]>1$, which implies

$$\frac{\int_S\max\left\{\xi_n-\sigma_n,\widehat{\xi}_n-\widehat{\sigma}_n\right\}ld\mu\left(s\,|D\right)}{\int_S\left(\widehat{\xi}_n-\widehat{\sigma}_n\right)ld\mu\left(s\,|D\right)}\nrightarrow 1.\ \|$$

The combination of Claims 5 and 6 is a direct violation of Axiom 1'. Hence, Axiom 1 implies that (ϕ,μ,U) is unique in the sense of Theorem 1'.

It remains to show that Axiom 1' 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 1' is violated. Then, there are sequences $\langle A_n\rangle,\langle B_n\rangle,\langle C_n\rangle\subseteq A$, 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 $D \in F$ and $\frac{p_{C_n,A_n}(D')}{p_{C_n,B_n}(D')} \to 1$ for all $D' \in F$. $\frac{p_{C_n,A_n}(D')}{p_{C_n,B_n}(D')} \to 1$ for all $D' \in F$ 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 $D' \in F$. $\frac{p_{C_n,A_n \cup B_n}(D)}{p_{C_n,B_n}(D)} \nrightarrow 1$ 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 $1 + \eta \frac{c_{A_n,B_n}(s)}{\int\limits_S c_{C_n,B_n}(s)\mu(s|D)} > 0$ for all $s \in S$. From here the argument is identical to the one in the finite case. Thus, Axiom 1' must hold. ■

7.5. Proof of Theorem 2':

Definition 13: Let $\{D_t\,|t \in \{1,..,T\}\}$ be a finite partition of I with $D_t \in F$. $\{D_t\}$ 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 $G^* := \bigcup\limits_{\{D_t\}} G_{\{D_t\}}$ 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 $\mathcal{A}$.

Lemma 2': *$\succ$ constrained to $\overline{G} \cap G^*$ satisfies Axioms 2-4, 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 $v(g) = \sum\limits_{t=1}^{T} v_{D_t}(g(D_t))$ 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 $v'(g) = \sum\limits_{t=1}^{T} v'_{D_t}(g(D_t))$ 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: That $v(g) = \sum\limits_{t=1}^{T} v_{D_t}(g(D_t))$ for $g \in \overline{G} \cap G_{\{D_t\}}$ represents $\succ$ confined to $\overline{G} \cap G_{\{D_t\}}$, is implied by Lemma 2.

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 $D_t = \bigcup\limits_{t\in\tau} D'_t$, the usual induction argument yields

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

and thus $v_{D_t}(A) = \sum\limits_{t\in\tau} v_{D'_t}(A)$. Therefore, $v(g) = \sum\limits_{t=1}^{T} v_{D_t}(g(D_t))$ 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 Lemma 2. That the representation implies continuity and linearity of v and, thus, the axioms is obvious. □

As suggested in the text, first establish the result of Theorem 2 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 7: If $\succ$ satisfies Axioms 2-6, 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 F, such that, for $g \in G_{\{D_t\}}$,

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

represents $\succ$ on G^*. Furthermore, there is a function $v : G \to \mathbb{R}$ as in Lemma 2 that agrees with V on G^*.

Proof: Just as in the proof of Theorem 2, establish that there is an event dependent, positive scaling factor $\pi'(D)$, such that

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

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 Lemma 2', $v_{D \cup D'}(A) = v_D(A) + v_{D'}(A)$. 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 $D, D' \in F$. 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 $\bigcup D_t \in F$. 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 Claim 7. $\|$

Claim 8: The simple acts G^* are dense in G in the topology defined on G.

Proof: I will 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 $G_i = A$ and let $B_\varepsilon(A) \subseteq A$ be an open ball of radius $\varepsilon > 0$ around $A \in A$,

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

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 $p_i^{-1}(B_\varepsilon(g(i))) \subseteq G$.[48] Let a finite set $I' \subseteq I$ index the relevant dimensions

[48] 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 $i \in I$.

for this intersection. I will establish that there is always a simple act h with

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

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 A' be the set of all subsets of L. Then $A' \subset A$, and for all $A \in A$ there is $A' \in A'$ with $d_h(A, A') < \varepsilon$ by the definition of $d_h(A, B)$. Thus, there is an act in $\bigcap_{I'} p_i^{-1}\left(B_\varepsilon\left(g\left(i\right)\right)\right)$ with support only in 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 $\bigcap_{I'} p_i^{-1}\left(B_\varepsilon\left(g\left(i\right)\right)\right)$. $\|$

Claim 8 implies that, if

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

on G^*, which can be guaranteed according to Claim 7, then the continuous function $v(g)$ is uniquely determined on all of G.

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

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

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

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

for a given act $g \in G$ and some measure $\mu_i(s)$. 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 3) imply that the set of acts with only singletons in their support has a worst element, $\underline{g}$. Axiom 6 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 2':

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

This completes the proof of the sufficiency statement in Theorem 2'. That the axioms are also necessary for the existence of the representation is straight forward to verify. ■

7.6. Proof of Proposition 2

The following lemma is at the heart of the proof of Proposition 2:

Lemma 4: If I is finite and (π, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ 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 $i, j \in I$.

Proof: For any given $i \in I$, (π, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ represent the same preference, $\succ_i$.

Then $\widehat{\mu}(s|i)$ must be a probability measure with $\widehat{\mu}(s|i) \propto \frac{l(s)}{\widehat{l}(s)}\mu(s|i)$ and consequently

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

At the same time (π, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ represent the same preference across I. It is easy to verify that this implies $\phi(i) \int\limits_S \widehat{l}(s) \max\limits_{\alpha \in g(i)} (s \cdot \alpha)\, d\widehat{\mu}(s|i) \propto \pi(i) \int\limits_S l(s) \max\limits_{g(j)} (s \cdot \alpha)\, d\mu(s|i)$ for all $g \in G$ and hence $\phi(i) = \pi(i) \int\limits_S \frac{l(s)}{\widehat{l}(s)} d\mu(s|i)$,which establishes Lemma 4. $\square$

I^* and S^* are assumed to have finite cardinality T. According to Lemma 4, $\widehat{l}(s)$ has to solve the system of equations $\phi(i) \propto \pi(i) \sum\limits_{S^*} \frac{l(s)}{\widehat{l}(s)} \mu(s|i)$ for all $i \in I^*$. We want to establish that there is a neighborhood of π, such that all ϕ in this neighborhood allow an alternative representation, $\left(\phi, \widehat{\mu}, \widehat{U}\right)$. Interpret π and ϕ as vectors in $\mathbb{R}^T_+$. Denote by $\mu(s) \in \mathbb{R}^T_+$ the vector with i-th component $\mu(s|i)$ and by $\pi \odot \mu(s) \in \mathbb{R}^T_+$ 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 5: Under the conditions of Proposition 2, $\{\mu(s)\}_{s \in S^*}$ 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 $\mu(s_T) = \sum\limits_{n \in \{1,...,T-1\}} c_n \mu(s_n)$. Then for some $\tau \in (0, \infty) \setminus \{1\}$, one 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 for $U'_s(\alpha) = l'(s)\, s \cdot \alpha$ the CPF representation (π, μ', U') is numerically identical to the representation (π, μ, U). This contradicts Theorem 1. $\square$

$\pi \in \mathbb{R}^T_+$. Thus, $\{\pi \odot \mu(s)\}_{s \in S^*}$ must also be linearly independent. 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: $\pi = \sum\limits_{S^*} \pi \odot \mu(s)$. 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 2: 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, $\left(\phi, \widehat{\mu}, \widehat{U}\right)$. Since the solution of a linear system of equations is continuous in all parameters, it is continuous in π. This establishes the second part of Proposition 2. ■

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 (2009), "Subjective States: a more Robust Model." *Games and Economic Behavior,* forthcoming.

[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 (2009a), "A Theory of Bayesian Decision Making." mimeo

[14] Karni, Edi (2009b), "Subjective Probabilities on a State Space." mimeo

[15] Kopylov, Igor (2009), "Finite additive utility representations for preferences over menus." *Journal of Economic Theory*, Vol. 144, No. 1), 354-374.

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

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

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

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