Contingent Preference for Flexibility: Eliciting Beliefs from Behavior*

Philipp Sadowski†

December 8, 2011

Abstract

Following Kreps (1979), I consider a decision maker who is uncertain about her future taste. This uncertainty leaves the decision maker with a preference for flexibility: When choosing among menus containing alternatives for future choice, she weakly prefers menus with additional alternatives. Standard 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. I allow choice between menus to depend on objective states. My axioms provide a representation that uniquely identifies beliefs, provided objective states are sufficiently relevant for choice. I suggest this result as a choice theoretic foundation for the assumption, commonly made in the (incomplete) contracting literature, that contracting parties who know each others' ranking of contracts, also share beliefs about each others' future tastes in the face of unforeseen contingencies.

Keywords: Preference for Flexibility, Unique Beliefs, Unforeseen Contingencies, Incomplete Contracts

1. Introduction

The expected utility model of von Neumann and Morgenstern (1944, henceforth vNM) explains choice under risk by considering probabilities and taste (a ranking of outcomes) separately. In the context of choice under subjective uncertainty, the corresponding separation of beliefs and tastes is a central concern. For the extreme case where all subjective uncertainty can be captured by objective states of the world, the works of Savage (1954) and Anscombe and Aumann (1963, henceforth AA) achieve this separation. In the opposing extreme, where none of the subjective uncertainty can be captured by objective states, 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, Bart Lipman, Mark Machina, Eric Maskin, Tymon Tatur, Justin Valasek, numerous seminar audiences, and four anonymous referees for helpful suggestions.

†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; a relevant corrigendum is Dekel et al. [2007, henceforth DLRS])[1] find that the separation is not possible in this case. This is the standard indeterminacy of state dependent expected utility models.

In the general case, some, but potentially not all, subjective uncertainty can be captured by objective states. 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 tastes and beliefs over those tastes, provided that objective states are "relevant enough." The main identification result provides a tight behavioral characterization of relevant enough.

The timing of choice is as follows: In period 1, the decision maker (DM) chooses an opportunity act. An opportunity act specifies a menu of alternatives for future choice contingent on the objective state. Between periods 1 and 2 an objective state realizes. In period 2 the act is evaluated and DM gets to choose from the resulting menu. Only period 1 choice is observed. If objective states do not account for all subjective uncertainty that resolves between periods 1 and 2, then DM has contingent uncertainty about her future taste. In that case, commitment to a contingent plan of period 2 choice is costly and one should observe contingent preference for flexibility: All else being equal, DM prefers an act that assigns a menu with additional alternatives to any particular state.

This paper provides a representation of such preferences, labeled a representation of Contingent Preference for Flexibility (CPF). As in DLR, subjective uncertainty is modeled via 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 behaves as if the objective state may be informative about her future tastes, and so conditions her beliefs about future tastes on the objective state. Contingent on the state, choice over menus has a subjective expected utility representation, as in DLR. I show that a central new axiom, *Relevant Objective States*, is equivalent to the unique identification of utilities and conditional beliefs in this representation.

To be more specific, let I be the objective state space. An opportunity act, g, assigns a contingent menu of lotteries over prizes, $g(i)$, to every objective state, i in I. The taste space, S, collects all possible vNM rankings of lotteries over prizes. In the case of finite I,

[1]Throughout the paper I refer to the version of their model that represents preference for flexibility.

[2]In the 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.

choice over acts has a CPF representation, if it can be represented by

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

where ϕ is a probability measure on I, the realized vNM utility function U_s represents taste s in S, and $\mu_i(s)$ is a probability measure on S. The representation suggests that, while the menu of alternatives DM expects to choose from in stage 2 depends on i and not s, she anticipates a utility function that depends on s and not i. She also expects to learn s and i prior to choosing an alternative. The measure ϕ is interpreted as DM's prior over I and $\mu_i(s)$ is interpreted as the belief that taste s occurs, contingent on i.

Theorem 1 takes the CPF representation and the distribution ϕ as given.[3] It establishes that conditional beliefs $\mu_i(s)$ are unique and utilities U_s are unique in an appropriate sense if and only if choice between opportunity acts satisfies the Relevant Objective States axiom. The axiom is formulated in terms of DM's ranking of menus contingent on the objective state, which is derived from her choice over acts. Say that two menus are *equivalent for DM*, if for every contingent ranking the union of those menus is indifferent to each of the menus individually. Objective states are relevant, if for any two menus that are not equivalent for DM, there is an objective state contingent on which one is strictly preferred over the other.

Theorem 2 states that choice over opportunity acts has a CPF representation if and only if it satisfies the immediate extensions of the AA and DLR axioms. These axioms are necessary for a more general representation, where both beliefs and utilities depend on objective states. For the separation of beliefs and tastes, however, it is important that only beliefs condition on objective states. Theorem 2 implies that this interpretation is always possible, as it does not constrain period 1 choice.

Even though the model does not capture period 2 choice from a menu explicitly, a researcher may want to forecast period 2 choice behavior. The CPF representation describes choice over menus in period 1, as if the DM held beliefs about the tastes that might govern her period 2 choice. Theorem 1 uniquely identifies those beliefs, which are parameters of the representation, from period 1 choice. Therefore, the natural inductive step is to employ the DM's beliefs about future tastes to forecast period 2 choice behavior. There are good arguments against this inductive step. For example, one could instead make period 2 choice part of the domain, leaving less room for erroneous modeling assumptions.[4] However, an

[3] ϕ could be objective. If ϕ is subjective as suggested above, it must also be elicited from choice. I address this case in Theorem 3.

[4] Ahn and Sarver (2011) provide a model that connects preference for flexibility in period 1 to choice frequencies in period 2.

essential reason for the use of scientific models is to make predictions about the world based on limited data. Choice Theory is well positioned to supply such models for economic applications: Axiomatization translates limited data (here period 1 choice data) to a model, and identification establishes those parameters of the model (here the beliefs) that one might base inferences on.

Being able to forecast behavior can be important in strategic situations, for example when one party's valuation of a contract depends on future actions taken by the other party. Contracting models usually do not only assume that parties know each others' ranking of contracts, but require the stronger assumption that 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 others' ranking of contracts; this question is usually not formally addressed in applied models and is not my focus here. Instead, I am concerned with the second assumption. If two parties write a contract in the face of indescribable or unforeseen contingencies,[5] it seems natural that there might be asymmetric information about those contingencies. 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.[6]

As an illustrative example of a CPF representation, consider a retailer, who writes a contract with her supplier today about tomorrow's order. 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. (While demand is observable in many situations, unobservable demand levels here simply serve as convenient labels for the different unobservable profit functions the retailer can conceive of.) 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. The most efficient contract might give the retailer some choice of supply quantities, q, contingent on consumer confidence; consider this type of contract. From the retailer's perspective, the contract is an act in the terminology of this paper. Routinely one might write down the following objective function for the retailer's

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

[6] 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."*

choice between contracts:

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

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, $\mu_i(s)$. She values the menu $g(i)$ at its expected value, $\sum_{s\in\{h,l\}} \mu_i(s) \max_{q\in g(i)} (U_s(q))$. 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.[7]

The example also speaks to the possible strategic value of uniquely identified beliefs. Suppose the retailer has private knowledge about tomorrow's demand, contingent on consumer confidence. Demand may affect the supplier's profit indirectly, through tomorrow's choice of quantity by the retailer. Therefore, when evaluating contracts, the supplier would like to forecast demand based on the retailer's beliefs. If the retailer's preferences have a CPF representation with uniquely identified beliefs, $\mu_i(s)$, then the supplier is able to infer those beliefs from the retailer's preferences.

Section 2 investigates under what conditions beliefs are identified in the example above. In the main part of the paper the objective state space, I, is assumed to be finite. Section 3 lays out the model and establishes Theorems 1 and 2. Section 4 contains Theorem 3, which combines the two results and elicits beliefs ϕ on I from choice. Section 5 discusses related literature. Section 6 comments in more detail on possible implications for contracting. Appendix A characterizes partial failures of the Relevant Objective States axiom and provides existence and identification results for the case of a general measurable objective state space. Most proofs are relegated to Appendix B.

2. Illustration of Identification of Beliefs

In this section, I consider three cases of a CPF representation: when none of the subjective uncertainty can be captured by objective states (irrelevant objective states); when all of the subjective uncertainty can be captured by objective states (no preference for flexibility); and the general case, where some, but not all, of the subjective uncertainty can be captured by

[7] The CPF representation also evaluates more general contracts, where, contingent on consumer confidence, the retailer is given some choice between lotteries over different quantities. For example, the contract might specify an action which has probabilistic consequences.

objective states (preference for flexibility and relevant objective states). To illustrate these cases, I use the setup of the above example, but where final outcomes are lotteries, α, over quantities.

• Irrelevant objective states: Suppose that the retailer's beliefs are independent of consumer confidence; that is $\mu_H(h) = \mu_L(h) = \mu(h)$ and

$$V(g) = \sum_{i \in \{H,L\}} \phi(i) \left[\sum_{s \in \{h,l\}} \mu(s) \max_{q \in g(i)} (U_s(q)) \right].^8$$

In this case, her induced ranking of menus is independent of consumer confidence and it is without loss of generality to consider only contracts with $g(H) = g(L)$. If g is such a non-contingent contract, then

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

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 models.

• No preference for flexibility: Suppose that $\mu_H(h) = 1$ and $\mu_L(h) = 0$. Now subjective uncertainty is perfectly captured by the objective states, and it is without loss of generality to identify h with H and l with L. This implies that none of the contingent rankings exhibit preference for flexibility. 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 state-dependent representation.

[8] Ozdenoren (2002) provides a model that generalizes this example, as discussed in Section 5.

• Preference for flexibility and relevant objective states: Lastly, suppose the retailer believes that the probability of high demand is increasing with consumer confidence; that is, $1 > \mu_H(h) > \mu_L(h) > 0$. Further suppose that there is another representation of the same ranking of contracts based on a prior over objective states, $\widehat{\phi}$, beliefs, $\widehat{\mu}_i(s)$, and tastes, $\widehat{U}_s$:

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

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 the menu that contains lotteries α and β by $\{\alpha, \beta\}$. Suppose for some $\beta < \alpha$ 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

$$\phi(H)\mu_H(h)\delta(U_h(q_h) - U_h(q_l)) = \phi(L)\mu_L(h)\varepsilon(U_h(q_h) - U_h(q_l)).$$

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

$$\frac{\mu_H(h)}{\mu_L(h)} = \frac{\varepsilon\phi(L)}{\delta\phi(H)} \text{ and } \frac{\widehat{\mu}_H(h)}{\widehat{\mu}_L(h)} = \frac{\varepsilon\widehat{\phi}(L)}{\delta\widehat{\phi}(H)}.$$

If probabilities of objective states are objective, that is, $\widehat{\phi} = \phi$, then

$$\frac{\mu_H(h)}{\mu_L(h)} = \frac{\widehat{\mu}_H(h)}{\widehat{\mu}_L(h)},$$

and similarly

$$\frac{\mu_H(l)}{\mu_L(l)} = \frac{\widehat{\mu}_H(l)}{\widehat{\mu}_L(l)}.$$

Since μ and $\widehat{\mu}$ are both probability measures, $\frac{\mu_H(h)}{\mu_L(h)} \neq 1$ immediately implies that $\mu \equiv \widehat{\mu}$. Standard arguments, applied to the comparison of contracts which disagree only under state i, imply that the expected utility functions $\widehat{U}_h$ and $\widehat{U}_l$ can only differ from their respective counterparts U_h and U_l by a common linear transformation and the addition of constants. This argument illustrates how identification relies crucially on the fact that beliefs ϕ over objective states are held fixed. More generally, it makes clear why it is necessary to observe the retailer's choice between contracts (opportunity acts), and not just her ranking of menus contingent on each objective state: the willingness to tradeoff payoffs across objective states (captured by the indifference between contracts g and g') determines the relative weight assigned to taste h under objective state H versus L.[9]

The above reasoning can be generalized to any finite state space, I. If a CPF representation has the feature that there are 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 Objective States must have this feature.

3. A Model with Unique Beliefs

Consider a two-stage choice problem, where an objective state realizes between the two stages. In period 2 DM chooses a lottery over prizes. This choice is not modelled explicitly. In period 1 DM chooses an opportunity act. Such an act is a state contingent specification of a set of lotteries (a menu) that contains the feasible alternatives for period 2 choice.

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. When there is no risk of confusion, x also denotes the degenerate lottery that assigns unit weight to x. Let $\mathcal{A}$ be the collection of all compact subsets of $\Delta(Z)$ with menus A, B, C as typical elements.[10] Further, let I be a finite objective state space with typical elements i, j. Let $\mathcal{F}$ be the σ-algebra generated by the power set of I, where $i, j \in I$ also denote elementary events.[11]

Let G be the set of all opportunity acts with typical elements g, h. An opportunity act is a measurable function $g : I \to \mathcal{A}$. If state i realizes, DM gets to choose an alternative from

[9] More specifically, my identification strategy relies on the linear aggregation of objetive states. Ozdenoren (2002) provides a representation of preferenes over opportunity acts that can accomodate ambiguity aversion with respect to objective states.

[10] 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.

[11] The case of a general measurable space $(I, \mathcal{F})$ is relegated to Appendix A.

the menu $g(i) \in \mathcal{A}$. This choice is not explicitly modeled. Instead, $\succ$ is a binary relation on $G \times G$; $\succcurlyeq$ and $\sim$ are defined the usual way.

The following concepts are important throughout the paper.

Definition 1: The convex combination of menus is defined as

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

The convex combination of opportunity acts is defined, such that

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

To define DM's induced ranking of menus A and B contingent on state $i \in I$, consider acts g_i^A and g_i^B that give menu A and B, respectively, in state i and some arbitrary but fixed default menu, A^*, in every other state. Comparing g_i^A and g_i^B induces a ranking $\succ_i$ over menus. In the context of the model below, $\succ_i$ is independent of A^*.

Definition 2: Fix an arbitrary menu $A^* \in \mathcal{A}$. For $i \in I$ and $A \in \mathcal{A}$, define g_i^A by

$$g_i^A(j) := \begin{cases} A \text{ for } j = i \\ A^* \text{ otherwise} \end{cases}.$$

Let the *contingent ranking* $\succ_i$ be the induced binary relation on $\mathcal{A} \times \mathcal{A}$, $A \succ_i B$ if and only if $g_i^A \succ g_i^B$; $\succcurlyeq_i$ and $\sim_i$ are defined the usual way. A state $i \in I$ is *nonnull*, if there are $A, B \in \mathcal{A}$ with $A \succ_i B$.

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: The set

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

is the *taste space*.[12] Let $\mathcal{B}$ be the Borel σ-algebra on S.

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

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: Call (ϕ, μ, U) a *Contingent Preference for Flexibility (CPF) representation* of the preference relation $\succ$, if ϕ is a probability measure on I, $\mu = \{\mu_i\}_{i \in I}$ is a probability kernel from $(I, \mathcal{F})$ to $(S, \mathcal{B})$, and $U = \{U_s\}_{s \in S}$ is a family of vNM utilities on $\Delta(Z)$, integrable in s, where U_s represents taste s and the objective function

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

represents $\succ$.

If U_s is a vNM representation of taste s, then it must have the form $U_s(\alpha) = l(s)(s \cdot \alpha) + b_s$, where $s \cdot \alpha$ is the dot product of state s and lottery α, $l(s)$ is the "intensity" of taste s and b_s is a constant. The relative intensity of utilities together with beliefs determines how DM trades off gains across tastes. The constants b_s have no behavioral content. This motivates the next definition.

Definition 5: Let (ϕ, μ, U) be a CPF representation of $\succ$.
i) The *space of relevant objective states,* $I^* \subseteq I$, is the minimal set with $\phi(I^*) = 1$. The *space of relevant tastes is* $S^* := \bigcup_{i \in I^*} \text{supp}(\mu_i)$.[14]
ii) μ and U are *unique given* ϕ, if for any other CPF representation $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ of $\succ$, and for all $i \in I^*$ and all $S' \subseteq S^*$, $\widehat{\mu}_i(S') = \mu_i(S')$ and there are $a > 0$ and $\{b_s\}_{s \in S^*} \subset \mathbb{R}$, such that $\int_{S'} \widehat{U}_s d\mu_i(s) = \int_{S'} (aU_s + b_s)\, d\mu_i(s)$.

S^* can be thought of as the set of tastes DM considers possible. 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 tastes U, *provided* that DM's choice over acts has a CPF representation for a given distribution ϕ.

[13] In the context of the representation theorem in DLRS, as in the theorems that follow, there is clearly always a larger taste space, also allowing a representation of $\succ_D$, in which multiple tastes represent the same ranking of lotteries.

[14] The support of a measure is the closure of the collection of points, for which every neighborhood in $\mathcal{B}$ has positive measure.

Axiom 1 *(Relevant Objective States)*: *If* $A \cup B \nsim_i B$ *for some* $i \in I$, *then there is* $j \in I$ *with* $A \nsim_j B$.

Two menus that are distinct elements of $\mathcal{A}$ might still be *equivalent for DM*, in the sense that they provide her with the same utility under every relevant taste. If A and B were equivalent for DM, then she should be willing to choose from $A \cup B$ by simply ignoring A.[15] This cannot be the case if $A \cup B \nsim_i B$ for some $i \in I$. If $A \nsim_i B$, then Axiom 1 is empty. If $A \sim_i B$, then $A \cup B \nsim_i B$ implies that, contingent on i, the item chosen from $A \cup B$ must sometimes be in A and sometimes in B. Axiom 1 requires that there exists a contingent ranking for which either one or the other case becomes more important, namely that there is $j \in I$ with $A \nsim_j B$. To paraphrase Axiom 1: If two menus are not equivalent for DM, then there exists an objective state contingent on which, one is preferred over the other.[16] Axiom 1 is not a strong assumption in the sense that it is local; it only requires breaking indifferences. For comparison, suppose instead that the state $i \in I$ was required to provide a complete description of all relevant aspects of the world, as in AA. Then the stronger assumption of state contingent strategic rationality would have to hold: If $A \cup B \succ_i B$, then $A \sim_i A \cup B$.[17] In terms of the example from the introduction, consumer confidence (the objective state) may be relevant for the retailer's beliefs about her desire to order a large or a small quantity, but it is conceivable that the retailer prefers a large quantity even when confidence is low and vice versa. This notion is weaker than the assumption of state contingent strategic rationality, according to which the retailer always prefers the large quantity when confidence is high and the small quantity when confidence is low.

Theorem 1: *If* $\succ$ *has a CPF representation* (ϕ, μ, U), *then statements i) - iii) below are equivalent:*
i) $\succ$ *satisfies Axiom 1,*
ii) μ *and* U *are unique given* ϕ,
iii) the cardinality of S^* *equals the number of linearly independent elements in* $\{\mu_i\}_{i \in I^*}$.

Proof: See Appendix B.

If a decision maker behaves as if she has preference for flexibility because of uncertainty

[15] Implicit in the interpretation is that, ultimately, only the chosen item matters for the value of a menu.

[16] To see the content of the axiom in terms of the representation, suppose for simplicity that there are only two subjective states, s_A and s_B, where in state s_A the best element of $A \cup B$ lies in A, and in state s_B it lies in B. Suppose both states are supported by μ_i. This immediately implies $A \cup B \nsim_i B$. If $A \sim_i B$, then the axiom requires that there is $j \in I$ with $\mu_j(s_A) \neq \mu_i(s_A)$.

[17] Axiom 1 is immediately satisfied: $A \cup B \nsim_i B$ implies $A \nsim_i B$.

about her future taste, updates her beliefs over tastes when learning the objective state, and maximizes her expected utility according to objective probabilities over those states, then her preferences satisfy Axiom 1 if and only if her beliefs over future tastes are determined uniquely. This identification gives meaning to the description of beliefs and tastes as distinct concepts. 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. Item iii) in Theorem 1 conveniently constrains the space of relevant tastes, S^*, to be smaller than I, which is 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. Theorem 1′ in the appendix generalizes the result and considers I to be a general topological space, lifting the constraint on the cardinality of S^*.

Finally, given a particular CPF representation of $\succ$, item (iii) in Theorem 1 provides a criterion to check whether or not $\succ$ satisfies Axiom 1. This criterion is illustrated in the example in Section 2.

Remark: A remark on the interpretation of tastes, or subjective states, is in order. Suppose for a moment that there is an underlying state space Ω, which provides a complete description of all relevant aspects of the world. That is, $\omega \in \Omega$ even determines DM's taste, $s \in S$. In that case, S generates a sub σ-algebra on Ω. The question is to what extent Ω is *observable.* Let I be the collection of observable events $i \subset \Omega$, where I generates another sub σ-algebra on Ω. Now consider a probability measure μ on Ω representing DM's beliefs. If there is no correlation between events in I and events in S, then the induced marginal distribution $\mu_i(s)$ is independent of i, and the objective state space Ω can be dropped from the description of the model, as in DLR. For example, Ω could be the product space $I \times S$ and μ a product measure. If, in the other extreme, there is perfect correlation between events in I and events in S, then I itself can play the role of the complete objective state space in (the state dependent version of) the AA model. Theorem 1 is concerned with the general case of some, but not perfect correlation. While I is naturally interpreted as the collection of all observable contingencies, I will call events that are not in I "unobservable contingencies."

To see how relevant objective states imply unique beliefs and utilities, fix the distribution of objective states, ϕ, and suppose there were two CPF representations of the same preference relation, (ϕ, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ with corresponding value functions $V(g) = \sum_{i \in I} \phi(i) V_i(g(i))$ and $\widehat{V}(g) = \sum_{i \in I} \phi(i) \widehat{V}_i(g(i))$, where $V_i(A) = \int_S \max_{\alpha \in A} (U_s(\alpha)) \, d\mu_i(s)$ and analogously for $\widehat{V}_i(A)$. Neglecting additive constants, additive separability of the representations implies that $\phi(i) V_i(.) = \lambda \phi(i) \widehat{V}_i(.)$ for all $i \in I^*$ and for some $\lambda > 0$. Suppose further that for the contingent ranking $\succ_i$ one could construct menus $K \sim_i \widehat{K}$, such that K generates constant utility payoff across tastes according to (ϕ, μ, U) and $\widehat{K}$ according to $\left(\phi, \widehat{\mu}, \widehat{U}\right)$. On the one hand, $K \sim_i \widehat{K}$ would imply $V_i(K) = V_i\left(\widehat{K}\right) = \lambda \widehat{V}_i\left(\widehat{K}\right)$. On the other hand, changing the objective state from i to j only changes DM's beliefs about her future tastes. If a menu generates the same utility payoff for every taste, then the conditional value of the menu is independent of the objective state: $V_j(K) = V_i(K)$ and $\widehat{V}_j\left(\widehat{K}\right) = \widehat{V}_i\left(\widehat{K}\right)$ for all $j \in I^*$. Hence, $V_j(K) = \lambda \widehat{V}_j\left(\widehat{K}\right)$ or $K \sim_j \widehat{K}$ would have to hold for all $j \in I^*$. At the same time, if (ϕ, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ were distinct, $\widehat{K}$ would not generate constant utility payoffs across tastes according to (ϕ, μ, U), because utility payoffs depend on the intensities of U and $\widehat{U}$, respectively. Therefore $K \cup \widehat{K} \succ_{j'} K$ for some $j' \in I$. Relevant Objective States 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 utility payoff for every taste is not always possible. Because $S^* \subset S$ is finite, however, 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 utility 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 $K \sim_i \widehat{K}$ implies that $K \sim_j \widehat{K}$ for all $j \in I$. By the type of argument laid out above, $K \cup \widehat{K} \succ_{j'} K$ for some $j' \in I$. This contradicts Axiom 1.

Ozdenoren (2002) analyzes the case where Axiom 1 fails completely, in the sense that objective states are irrelevant to the decision maker. Then, only the support of the probability measures μ_i which allow a representation can be identified. This is the same indeterminacy encountered in DLR. Partial failures of the axiom are considered in Appendix A.1.

Both types of exogenous uncertainty in my domain are essential for the uniqueness result: On the one hand, DLR find that preferences over menus of lotteries alone do not allow the separate identification of tastes and beliefs μ. There has to be some possibility of varying one, but not the other. In the CPF representation, only beliefs, μ_i, condition on objective states. On the other hand, Nehring (1999) finds that acts with menus of prizes as outcomes do not allow the separate identification of tastes and beliefs in the axiomatic setup developed by Savage (1954). To establish the uniqueness result, the payoff generated by a menu must be varied independently for different tastes. This is possible only because DM can be offered lotteries over prizes.

I now establish existence of a CPF representation. As mentioned above, the axioms are direct extensions of familiar assumptions.

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

Axiom 3 *(vNM Continuity): If $g \succ h \succ g'$,* then there exist $p, q \in (0,1)$ such that $pg + (1-p)\,g' \succ h \succ qg + (1-q)\,g'$.

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."

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

The next axiom considers DM's contingent ranking of menus, $\succ_i$. As long as some subjective uncertainty is not captured by objective states, $\succ_i$ 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): $A \cup B \not\prec_i A$ for all $A, B \in \mathcal{A}$ and all $i \in I$.*

Lemma 1: *If $\succ$ satisfies Axioms 2-6, then $\succ_i$ is a preference relation and satisfies the appropriate variants of vNM Continuity, Independence and Monotonicity for all $i \in I$. Furthermore, there is a nonnull event $i \in I$.*

The proof is immediate.

Theorem DLRS *(Theorem 2 in DLRS)*: *For $i \in I$ nonnull, $\succ_i$ is a preference that satisfies the appropriate variants of vNM Continuity, Independence and Monotonicity if and only if there is a subjective state space S_i, a positive countably additive measure μ_i on S_i, and a set of non-constant and continuous expected utility functions $U_{s,i} : \Delta(Z) \to \mathbb{R}$, such that*

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

represents $\succ_i$ and every vNM ranking of lotteries in $\Delta(Z)$ *corresponds to at most one state in S_i.*[18]

Because $U_{s,i}(\alpha)$ are realized vNM utility functions, the subjective state space S_i can be replaced by the taste space S for all $i \in I$. Note that the taste space does not include the taste where DM is indifferent between all prizes, implicitly assuming nontriviality of the ex-post preferences over prizes.

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

Proof: See Appendix B.

Let $\overline{G}$ denote the collection of acts with support in $\overline{\mathcal{A}}$, the convex subsets of $\Delta(Z)$. The proof first employs the Mixture Space Theorem to establish an additively separable representation of $\succ$ constrained to $\overline{G}$. That is, $\sum_{i \in I} v_i(g(i))$ represents $\succ$ on $\overline{G}$ for some family of utility functions, $\{v_i\}_{i \in I}$, on $\overline{\mathcal{A}}$, where v_i are unique up to a common positive linear transformation and the addition of constants. Now consider some linear representation, V_i, of $\succ_i$ on $\mathcal{A}$. Since v_i must represent $\succ_i$ on $\overline{\mathcal{A}}$, the Mixture Space Theorem implies that v_i agrees with V_i up to scaling on $\overline{\mathcal{A}}$. The scaling is absorbed by $\phi(i)$, which is then normalized to be a probability distribution. Thus,

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

represents $\succ$ on $\overline{G}$. Next, for any act $g \in G$ there is an act $\overline{g} \in \overline{G}$, such that $g(i) \sim_i \overline{g}(i)$ for all $i \in I$ and, by Independence, $g \sim \overline{g}$. Since V_i represents $\succ_i$ on $\mathcal{A}$, $V_i(g(i)) = V_i(\overline{g}(i))$

[18] See footnotes 3 and 5 in DLRS.

and, because V is additively separable, $V(g) = V(\overline{g})$. Thus, V represents $\succ$ on G. Note that this is AA's state-dependent representation, with the exception that opportunity acts have menus as outcomes, while AA acts have lotteries as outcomes. Indeed, Axioms 2-4 imply AA's axioms. Furthermore, Axioms 2-6 imply DLRS' axioms, as shown in Lemma 1. According to Theorem DLRS, $\succ_i$ can then be represented by

$$\widehat{V}_i(A) = \int_S \max_{\alpha \in A} \left(\widehat{U}_{s,i}(\alpha)\right) d\widehat{\mu}_i(s),$$

where $\widehat{\mu}_i$ is a probability measure on S and $\widehat{U}_{s,i}$ is a vNM utility function that represents taste $s \in S$, that is, $\widehat{U}_{s,i}$ and $\widehat{U}_{s,j}$ are identical up to a positive affine transformation. Pick any $j \in I$ and define $U_s := \widehat{U}_{s,j}$. The lack of identification in DLRS implies that there is a measure μ_i on S, such that $\mu_i(s) U_s \propto \widehat{\mu}_i(s) \widehat{U}_{s,i}$. Therefore, $\succ_i$ can be represented by

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

for all $i \in I$. Since V_i is linear, there is a CPF representation (ϕ, μ, U); that is,

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

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

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 objective states impact not only probabilities, μ_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. Therefore, the assumption that only beliefs condition on objective states does not constrain period 1 choice.

4. Probabilities over Objective States

Theorem 1 takes the distribution ϕ on I and a CPF representation (ϕ, μ, U) as given and establishes that μ and U are unique in the appropriate sense if and only if objective states are relevant. ϕ might be objective in the sense that it corresponds to observed frequencies of objective states, or it might be subjective, in which case it must also be elicited from behavior. The unique identification of ϕ is analogous to the classical problem addressed by AA. There, the unique identification of probabilities of observable 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 *objective state independence* and Axiom 1 would rule out any preference for flexibility. Thus, the independence assumption has to be confined to a proper subset $\Psi \subset \mathcal{A}$ to be useful here. Having assumed state independent rankings, AA 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 result. For the CPF representation it would amount to requiring that $\int_S \max_{\alpha \in A} U_s(\alpha)\, d\mu_i(s)$ 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.[19] Thus, the requirement must be confined to a *particular* collection of menus.

Definition 6: Let $X \subseteq Z$ denote a set of prizes and $\Delta(X)$ the set of all lotteries with support in X. Let $\Psi(\Delta(X)) \subseteq \mathcal{A}$ be the set of all menus of lotteries that have support in X.

Definition 7: A CPF representation, (ϕ, μ, U), is *state-independent with respect to* $X \subseteq Z$, if $\int_S \max_{\alpha \in A} U_s(\alpha)\, d\mu_i(s) = \int_S \max_{\alpha \in A} U_s(\alpha)\, d\mu_j(s)$ for all $A \in \Psi(\Delta(X))$ and all $i, j \in I^*$. Further, (ϕ, μ, U) is the *unique CPF representation that is state-independent with respect to* X, if for any other CPF representation $\left(\widehat{\phi}, \widehat{\mu}, \widehat{U}\right)$ that is state-independent with respect to X, $\widehat{\phi} \equiv \phi$, $\widehat{\mu}_i \equiv \mu_i$ for all $i \in I^*$, and there are $a > 0$ and $\{b_s\}_{s \in S^*} \subset \mathbb{R}$, such that $\widehat{U}_s \equiv aU_s + b_s$ for all $s \in S^*$.

Axiom 7 *(Partial Objective State Independence): There is a non-degenerate $X \subseteq Z$, such that for $A, B \in \Psi(\Delta(X))$, $A \succ_i B$ for some $i \in I$ implies $A \succ_j B$ for all nonnull $j \in I$. If $\succ$ satisfies the same condition for $Y \subseteq Z$, then it also satisfies the condition for $X \cup Y$.*

To illustrate Axiom 7, consider $X = \{\$1, \$0\}$ to consist of the prizes "1 Dollar" and "nothing." The first part of Axiom 7 then requires that the ranking of menus that consist only of

[19] For a simple example of such inconsistency consider $\Psi = \{\{\alpha\}, \{\beta\}, \{\gamma\}\}$ but, for some $p \in (0,1)$ and $i, j \in I$, $\{p\alpha + (1-p)\gamma\} \succ_i \{\beta\} \succ_j \{p\alpha + (1-p)\gamma\}$. Since $\int_S \max_{\alpha \in A} U_s(\alpha)\, d\mu_i(s)$ is linear, it can not be independent of $i \in I$.

lotteries that pay out either \$1 or nothing must be state-independent. To motivate the requirement, it is sufficient to assume that the value of \$1 (versus nothing) is state-independent.

Once AA restrict attention to representations with state-independent utilities, there is no arbitrariness in their model. In contrast, preference for flexibility implies that X is a proper subset of Z. Hence, $\succ$ could satisfy the first part of Axiom 7 for some X and Y with $X \neq Y$, but not for $X \cup Y$. Either those menus with support in X or those with support in Y could then be assigned a cardinal ranking, which is state-independent. 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 7 rules out this scenario, suggesting the following definition:

Definition 8: Let $X^* \subseteq Z$ be the largest set for which $\succ$ satisfies the condition in Axiom 7.

Theorem 3: *The binary relation $\succ$ satisfies Axioms 1-7 if and only if it has a unique CPF representation that is state independent with respect to X^*, (ϕ, μ, U). In this representation $U_s(x)$ constant across S^* for all $x \in X^*$.*[20]

Proof: For a CPF representation where $\int_S \max_{\alpha \in A} U_s(\alpha)\, d\mu_i(s)$ does not depend on $i \in I$ for any $A \in \Psi(\Delta(X^*))$, the uniqueness of ϕ follows in complete analogy to the corresponding result in AA. Given this unique ϕ, Theorem 1 implies uniqueness of μ and uniqueness of U_s up to a common rescaling and the addition of constants. The existence of a representation where $U_s(x)$ is constant across S for all $x \in X^*$ is established in Appendix B. Hence, the unique representation must have this feature. ■

5. Related Literature

Ozdenoren (2002) also considers Preference for Flexibility in the presence of objective states of the world. Instead of Relevant Objective States (Axiom 1), which ensures that contingent rankings are sufficiently different, he assumes that all contingent rankings are the same.

[20] In the case where ϕ is objective, it is possible to strengthen Axiom 8, such that the *unique* CPF representation in Theorem 3 is based on ϕ:
Axiom *(Objective Probabilities): There is $X \subseteq Z$, such that for $A, B \in \Psi(\Delta(X))$ and nontrivial D and $D' \in \mathcal{F}$:*
$$\frac{\phi(D')}{\phi(D)+\phi(D')} h_D^A + \frac{\phi(D)}{\phi(D)+\phi(D')} h_{D'}^B \sim \frac{\phi(D)}{\phi(D)+\phi(D')} h_{D'}^A + \frac{\phi(D')}{\phi(D)+\phi(D')} h_D^B.$$
If $\succ$ satisfies the same condition for $X' \subseteq Z$, then it alo satisfies the condition for $X \cup X'$.
This implies Axiom 8. It also implies that $V(g_D^A) - V(g_D^B) = \left(V(g_{D'}^A) - V(g_{D'}^B)\right) \frac{\phi(D)}{\phi(D')}$ for $A, B \in \Psi(\Delta(X))$.

Consequently, beliefs are not identified in his model.

I know of three other identification results that deliver unique beliefs over future tastes for consumption in models of preference for flexibility. First, note that AA's identification of unique beliefs over objective states does not require *full* state independence of preferences.[21] In analogy to AA's argument, beliefs over tastes in the DLR model can be identified uniquely, as long as DM has no preference for flexibility with respect to *part* of the prize space. As an example, DLR suggest to consider a DM without preference for flexibility on one dimension of a product prize space (Shenone (2010) provides details.) Second, Ahn and Sarver (2011) provide a model that requires both choice between as well as random choice from menus to be observable. Their model restricts the beliefs that feature in the representation of choice between menus to correspond to the choice frequencies that describe choice from menus. Finally, in a dynamic model of preference for flexibility, Krishna and Sadowski (2011) show that the DM's attitude towards intertemporal tradeoff can also uniquely identify beliefs. They proceed to characterize a behavioral comparison of "greater preference for flexibility" in terms of a stochastic dominance condition on the beliefs. Without identification of beliefs, such a comparison is not possible.[22]

The domain of opportunity acts is first analyzed by Nehring (1999), and the notion of contingent menus appears in Epstein (2006). Following Nehring (1996), a companion paper to Nehring (1999), Epstein and Seo (2009) consider a domain of random menus, which are lotteries with menus as outcomes. On this domain they establish unique induced probability distributions over ex post upper contour sets as the strongest possible uniqueness statement.

Theorem 1 does not only provide unique beliefs, but also establishes the finiteness of the collection of relevant tastes, S^*. Dekel, Lipman and Rustichini (2009) and Kopylov (2009) generate finiteness of S^* in the absence of objective states by basically assuming that the number of lotteries DM can appreciate in any given menu is limited.

Finally, note that the state-independent version of AA's representation can be viewed as a special case of a unique CPF representation, where there is only one taste and the intensity of this one taste is independent of the objective state. Karni and coauthors, for example Grant and Karni (2005), Karni (2008), and Karni (2011a and 2011b), elaborate the point that interpreting AA's or Savage's (1954) unique subjective probabilities of observable states as DM's true beliefs may be misleading, in case the true intensity of her only taste is actually not state independent. The CPF model is not immune to this concern: Even if choice has a CPF representation, DM's true intensities of tastes might not actually be state independent.

[21] This insight also underlies the elicitation of beliefs, ϕ, over objective states in Section 4 of this paper.

[22] Limited by the lack of identification in their model, DLR suggest an alternative notion which can be characterized in terms of the support of the beliefs.

Similarly, the DM might not actually use the expected utility criterion to evaluate uncertain prospects, or alternatives other than the one that is ultimately chosen might also generate utility. None of those modeling assumptions remain innocuous, once the natural inductive step of forecasting period 2 choice is taken.

6. Asymmetric Information and Contracts

The CPF model interprets choice between acts in terms of unique state contingent beliefs over future tastes. Considering those beliefs as predictors of future choice[23] is an additional assumption. As I argue in the introduction, I view this type of assumption as the essence of model based forecasting. Axiomatizations translate limited choice data to models, and identification results establish those parameters of the models that one might base inferences on. For example, in the framework of expected utility, one may identify a person's risk aversion in one context with the curvature of their utility function and proceed to make predictions about their behavior in an entirely different context. The advantage when observing choice between menus is that the unobserved choice from the menu is implicitly part of the domain. This should strengthen the case for inference, not weaken it. Therefore, rather than categorically rejecting the assumption that beliefs are good predictors of future choice, one should critically assess it in the context of a particular application.

Consider, then, a situation where it is commonly assumed that the DM's beliefs over tastes are correct in the sense that they provide an accurate forecast of her period 2 choice. Conceptually an observer could agree on those beliefs, as the DM's period 2 choice is potentially observable.[24] The identification of beliefs from behavior is necessary to render such agreement behaviorally meaningful. In this sense my model can be viewed as a foundation for "common priors over subjective states," just like the identification of beliefs in the work of Savage (1954) is necessary to talk about common priors over objective states.[25] In fact, once beliefs that are assumed to forecast future choice become identified, the observer *should* agree on them.

If beliefs are a good predictor of future choice, then the elicitation of those beliefs can be of strategic value in situations where each party would otherwise be more informed about its own future choice. I now describe in more detail a strategic situation where this seems plausible.

[23] That is, beliefs correspond to choice frequencies.

[24] The interpretation of sujective states as tastes (vNM-rankings) is essential for this argument. In contrast, it is hard to conceptualize agreement on the intensities of the vNM-utilities.

[25] The assumption that beliefs are meaningful beyond their role in the representation of individual choice also underlies the notion of "objective probabilities" on which all agents can agree, even if they behave differently.

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 space of observable contingencies, I, is describable. In addition there are unobservable or indescribable contingencies that are more relevant for one party than for the other. It seems natural that information about those contingencies is 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, and that contingencies that are foreseen by both parties are observable, and are therefore in I.

Consider a principal and an agent who want to write a contract. 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.[26] Only the principal's valuations depend on unobservable contingencies, which are unforeseen only by the agent. Let S denote the principal's taste space. The contract can fully address uncertainty about the agent's payoff, but not about the principal's payoff. Therefore, an efficient contract generically assigns some control rights to the principal: it specifies a collection of action pairs for every observable contingency $i \in I$, from which the principal can choose at a later time. Whether such a contract is considered incomplete is a definitional question.[27] The reduced form of the contract, $g : I \rightarrow \mathcal{A}$, specifies a menu of lotteries over outcomes for every contingency $i \in I$. The principal chooses from $g(i)$, after i arises and after uncertainty about the unobservable contingencies that are relevant for her taste over outcomes, $s \in S$, is resolved. To agree on an efficient contract, both parties must be able to rank all contracts.

From the principal's point of view, the contract is an act in the terminology of the previous sections. Suppose that the principal's ranking of contracts satisfies Axiom 1 and has a CPF representation based on the objective probabilities of events, ϕ. Her choice of an alternative, α, depends only on her taste, s, not on the intensity of the utility, U_s, that represents it: $\alpha_s^*(A) := \arg\max_{\alpha \in A} (\alpha \cdot s)$ is the choice under taste s.[28] The CPF representation can be written as

$$V(g) = \sum_{i \in I} \left[\int_S U_s \left(\alpha_s^* \left(g(i) \right) \right) d\mu_i(s) \right] \phi(i),$$

where μ is uniquely identified. It captures the principal's state contingent beliefs about her own tastes, which correspond to her beliefs about those relevant contingencies, that are

[26] Contingencies that impact the effect of actions on the probabilities of outcomes are considered directly relevant for both parties and are, therefore, in I.

[27] See Section 5 in Hart and Moore (1999) for a discussion.

[28] As before, $\alpha \cdot s$ denotes the dot product between lottery α and taste s. The $\arg\max$ exists, because menus are compact. If it is not unique, ties can be broken in favor of the agent.

unforeseen by the agent.

The agent assigns a contingent 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 assesses probabilities of observable contingencies according to the probability distribution ϕ. While the agent can not foresee all the contingencies underlying the formation of the principal's taste, he does know the principal's ranking of contracts, and therefore μ. Hence, the agent can rank contracts according to

$$W(g) = \sum_{i \in I} \left[\int_S \left(\alpha_s^* (g(i)) \cdot c(i) \right) d\mu_i(s) \right] \phi(i).$$

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.

In order to allow both parties to rank all contracts, the (incomplete) contracting literature has to assume that both parties believe in the same probability distribution over utility-payoffs, ex ante.[29] It follows immediately that they also know each other's ranking of contracts. The converse is not true, unless beliefs are uniquely identified from the ranking. The weaker assumption of commonly known rankings is usually justified by some informal story of learning from past observations. This assumption is not my focus, and I make it without doing the game theoretic complexity of the contracting problem justice. Instead, I address the stronger assumption of common beliefs. 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 it seems natural that each party has an informational advantage with regards to their own future taste, even once rankings are known. 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.

7. Appendix A

Subsection 1 of the Appendix specifies the indeterminacy implied by partial failures of Axiom 1. Subsection 2 provides Theorems $1'$ and $2'$, which generalize the respective theorems from the text to the case where $(I, \mathcal{F})$ is a general measurable space.

[29] Section 3 in Maskin and Tirole (1999) elaborates this point.

7.1. Partial Failures of Axiom 1

Let I be finite and 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 state-independent. To say this more precisely:

Definition 9:

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

is the *cost of choosing 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

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

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 $\widehat{\mu}_i(s) = \frac{(1+\eta c_{A,B}(s))\mu_i(s)}{\sum_{S^*}(1+\eta c_{A,B}(s))\mu_i(s)}$ and $\widehat{U}_s = \frac{U_s}{1+\eta c_{A,B}(s)}$, indexed by $\eta > -\frac{1}{c_{A,B}(s)}$.*

Another pair of menus $A', B' \in A$ satisfying i) adds additional indeterminacy if and only if, for some $s, s' \in S$,

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

Proof: 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 another parameter is required to index the 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 $\widehat{\mu}_i(s)\neq\frac{\left(1+\eta c_{A,B}(s)\right)\mu_i(s)}{\sum_{S^*}\left(1+\eta c_{A,B}(s)\right)\mu_i(s)}$ for all η. There must be some non-constant function $c:S\rightarrow\mathbb{R}$, such that $\widehat{\mu}_i(s)\equiv\frac{(1+\eta c(s))\mu_i(s)}{\sum_{S^*}(1+\eta c(s))\mu_i(s)}$ for some $\eta>0$ and $c(s)\neq c_{A,B}(s)$. The representation of $\succ_i$ 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), $\sum_{S^*}c(s)\mu_i(s)$ must be constant across I. Hence, there is some non-constant function $\widetilde{c}:S\rightarrow\mathbb{R}$, with $\sum_{S^*}\widetilde{c}(s)\mu_i(s)=0$ for all $i\in I$. Let $\widetilde{c}^+(s)$ and $\widetilde{c}^-(s)$ be the positive and negative part of $\widetilde{c}(s)$, respectively. By Claim 2 in the proof of Theorem 1 below, it is possible to choose $\varepsilon,\alpha>0$ and , $\xi^+,\xi^-:S\rightarrow\mathbb{R}_+$, such that $\xi^+-\varepsilon=\alpha\widetilde{c}^+$ and $\xi^--\varepsilon=\alpha\widetilde{c}^-$ on S^*. 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 $A':=A_{\xi^+}$ and $B':=A_{\xi^-}$ 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. ■

7.2. Existence and Conditional Uniqueness for a Measurable Objective State Space

If the objective state space I is finite as in the body of the paper, then Axiom 1 limits the cardinality of the space of relevant tastes, S^*. In many standard models the state space is infinite, for example the objective state space in Savage (1954) and generically also the space of relevant subjective states in DLR. In most applications this is a disadvantage, as a finite subjective state space is interpretationally appealing and analytically convenient. However, for those applications where an infinite subjective state space is necessary, my results can be extended to the case when I is infinite, thereby removing the constraint on the cardinality of S^*.

Let $\mathcal{F}$ be a σ-algebra on I. In this context, let G be the collection of simple opportunity acts with typical elements g,h. A simple opportunity act is a measurable function $g:I\rightarrow\mathcal{A}$, such that there exists a finite and measurable partition $\{D_t\,|t\in\{1,..,T\}\}$ of I with $g(i)=g(j)$ if there is $D\in\{D_t\,|t\in\{1,..,T\}\}$ with $i,j\in D$. $\succ$ is a binary relation on $G\times G$.[30] The definition of $\succ_D$ for $D\in\mathcal{F}$ is analogous to the definition of $\succ_i$, Definition 2. Definition 4 of

[30] One could, instead, consider any measurable function $g:I\rightarrow\mathcal{A}$ as an opportunity act. Analyzing choice on this larger domain is technically more inolved. In particular, it requires a strengthening of Axiom 3 (Continuity). I focus on the smaller domain of simple opportunity acts, as it is sufficient to identify beliefs over tastes with arbitrary support.

the CPF representation remains valid, where the value function now takes the form

$$V(g) = \int_I \left[\int_S \left(\max_{\alpha \in g(D_t)} U_s(\alpha) \right) d\mu_i(s) \right] d\phi(i),$$

where ϕ is a countably additive probability measure on $(I, \mathcal{F})$ and where μ is a well defined stochastic kernel between $(I, \mathcal{F})$ and $(S, \mathcal{B})$. Definitions and results that generalize those in Section 3 are distinguished by a prime on their label.

Definition 5′: Let (ϕ, μ, U) be a CPF representation of $\succ$.
i) Let $I^* := supp(\phi)$ and let $\mathcal{F}^*$ be the σ-algebra on I^* that corresponds to $\mathcal{F}$. The *space of relevant tastes is* $S^* := \bigcup_{D \in \mathcal{F}^*} supp(\mu_D)$ with Borel σ-algebra $\mathcal{B}^*$.
ii) μ and U are *unique given* ϕ, if for any other CPF representation $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ of $\succ$ the functions $\widehat{\mu}$ and μ induce the same kernel between the measurable spaces $(I^*, \mathcal{F}^*)$ and $(S, \mathcal{B})$, and there is $a > 0$ and an integrable function $b : S^* \to \mathbb{R}$, such that $\int_{S'} \widehat{U}_s d\mu_D(s) = \int_{S'} (aU_s + b(s)) d\mu_D(s)$ for all $D \in \mathcal{F}^*$ and all $S' \in \mathcal{B}^*$.

The next definition provides a measure of how much 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 10: Given $D \in \mathcal{F}$, let $\underline{z}$ be the worst prize: $A \succcurlyeq_D \{\underline{z}\}$ for all $A \in \mathcal{A}$. 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

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

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

If $\succ$ can be represented by a CPF representation, then the prize $\underline{z}$ must exist because Z is finite and because $\succ_D$ must obviously satisfy *Monotonicity*. Note that $p_{A,B}(D) \neq 0$ implies that D is nonnull. Endow $\mathcal{A}$ with the topology generated by the Hausdorff metric

$$d_h(A, B) = \max\left\{ \max_{\alpha \in A} \min_{\beta \in B} \|\alpha - \beta\|, \max_{\beta \in B} \min_{\alpha \in A} \|\alpha - \beta\| \right\}.$$

If two sequences of menus, $\langle A_n \rangle$ and $\langle B_n \rangle$, converge to the same limit in the Hausdorff topology, then the cost of choosing from B_n instead of A_n vanishes under every event. However, the ratio of such costs may have a well defined limit.

Axiom 1′ *(Relevance and Tightness of Objective States): 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. If $p_{C_n, B_n}(D) \nrightarrow 0$, then Axiom 1 also trivially implies Axiom 1′. Thus, Axiom 1′ is only stronger than Axiom 1 for $p_{C_n, B_n}(D) \rightarrow 0$.

Theorem 1′: *If $\succ$ has the CPF representation (ϕ, μ, U), then μ and U are unique given ϕ if and only if $\succ$ satisfies Axiom 1′.*

Proof: See Appendix B.

The discussion of Theorem 1 applies here. 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, however, then making a menu less preferred by a finite amount under one taste will invariably make it worse under similar tastes (where tastes are viewed as vectors in $\mathbb{R}^k_+$,) 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 choosing 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 of Theorem 1′ is the same as for Theorem 1.

In addition to Axioms 2-6, an axiomatization of the CPF representation requires that "small" events do not matter too much for the ranking of acts.

Definition 11: For $f, g \in G$ and $D \in \mathcal{F}$, let fDg be the act, such that

$$fDg\,(i) := \begin{cases} f(i) \text{ for } i \in D \\ g\,(i) \;\text{ otherwise} \end{cases}.$$

Axiom 8 *(Event-Continuity): For any three acts $f, g, h \in G$ with $h \succ g$ and any sequence $\{D_t\}$ in $\mathcal{F}$ with $D_{t+1} \subset D_t$ and $\bigcap_t D_t = \emptyset$, there exists T, such that $h \succ fD_t g$ for all $t > T$.*

Theorem 2′: *The binary relation $\succ$ satisfies Axioms 2-6 and 8 if and only if it has a CPF representation.*

Proof: See Appendix B.

I do not provide a generalization of Theorem 3 here. It would have to be based on a theory that generalizes AA's results to the case of an infinite objective state space. Fishburn (1979) provides such a generalization in Section 13.3.

8. Appendix B

After collecting some useful properties of support functions,[31] results are established in the order they appear in the text.

8.1. Support Functions

Definition 12: Call $\sigma_A : S \to \mathbb{R}$ with $\sigma_A\,(s) := \max_{\alpha \in A} (\alpha \cdot s)$ the *support function* of A.

Support functions have the following properties:

(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 $\sigma_{(A \cup B)} = \sigma_A \vee \sigma_B$

(iv) $\sigma_A = \sigma_{conv(A)}$, where $conv\,(A)$ is the convex hull of A.

[31] The introduction of support functions to the analysis of choice over menus is a major contributuion of DLR. For a comprehensive treatment of support functions in this context, see Chatterjee and Krishna (2009).

Denote by A_σ the maximal menu supported by σ, $A_\sigma = \bigcap_{s \in S} \{\alpha \in \Delta(Z) \,|\alpha \cdot s \leq \sigma(s)\}$. Let $\overline{\mathcal{A}}$ be the collection of all convex subsets of $\Delta(Z)$. Note that $A \in \overline{\mathcal{A}}$ if and only if A is maximal with respect to some support function. 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 2: For $\varepsilon \geq 0$ 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 $\sigma_{B_\varepsilon} \equiv \varepsilon$. In particular, the degenerate menu B_0 that contains only the center of the simplex (the lottery that assigns weight $1/k$ to every prize), has support function 0. $\square$

8.2. Proof of Theorem 1

Proof of Theorem 1, i)⇒iii): Let $I^\mu \subseteq I^*$ be a largest (in terms of cardinality) subset of linearly independent elements in $\{\mu_i\}_{i \in I^*}$. Then $\#S^* \geq \#I^\mu$ must trivially hold. It has to be shown that $\#S^* = \#I^\mu$. Suppose to the contrary that $\#S^* > \#I^\mu$. The definition of S^* implies that one can find at least $\#I^\mu + 1$ distinct Borel sets with non-empty interior, $\{S_t\}_{t=1}^{\#I^\mu+1}$, such that for all $t \leq \#I^\mu + 1$ there exists $i \in I^\mu$ with $\mu_i(\mathrm{int}(S_t)) > 0$. Since μ_i can have at most countably many atoms, one can further guarantee $\mu_i(Cl(S_t) \cap Cl(S_{t'})) = 0$ for all $t, t' \leq \#I^\mu + 1$ and all $i \in I^\mu$.

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 $x_t(i) := \int_S [\xi_t(s) - \varepsilon]\, d\mu_i(s) > 0$ 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 $\alpha \cdot s \leq \varepsilon$ for all $s \in S \backslash S_t$. This implies $\max_{\alpha \in B}(\alpha \cdot s) > \varepsilon$ for all s in the non-empty interior of S_t. Hence, $\sigma_B > \sigma_\varepsilon$ must hold on int(S_t). Let $\xi_t := \sigma_B$. $\|$

For any $\vartheta \in \mathbb{R}$ and $t' \in \{1, \#I^\mu + 1\}$, the following system of $\#I^\mu + 1$ independent linear equations with variables $\{p_t\}_{t \in \{1, \#I^\mu+1\}}$ has a solution:

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

where x_t is as defined in Claim 1. Choose $\vartheta \neq 0$ such that $\sum |p_t| = 1$. The convex combination of finitely many menus is well defined, and by property (ii) in the previous sub-section, the convex combination of finitely many support functions is, too. Thus one can define two support functions

$$\xi \;:=\; \sum_{t=1}^{\#I^{\mu}+1} |p_t|\left(\mathbf{1}_{p_t>0}\xi_t + \mathbf{1}_{p_t<0}\varepsilon\right)$$
$$\sigma \;:=\; \sum_{t=1}^{\#I^{\mu}+1} |p_t|\left(\mathbf{1}_{p_t>0}\varepsilon + \mathbf{1}_{p_t<0}\xi_t\right)$$

where ξ_t and ε are as in Claim 1. Then $\sum_{t=1}^{\#I^{\mu}+1} x_t(i)\,p_t = 0$ for all $i \in I^{\mu}$ immediately implies that $A_\xi \sim_i A_\sigma$ for all $i \in I^{\mu}$. Since I^{μ} indexes a largest (in terms of cardinality) subset of linearly independent elements in $\{\mu_i\}_{i\in I^*}$, the same must hold for all $i \in I^*$. For $i \in I\backslash I^*$, $A_\xi \sim_i A_\sigma$ trivially holds. At the same time, $p_{t'} \neq 0$ implies that $A_\xi \cup A_\sigma \succ_i A_\xi$ for some $i \in I^{\mu}$, which contradicts Axiom 1.

Proof of Theorem 1, i)⇒ii):

Claim 2: Fix $\varepsilon > 0$ such that B_ε lies in the interior of the simplex. For any positive function f on S^* there is $\overline{p} > 0$ small enough, such that for any $0 < p < \overline{p}$ there is a support function ξ with $\xi - \varepsilon|_{S^*} = pf$.

Proof of Claim 2: By (iii) S^* is finite, $S^* = \{s_1, s_2, ...\}$. Consider $\{S_t\}_{t=1}^{\#S^*}$ with $S_t \subseteq S$, $s_t \in S_t$ and $s_{t'} \notin S_t$ for $t' \neq t$. Construct $\xi_t(s)$ as in Claim 1. Let $x_t := \xi_t(s_t) - \varepsilon$. Choose $\{p_t\}_{t=1}^{\#S^*}$ such that $p_t x_t \propto f(s_t)$ and $\sum p_t = 1$. Define $\overline{\xi} := \sum p_t \xi_t$. Then $\overline{\xi} - \varepsilon\big|_{S^*} \equiv \overline{p}f$ for some $\overline{p} > 0$. For $p < \overline{p}$ let $\xi := \frac{p}{\overline{p}}\overline{\xi} + \frac{(1-p)}{\overline{p}}\varepsilon$. Then $\xi - \varepsilon|_{S^*} \equiv pf$. $\|$

Suppose (ϕ, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ are two distinct CPF representations of $\succ$ with S^* and $\widehat{S^*}$ as the corresponding relevant taste spaces. Up to a constant, the vNM expected utility $U_s(p)$ can be written as $l(s)(s \cdot p)$. Then $\max_{\alpha\in A} U_s(\alpha) = l(s)\,\sigma_A(s)$. As in the text, $l(s)$ captures the "intensity" of taste s. Let $f(s) \propto \frac{1}{l(s)}$ on S^*. Analogously let $\widehat{f}(s) \propto \frac{1}{\widehat{l}(s)}$ on $\widehat{S^*}$. Claim 2 implies that, for p' and $\widehat{p}$ small enough, there are ξ' and $\widehat{\xi}$ with $\xi' - \varepsilon|_{S^*} = p'f$ and

$\left.\widehat{\xi}-\varepsilon\right|_{\widehat{S^*}}=\widehat{p}\widehat{f}$. Suppose, without loss of generality, that $\xi' \succcurlyeq_i \widehat{\xi}$. Claim 2 and continuity of the CPF representation imply that there are also $p<p'$ and ξ, such that $\left.\xi-\varepsilon\right|_{S^*}=pf$ and $\xi\sim_i\widehat{\xi}$. As in the text, let V and $\widehat{V}$ denote the value functions that correspond to (ϕ,μ,U) and $\left(\phi,\widehat{\mu},\widehat{U}\right)$, respectively. Make the following three observations:

1) Neglecting additive constants, $\xi\sim_i\widehat{\xi}$ implies

$$\sum_{S^*}\xi(s)\,l(s)\,\mu_i(s)=\sum_{S^*}\widehat{\xi}(s)\,l(s)\,\mu_i(s)$$

and consequently

$$\sum_{S^*}(\xi(s)-\varepsilon)\,l(s)\,\mu_i(s)=\sum_{S^*}\left(\widehat{\xi}(s)-\varepsilon\right)l(s)\,\mu_i(s)\,.$$

2) Because $f(s)\propto\frac{1}{l(s)}$ it must be true that $(\xi(s)-\varepsilon)\,l(s)$ is constant across S^*. Consequently,

$$\sum_{S^*}(\xi(s)-\varepsilon)\,l(s)\,\mu_i(s)=\sum_{S^*}(\xi(s)-\varepsilon)\,l(s)\,\mu_j(s)$$

for all $i,j\in I^*$.

3) An analogous argument implies that $\sum_{\widehat{S^*}}\left(\widehat{\xi}(s)-\varepsilon\right)\widehat{l}(s)\,\widehat{\mu}_i(s)=\sum_{\widehat{S^*}}\left(\widehat{\xi}(s)-\varepsilon\right)\widehat{l}(s)\,\widehat{\mu}_j(s)$ for all $i,j\in I^*$. Therefore,

$$\begin{aligned}
&\widehat{V}\left(\frac{\phi(j)}{\phi(i)+\phi(j)}g_i^{A_{\widehat{\xi}}}+\frac{\phi(i)}{\phi(i)+\phi(j)}g_j^{A_\varepsilon}\right)-\widehat{V}\left(\frac{\phi(j)}{\phi(i)+\phi(j)}g_i^{A_\varepsilon}+\frac{\phi(i)}{\phi(i)+\phi(j)}g_j^{A_{\widehat{\xi}}}\right)\\
=&\frac{1}{\phi(i)+\phi(j)}\left[\phi(j)\,\phi(i)\sum_{\widehat{S^*}}\left(\widehat{\xi}(s)-\varepsilon\right)\widehat{l}(s)\,\widehat{\mu}_i(s)-\phi(i)\,\phi(j)\sum_{\widehat{S^*}}\left(\widehat{\xi}(s)-\varepsilon\right)\widehat{l}(s)\,\widehat{\mu}_j(s)\right]\\
=&\frac{\phi(j)\,\phi(i)}{\phi(i)+\phi(j)}\left[\sum_{\widehat{S^*}}\left(\widehat{\xi}(s)-\varepsilon\right)\widehat{l}(s)\,\widehat{\mu}_i(s)-\sum_{\widehat{S^*}}\left(\widehat{\xi}(s)-\varepsilon\right)\widehat{l}(s)\,\widehat{\mu}_j(s)\right]=0
\end{aligned}$$

and hence

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

for all $i,j\in I^*$. Since the measure ϕ on I is the same in both representations, evaluating this indifference in terms of the value function V implies

$$\sum_{S^*}\left(\widehat{\xi}(s)-\varepsilon\right)l(s)\,\mu_i(s)=\sum_{S^*}\left(\widehat{\xi}(s)-\varepsilon\right)l(s)\,\mu_j(s)$$

for all $i, j \in I^*$.

Combining observations 1-3,

$$\sum_{S^*} (\xi(s) - \varepsilon)\, l(s)\, \mu_j(s) = \sum_{S^*} \left(\widehat{\xi}(s) - \varepsilon\right) l(s)\, \mu_j(s)$$

for all $j \in I^*$ and hence, $\widehat{\xi} \sim_j \xi$ for all $j \in I^*$. At the same time, (ϕ, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ are distinct. DLR establish that $\widehat{\mu}_i(s)\, \widehat{l}(s) \propto \mu_i(s)\, l(s)$ for the case of a finite taste space (their Theorem 1). This implies that $\widehat{p}\widehat{f}(s)$ is not identical to $pf(s)$ on S^* or on $\widehat{S^*}$. Without loss of generality suppose they disagree on S^*. Then, $\left(\widehat{\xi}(s) - \varepsilon\right) l(s)$ is not constant across S^*. Because $\widehat{\xi} \sim_j \xi$, there must be $s', s'' \in S^*$ with $\widehat{p}\widehat{f}(s') > pf(s')$ and $\widehat{p}\widehat{f}(s'') < pf(s'')$. Hence, $A_{\widehat{\xi}} \cup A_\xi \succ_j A_\xi$ for all $j \in I$ with $\mu_j(s') > 0$. This contradicts Axiom 1. Therefore, $S^* = \widehat{S^*}$ and $l(s) \propto \widehat{l}(s)$ on S^*. This establishes that there are $a > 0$ and $\{b_s\}_{s \in S^*} \subset \mathbb{R}$, such that $\widehat{U}_s = aU_s + b_s$.

That the measure μ_i is unique for all $i \in I$ with $\phi(i) > 0$ then follows immediately from $\widehat{\mu}_i(s)\, \widehat{l}(s) \propto \mu_i(s)\, l(s)$.

Proof of Theorem 1, ii)⇒i): It has to be established 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 $A, B \in \mathcal{A}$, such that $A \sim_j B$ for all $j \in I$ and $A \cup B \succ_i B$ for some $i \in I$. $A \sim_j B$ for all $j \in I$ implies $\sum_{S^*} c_{A,B}(s)\, \mu_j(s) = 0$ for all $j \in I$ and for $c_{A,B}(s)$ as defined in Definition 9. $A \cup B \succ_i B$ 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 $\widehat{\mu}(s\,|i) := (1 + \eta c_{A,B}(s))\, \mu_i(s)$, where $\eta \neq 0$ is small enough, such that $1 + \eta c_{A,B}(s) > 0$ for all $s \in S^*$. Accordingly define $\widehat{l}(s) := \frac{l(s)}{1+\eta c_{A,B}(s)}$. The value function $\widehat{V}$ that corresponds to $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ is numerically identical to V, and therefore $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ also represents $\succ$. This contradicts the uniqueness statement in Theorem 1 ii). Thus, Axiom 1 is necessary for this uniqueness statement.

Proof of Theorem 1, iii)⇒ii): Suppose (ϕ, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ both represent $\succ$. Given the finiteness of S^*, the argument from the third example in Section 2 trivially generalizes to imply that $\frac{\mu_i(s)}{\mu_j(s)} = \frac{\widehat{\mu}_i(s)}{\widehat{\mu}_j(s)}$ for all $i, j \in I^*$ and $s \in S^*$. Hence, if, for $i \in I^*$ fixed, $\mu_i(s) = \widehat{\mu}_i(s)$ for all $s \in S^*$, then $\mu_j(s) = \widehat{\mu}_j(s)$ for all $s \in S^*$ and all $j \in I^*$. That is, uniqueness of μ requires the determination of $\#S^*$ variables. By (iii) the normalizations

$$\sum_{s \in S^*} \mu_j(s) = 1 \text{ for all } j \in I^*$$

provide a system of $\#S^*$ linearly independent equations. Given the uniqueness of μ, the uniqueness of U up to a common rescaling and the addition of constants follows from the identification result in DLR. ■

8.3. Proof of Theorem 2

Definition 13: As in the proof of Theorem 1, 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 valued act.

The proof proceeds to establish that for every act g there is a convex valued $\overline{g}$, such that $g(i) \sim_i \overline{g}(i)$ for all $i \in I$ and thus, by Independence, $g \sim \overline{g}$. Additive separability across I is established for convex valued acts. Thus, the act g can be evaluated by finding the act $\overline{g}$ and calculating its value state by state. Finally, Theorem DLRS provides a representation of $\succ_i$ that allows replacing the value of $\overline{g}(i)$ with the subjective expectation of the value of $g(i)$.

Lemma 3: *$\succ$ constrained to $\overline{G}$ satisfies Axioms 2-4 if and only if there is a family of continuous linear functions $\{v_i\}_{i \in I}$, $v_i : \overline{\mathcal{A}} \to \mathbb{R}$, such that $v : \overline{G} \to \mathbb{R}$ with $v(g) = \sum_{i \in I} v_i(g(i))$, represents $\succ$ on $\overline{G}$.*

Moreover, if there is another family of continuous linear functions $\{v'_i\}_{i \in I}$, $v'_i : \overline{\mathcal{A}} \to \mathbb{R}$, such that $v'(g) = \sum_{i \in I} v'_i(g(i))$ represents $\succ$ on $\overline{G}$, then there are constants $a > 0$ and $\{b_i \,|i \in I\}$, such that $v'_i = b_i + a v_i$ for each $i \in I$.

Proof: The collection of convex valued acts $\overline{G}$ together with the convex combination of acts as a mixture operation is a mixture space. Lemma 3 is an application of the Mixture Space Theorem (Theorem 5.11 in Kreps (1988)), where additive separability across I follows from the usual induction argument and the continuity of v_i is a consequence of Axiom 2. □

According to Theorem DLRS, $\succ_i$ can be represented by

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

for all $i \in I^*$, where $U_{s,i}$ is a vNM utility function that represents taste s, that is, for any $i \in I^*$ there is a $\widehat{\mu}_i(s)$-measurable function $\lambda_i : S \to \mathbb{R}_+$, such that $\max_{\alpha \in A} U_{s,i}(\alpha) = \lambda_i(s)\, \sigma_A(s)$.

Defining $\mu_i(s) := \frac{\lambda_i(s)\widehat{\mu}_i(s)}{\int_S \lambda_i(s) d\widehat{\mu}_i(s)}$ allows $\succ_i$ to be represented by

$$V_i(A) = \int_S \sigma_A(s)\, d\mu_i(s).$$

Corollary 1: If $i \in I$ is nonnull, 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 $\mathcal{A}$. $\square$

Consequently, there is an event dependent, positive scaling factor $\pi(i)$, such that, up to a constant, $v_i(A) = \pi(i) V_i(A)$ for all $A \in \overline{\mathcal{A}}$, where $\pi(i) = 0$ if and only if $i \notin I^*$. For every $g \in G$, define $\overline{g} \in \overline{G}$, such that $\overline{g}(i) := conv(g(i))$ for all $i \in I^*$. Property (iv) of support functions (Appendix B.1) implies $V_i(\overline{g}(i)) = V_i(g(i))$ for all $i \in I^*$. Independence immediately implies that $\overline{g} \sim g$. Let V' represent $\succ$ on G and $V' \equiv v$ on $\overline{G}$. Then, $V'(g) = V'(\overline{g}) = \sum_{i \in I} v_i(\overline{g}(i)) = \sum_{i \in I} \pi(i) V_i(\overline{g}(i)) = \sum_{i \in I} \pi(i) V_i(g(i))$. Hence, $g \succ h$ if and only if $\sum_{i \in I} \pi(i) V_i(g(i)) > \sum_{i \in I} \pi(i) V_i(h(i))$. Therefore,

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

represents $\succ$. Since v is unique only up to positive affine transformations, $\pi(i)$ can be normalized to be a probability measure, $\phi(i)$. This establishes the sufficiency statement in Theorem 2. In this particular CPF representation, the non-uniqueness of the representation has been exploited to normalize the state independent utilities, U_s, as suggested in DLRS.

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

8.4. Proof of Theorem 3 (Existence)

If $\succ$ has a CPF representation, then Axiom 7 implies that there is no preference for flexibility on $\Delta(X^*)$. That is, $A \succ_i B$ implies $A \sim_i A \cup B$ for all $A, B \in \Delta(X^*)$ and for all $i \in I$. To see this, suppose to the contrary that there was preference for flexibility on $\Delta(X^*)$, that is, there are menus $A, B \subset \Delta(X^*)$ with $A \cup B \succ_i A$ and $A \sim_i B$ for some $i \in I$. But then Axiom 1 would imply that there exists $j \in I$, such that $A \nsim_j B$, which contradicts Axiom 7.

Consider the CPF representation from Theorem 2:

$$\widehat{V}(g) = \sum_{i \in I} \widehat{\phi}(i) \left[\int_S \left(\max_{\alpha \in g(i)} \widehat{U}_s(\alpha) \right) d\widehat{\mu}_i(s) \right].$$

Fix $s' \in S^*$. The fact that there is no preference for flexibility on $\Delta(X^*)$ implies that for any $s \in S^*$ there is $\lambda(s)$, such that $\widehat{U}_s(x) = \lambda(s)\widehat{U}_{s'}(x)$ on X^*, as otherwise one could easily construct $A, B \subset \Delta(X^*)$ with $A \cup B \succ A$. Let $U_s(.) := \frac{\widehat{U}_s(.)}{\lambda(s)}$ to ensure that indeed $U_s(x)$ is constant across S for all $x \in X^*$. Finally, let $\mu_i(s) := \frac{\lambda(s)\widehat{\mu}_i(s)}{\int_S \lambda(s) d\widehat{\mu}_i(s)}$ and $\phi(i) := \frac{\widehat{\phi}(i) \int_S \lambda(s) d\widehat{\mu}_i(s)}{\sum_{i \in I} \widehat{\phi}(i) \left[\int_S \lambda(s) d\widehat{\mu}_i(s) \right]}$ to ensure that

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

agrees with $\widehat{V}$ up to a positive affine transformation. ■

8.5. 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, I first establish the analogous claim to Claim 2. The definition of support functions (Definition 12) and all related notations remain relevant here. For notational convenience I omit the dependence of functions from S to $\mathbb{R}$ on $s \in S$, when there is no risk of confusion.

Consider the uninformative event $I \in \mathcal{F}$. Note that $\nu(S') := \int_{S'} l d\mu_I(s)$ is itself a positive measure:[32] It exists for any measurable $S' \subset S$, as it is bounded above by $\int_S l d\mu_I(s)$, which is finite because the value of the menu supported by σ_ε in Lemma 2

[32] If information is ignored, in the sense that DM only gets to choose between degenerate acts that do not condition on information, then preferences can be represented as in DLRS. 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 ν with respect to $\mu(.|I)$ evaluated in s is $l(s)$, the intensity of taste s.

is $\int_S \sigma_\varepsilon l d\mu_I(s) = \varepsilon \int_S l d\mu_I(s)$. It is positive, because the intensity of tastes $l : S \to \mathbb{R}^+$ is a strictly positive function.[33]

Lemma 4: There are support functions ξ and σ and a number $\alpha > 0$, such that $\mu_I(S') - \int_{S'} \alpha(\xi - \sigma) l d\mu_I(s) < \varepsilon$. For $\alpha' > \alpha$ there are also support functions ξ' and σ', such that $\mu_I(S') - \int_{S'} \alpha'(\xi' - \sigma') l d\mu_I(s) < \varepsilon$.

Proof of Lemma 4:

Claim 3: If f is positive and integrable under ν, then for any $\varepsilon > 0$, there is a continuous, bounded, positive function $g : S \to \mathbb{R}$, such that $\int_S |f - g| d\nu(s) < \varepsilon$.

Proof: As f and ν are both weakly positive, $\int_S |f\nu| ds$ exists. Thus, for every $\varepsilon > 0$, there exists a continuous function $g : S \to \mathbb{R}$ such that $\int_S |g - f| d\nu(s) < \varepsilon$. See, for example, Billingsley (1995), Theorem 17.1. Since f is positive, g can be chosen to be positive.‖

Note that $\frac{1}{l} : S \to \mathbb{R}_+$ is strictly positive, because l is. It is integrable under ν, because $\int_S \frac{1}{l} d\nu(s) = \int_S d\mu_I(s) = 1$. Given $\varepsilon > 0$, Claim 3 implies that there is a continuous, bounded, positive function g, such that

$$\int_{S'} \left| \frac{1}{l} - g \right| d\nu(s) < \frac{\varepsilon}{2}.$$

Claim 4 (Lemma 1.7.9. in Schneider (1993))**:** 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 $\mathbb{R}^k$.

Claim 4 implies that for every $\varepsilon > 0$ there are two support functions ξ and σ and a number $\alpha > 0$, such that

$$\int_{S'} |g - \alpha(\xi - \sigma)| d\nu(s) < \frac{\varepsilon}{2}$$

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

Hence,

$$\mu_I(S') - \int_{S'} \alpha(\xi - \sigma) l d\mu_I(s) \leq \int_{S'} \left| \frac{1}{l} - \alpha(\xi - \sigma) \right| l d\mu_I(s)$$

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

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

This establishes the first part of the lemma. To show the second part, consider $\alpha' = c\alpha$ with $c > 1$. 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 4. $\square$

Suppose (ϕ, μ, U) is a CPF representations of $\succ$. Following Lemma 4, one can define a sequence of support functions $\langle \xi_n \rangle$ and $\langle \sigma_n \rangle$ and a sequence of numbers $\langle \alpha_n \rangle$ with $\alpha_n \to \infty$, such that

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

for every measurable set $S' \subseteq S$ and for all $n > 0$. In particular, $\int_S (\xi_n - \sigma_n) \, l d\mu_I(s) \to 0$.

Now consider another CPF representation of $\succ$, $\left(\phi, \widehat{\mu}, \widehat{U}\right)$, and define corresponding sequences $\left\langle \widehat{\xi}_n \right\rangle$, $\langle \widehat{\sigma}_n \rangle$ and $\langle \widehat{\alpha}_n \rangle$. Obviously, also $\int_S \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) \widehat{l} d\widehat{\mu}_I(s) \to 0$. It follows immediately from the uniqueness statements in Theorems 3 and 4 in DLR, that μ_D and $\widehat{\mu}_D$ share the same support, and that $l(s)\,\mu_D(s)$ differs from $\widehat{l}(s)\, d\widehat{\mu}_D(s)$ at most by scaling for any $D \in \mathcal{F}$. In particular, $l(s)\,\mu_I(s) \propto \widehat{l}(s)\, d\widehat{\mu}_D(s)$, and, therefore, $\int_S \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) l d\mu_I(s) \to 0$. Continuity of the integral implies that it is possible to choose$\langle \xi_n \rangle$, $\langle \sigma_n \rangle$, $\left\langle \widehat{\xi}_n \right\rangle$ and $\langle \widehat{\sigma}_n \rangle$ such that

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

for all $n > 0$, 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 10 in terms of support functions yields $p_{A,B}(D) \propto \int_S (\sigma_A - \sigma_B) \, l d\mu_D(s)$. 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

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

By definition, $\mu_I(S') - \alpha_n \int_{S'} (\xi_n - \sigma_n)\, l d\mu_I(s) < \frac{1}{n}$ for every measurable set $S' \subseteq S$ and for all $n > 0$ implies that (i) $\lim_{n\to\infty} \left[\alpha_n \int_S (\xi_n - \sigma_n)\, l d\mu_I(s)\right] = 1$, because μ is a probability measure and (ii) $\alpha_n (\xi_n - \sigma_n)\, l \to 1$ almost everywhere according to $\mu_I(s)$. 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 measure μ_D is dominated by μ_I and $S' \subseteq S$ is μ_D measurable if and only if it is μ_I measurable. Hence,

$$\lim_{n\to\infty} \left[\alpha_n \int_S (\xi_n - \sigma_n)\, l d\mu_D(s)\right] = 1$$

for all $D \in \mathcal{F}$. Analogously

$$\lim_{n\to\infty} \left[\widehat{\alpha}_n \int_S \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) \widehat{l} d\widehat{\mu}_D(s)\right] = 1$$

for all $D \in \mathcal{F}$. As in the case of finite I it is easy to verify that the fact that the limits are independent of D is meaningful in terms of $\succ$.[34] That is, since (ϕ, μ, U) and $\left(\phi, \widehat{\mu}, \widehat{U}\right)$ both represent $\succ$, there is also a sequence of numbers $\langle \beta_n \rangle$, such that $\lim_{n\to\infty} \left[\beta_n \int_S \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) l d\mu_D(s)\right] = 1$ 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 observation (ii) above this implies that $\frac{\int_S (\xi_n - \sigma_n) l d\mu_D(s)}{\int_S \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) l d\mu_D(s)} \to 1$ 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 $\frac{p_{C_n, A_n \cup B_n}(I)}{p_{C_n, B_n}(I)} \nrightarrow 1$.

Proof: First, note that

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

Second, note that, $\lim_{n\to\infty} \left[\widehat{\alpha}_n \int_{S'} \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) l d\mu_I(s)\right] = \int_{S'} \frac{l}{\widehat{l}} d\mu_I(s)$. Hence, on the one hand, $\lim_{n\to\infty} \left[\widehat{\alpha}_n \int_S \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) l d\mu_I(s)\right] = \int_S \frac{l}{\widehat{l}} d\mu_I(s)$, and on the other hand $\lim_{n\to\infty} \left[\alpha_n \int_S \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) l d\mu_I(s)\right] = 1$, as established in the proof of Claim 5. It follows

[34] See observation 3 in the proof that item i) of Theorem 1 implies item ii).

that $\lim_{n\to\infty} \frac{\widehat{\alpha}_n}{\alpha_n} = \int_S \frac{l}{\widehat{l}} d\mu_I(s)$. Recall that $\mu_D(s)$ and $\widehat{\mu}_D(s)$ share the same support, and that $l(s)\mu_D(s)$ differs from $\widehat{l}(s)\, d\widehat{\mu}_D(s)$ at most by scaling for any $D \in \mathcal{F}$. Therefore, 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 in the sense that there is $S' \subset S$, such that $\int_{S'} \frac{l}{\widehat{l}} d\mu_I(s) \neq \mu_I(S') \int_S \frac{l}{\widehat{l}} d\mu_I(s)$. Without loss of generality suppose that $\int_{S'} \frac{l}{\widehat{l}} d\mu_I(s) > \mu_I(S') \int_S \frac{l}{\widehat{l}} d\mu_I(s)$. Taking all this together,

$$\lim_{n\to\infty} \left[\widehat{\alpha}_n \int_{S'} \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) l d\mu_I(s)\right] = \int_{S'} \frac{l}{\widehat{l}} d\mu_I(s) > \mu_I(S') \int_S \frac{l}{\widehat{l}} d\mu_I(s) = \mu_I(S') \lim_{n\to\infty} \frac{\widehat{\alpha}_n}{\alpha_n}$$

or

$$\lim_{n\to\infty} \left[\alpha_n \int_{S'} \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) l d\mu_I(s)\right] > \mu_I(S').$$

Therefore, $\lim_{n\to\infty} \left[\alpha_n \int_S \max\left\{\xi_n - \sigma_n, \widehat{\xi}_n - \widehat{\sigma}_n\right\} l d\mu_I(s)\right] > 1$, which implies

$$\frac{\int_S \max\left\{\xi_n - \sigma_n, \widehat{\xi}_n - \widehat{\sigma}_n\right\} l d\mu_I(s)}{\int_S \left(\widehat{\xi}_n - \widehat{\sigma}_n\right) l d\mu_I(s)} \nrightarrow 1. \parallel$$

The combination of Claims 5 and 6 provides 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

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

for all $D' \in F$. $\frac{p_{C_n, A_n \cup B_n}(D)}{p_{C_n, B_n}(D)} \nrightarrow 1$ implies that there is a set $S' \subseteq S$ with $\mu_D(S') > 0$ and

$$\frac{\int\limits_{S'} c_{A_n, B_n}(s)\, d\mu_D(s)}{\int\limits_S c_{C_n, B_n}(s)\, d\mu_D(s)} \nrightarrow 0.$$

In complete analogy to the finite case, define

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

where η is small enough such that $1+\eta\frac{c_{A_n,B_n}(s)}{\int\limits_S c_{C_n,B_n}(s)\mu_D(s)}>0$ for all $s\in S$. Another CPF representation $\left(\phi,\widehat{\mu},\widehat{U}\right)$ can then be defined in complete analogy to the finite case. Thus, Axiom 1' must hold. ■

8.6. Proof of Theorem 2′:

In analogy to Definition 13, let $\overline{G}$ be the collection of all simple convex valued acts. Let $G_{\{D_t\}}\subset G$ be the collection of all acts that are measurable with respect to the partition $\{D_t\}$ of I in $\mathcal{F}$.

Lemma 3′: *$\succ$ constrained to $\overline{G}$ satisfies Axioms 2-4 if and only if there are continuous linear functions $v_D:\overline{\mathcal{A}}\to\mathbb{R}$,* indexed by $D\in\mathcal{F}$, *that satisfy*

(i) $v:\overline{G}\to\mathbb{R}$ with $v\left(g\right)=\sum\limits_{t=1}^{T}v_{D_t}\left(g\left(D_t\right)\right)$ for $g\in\overline{G}\cap G_{\{D_t\}}$, represents $\succ$

(ii) if $\{D_t\}_{t=1}^{T}$ is a partition of I, $\tau\subseteq\{1,...,T\}$, and $D=\bigcup\limits_{t\in\tau}D_t$, then $v_D\left(A\right)=\sum\limits_{t\in\tau}v_{D_t}\left(A\right)$ for all $A\in\overline{\mathcal{A}}$.

Moreover, another collection of continuous linear functions, $v'_D:\overline{\mathcal{A}}\to\mathbb{R}$, satisfies (i) and (ii) if and only if there are constants $a>0$ and a finitely additive function $b:\mathcal{F}\to\mathbb{R}$, such that $v'_D=b\left(D\right)+av_D$ for each $D\in\mathcal{F}$.

Proof: That $v\left(g\right)=\sum\limits_{t=1}^{T}v_{D_t}\left(g\left(D_t\right)\right)$ for $g\in\overline{G}\cap G_{\{D_t\}}$ represents $\succ$ confined to $\overline{G}\cap G_{\{D_t\}}$, is implied by Lemma 3. 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 $\{D'_t\}_{t=1}^{T'}$. Let $\tau\subseteq\{1,...,T'\}$ be such that $D_t=\bigcup\limits_{t\in\tau}D'_t$, and let $\sharp\tau$ be the number of elements in τ. 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 $v_{D_t}\left(A\right)=\sum\limits_{t\in\tau}v_{D'_t}\left(A\right)$, which is item ii) of the lemma. This implies that

$v(g) = \sum_{t=1}^{T} v_{D_t}(g(D_t))$ for $g \in G \cap G_{\{D_t\}}$ represents $\succ$ on G, which is item i).

The uniqueness statement follows immediately from the uniqueness in Lemma 3, where finite additivity of b is implied by property (ii). That the representation implies continuity and linearity of v and, thus, the axioms is obvious. □

As in the proof of Lemma 2, let B_0 denote the degenerate menu that contains only the center of the simplex $\Delta(Z)$. Given a collection of functions v'_D as in Lemma 3′, let $b(D) := -v'_D(B_0)$ to find a collection of functions $v_D = b(D) + v'_D$ that satisfy (i) and (ii) and $v_D(B_0) = 0$ for all D. Next I establish that Axiom 8 implies that the functions v_D are countably additive in D.

Claim 7: *Suppose $\succ$ constrained to $\overline{G}$ satisfies Axioms 2-4 and Axiom 8 and that the functions v_D with $v_D(B_0) = 0$ for all D satisfy (i) and (ii) in Lemma 3′. For a countable collection of disjoint sets in F, $\{D_t\}_{t\geq 1}$, let $D := \bigcup_{t\geq 1} D_t$. Then $v_D(A) = \sum_{t\geq 1} v_{D_t}(A)$ for all $A \in \overline{\mathcal{A}}$.*

Proof: Given a set $A \in \overline{\mathcal{A}}$, let $f(i) = A$ and $g(i) = B_0$ for all $i \in I$. Given $\varepsilon > 0$, choose an act h such that $\varepsilon > v(h) - v(g) > 0$ (this is possible by continuity of the value function.) Axiom 8 implies that for any nested sequence $\{D_t\}$ in $\mathcal{F}$ with $\bigcap D_t = \emptyset$, there exists T, such that $h \succ fD_t g$ for all $t > T$. It follows immediately from Lemma 3′, that $v_{D_t}(A) - v_{D_t}(B_0) = v_{D_t}(A) < \varepsilon$ for all $t > T$. A symmetrical argument establishes that $-v_{D_t}(A) < \varepsilon$ for all $t > T$, and hence $v_{D_t}(A) \to 0$, which implies countable additivity as claimed (see, for example, Theorem 3.1.1. in Dudley (2002)). ‖

Corollary 1 in the proof of Theorem 2 still holds, where i is replaced with D. That is, on $\overline{\mathcal{A}}$ and up to a positive affine transformation, v_D agrees with the representation V_D of $\succ_D$ as provided by Theorem DLRS,

$$V_D(A) = \int_S \sigma_A(s)\, d\mu_D(s).$$

In complete analogy to the proof of Theorem 2, it can be established that there is an event dependent, positive scaling factor $\pi(D)$ such that

$$V(g) = \sum_{t=1}^{T} \pi(D_t) \int_S \sigma_{g(D_t)}(s)\, d\mu_{D_t}(s)$$

for $g \in G_{\{D_t\}}$, where V represents $\succ$. $\pi(D) = 0$ if and only if D is trivial. For $D \in \mathcal{F}$ and $S' \in \mathcal{B}$, let

$$\eta(D \times S') := \frac{\pi(D) \int_{S'} d\mu_D(s)}{\pi(I) \int_{S} d\mu_I(s)}.$$

Claim 8: η is a countably additive probability measure on $\{D \times S' \,|D \in \mathcal{F},\ S' \in \mathcal{B}\}$.

Proof: Countable additivity with respect to $\mathcal{F}$ follows immediately from the expression for V above, when choosing g such that $\sigma_{g(D)}(s) = \varepsilon$ for all $s \in S^*$ and all $D \in \mathcal{F}$, which is possible by Lemma 2. Fixing $D \in \mathcal{F}$, countable additivity with respect to $\mathcal{B}$ follows from Theorem DLRS, which implies that μ_D is a countably additive measure (not necessarily a probability measure). $\|$

Claim 8 states that η is countably additive on $\{D \times S' \,|D \in \mathcal{F},\ S' \in \mathcal{B}\}$, which is a semi-ring of sets. It can, therefore, be extended to a countably additive probability measure on the product σ-algebra $\mathcal{F} \otimes \mathcal{B}$. See, for example, Proposition 3.2.4. in Dudley (2002). Then,

$$V(g) = \int_{I \times S} \sigma_{g(i)}(s)\, d\eta(i, s).$$

The measure $\eta(i, s)$ can be decomposed into a countably additive marginal distribution $\phi(i) := \eta(i, S)$ on I and a countably additive conditional distribution $\mu_i(s)$ on S given i. The existence of the conditional distribution, $\mu_i(s)$, is implied by, for example, Theorem 10.2.8 in Dudley (2002), after observing that $(I \times S, \mathcal{F} \otimes \mathcal{B}, \eta)$ is a probability space and S is a Polish space with the standard metric on $\mathbb{R}^k$. Theorem 10.2.1 in Dudley (2002) establishes that

$$V(g) = \int_I \int_S \sigma_{g(i)}(s)\, d\mu_i(s)\, d\phi(i)$$

as desired.[35]

This completes the proof of the sufficiency statement in Theorem 2′. That Axioms 1-6 are also necessary for the existence of the representation follows as in the case of finite I. The necessity of Axiom 8 follows immediately from the countable additivity of the measure ϕ. ■

[35] As in the case of finite I, the non-uniqueness of the representation has been exploited to normalize all utility functions as suggested in DLRS.

References

[1] Ahn, David and Todd Sarver (2011), "Preference for Flexibility and Random Choice." mimeo

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

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

[4] Chatterjee, Kaylan and R. Vijay Krishna (2009), "On Preferences with Infinitely Many Subjective States." *Economic Theory*, October 2009.

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

[6] —— (1998-b), "Standard State-Space Models Preclude Unawareness." *Econometrica*, Vol 66, No. 1, 159-173.

[7] —— (2001), "Representing Preferences with a Unique Subjective State Space." *Econometrica*, Vol 69, No. 4, 891-934.

[8] —— (2009), E. Dekel, B. L. Lipman, and A. Rustichini, "Temptation driven preferences." *Review of Economic Studies*, Vol. 76, No. 3, 937 – 971.

[9] 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.

[10] Dudley, R. (2002), *Real Analysis and Probability.* Cambridge University Press, Cambridge.

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

[12] Epstein, Larry G. and Kyoungwon Seo (2009), "Subjective States: a more Robust Model." *Games and Economic Behavior,* Vol 67, No. 2, 408-427.

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

[14] 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.

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

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

[17] Higashi, Youichiro, Kazuya Hyogo and Norio Takeoka (2009), "Subjective Random Discounting and Intertemporal Choice," *Journal of Economic Theory*, Vol. 144, 1015-1053.

[18] Karni, Edi (2008), "On Optimal Insurance in the Presence of Moral Hazard." EGRIE Keynote address, *The Geneva Risk and Insurance Review*, Vol. 33, 1-18.

[19] —— (2011a), "A Theory of Bayesian Decision Making with Action-Dependent Subjective Probabilities." *Economic Theory,* Vol. 48, No. 1, 125-146.

[20] —— (2011b), "Subjective Probabilities on a State Space." *American Economic Journal: Microeconomics*, Vol. 3, No. 4, 172-185.

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

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

[23] —— (1988), *Notes on the Theory of Choice.* Westviev Press, Boulder, CO.

[24] —— (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.

[25] Krishna, R. Vijay and Philipp Sadowski (2011), "Dynamic Preference for Flexibility." mimeo

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

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

[28] —— (1999), "Preference for Flexibility in a Savage Framework." *Econometrica*, Vol. 67, No.1, 101-119.

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

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

[31] Schneider, Rolf (1993), *Convex Bodies: the Brunn-Minkowski theory.* Cambridge University Press, Cambridge.

[32] Shenone, Pablo (2010), "Identifying subjective beliefs in subjective state space models." mimeo

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

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