ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

http://www.econometricsociety.org/

Econometrica, Vol. 78, No. 4 (July, 2010), 1285–1339

A UNIQUE COSTLY CONTEMPLATION REPRESENTATION

HALUK ERGIN
Washington University in Saint Louis, Saint Louis, MO 63130, U.S.A.

TODD SARVER
Northwestern University, Evanston, IL 60208, U.S.A.

Econometrica, Vol. 78, No. 4 (July, 2010), 1285–1339

A UNIQUE COSTLY CONTEMPLATION REPRESENTATION

BY HALUK ERGIN AND TODD SARVER[1]

We study preferences over menus which can be represented as if the individual is uncertain of her tastes, but is able to engage in costly contemplation before selecting an alternative from a menu. Since contemplation is costly, our key axiom, aversion to contingent planning, reflects the individual's preference to learn the menu from which she will be choosing prior to engaging in contemplation about her tastes for the alternatives. Our representation models contemplation strategies as subjective signals over a subjective state space. The subjectivity of the state space and the information structure in our representation makes it difficult to identify them from the preference. To overcome this issue, we show that each signal can be modeled in reduced form as a measure over ex post utility functions without reference to a state space. We show that in this reduced-form representation, the set of measures and their costs are uniquely identified. Finally, we provide a measure of comparative contemplation costs and characterize the special case of our representation where contemplation is costless.

KEYWORDS: Costly contemplation, aversion to contingent planning, subjective state space.

1. INTRODUCTION

IN MANY PROBLEMS OF INDIVIDUAL CHOICE, the decision-maker faces some uncertainty about her preferences over the available alternatives. In many cases, she may be able to improve her decision by first engaging in some form of introspection or contemplation about her preferences. However, if this contemplation is psychologically costly for the individual, then she will not wish to engage in any unnecessary contemplation. This will lead a rational individual to exhibit what we will refer to as an aversion to contingent planning.

To illustrate, consider a simple example. We will take an individual to one of two restaurants. The first one is a seafood restaurant that serves a tuna (t) and a salmon (s) dish, which we denote by $A = \{t, s\}$. The second one is a steak restaurant that serves a filet mignon (f) and a ribeye (r) dish, which we denote by $B = \{f, r\}$. We will flip a coin to determine to which restaurant to go. If it comes up heads, then we will buy the individual the meal of her choice in A, and if it comes up tails, then we will buy her the meal of her choice in B. We consider presenting the individual with one of the two following decision problems:

DECISION PROBLEM 1: We ask the individual to make a complete contingent plan listing what she would choose conditional on each outcome of the coin flip.

[1]We thank Eddie Dekel, Faruk Gul, Bart Lipman, Massimo Marinacci, numerous seminar participants, the editor, and the three anonymous referees for helpful comments and suggestions.

 DOI: 10.3982/ECTA7801

DECISION PROBLEM 2: We first flip the coin and let the individual know its outcome. She then selects the dish of her choice from the restaurant determined by the coin flip.

It is conceivable that the individual prefers facing the second decision problem rather than the first one. In this case, we say that her preferences (over decision problems) exhibit an *aversion to contingent planning* (ACP). Our explanation of ACP is that the individual finds it psychologically costly to figure out her tastes over meals. Because of this cost, she would rather not contemplate an inconsequential decision: She would rather not contemplate about her choice out of A were she to know that the coin came up tails and her actual choice set is B. In particular, she prefers to learn which choice set (A or B) is relevant before contemplating her choice.

Our main results are a representation and a uniqueness theorem for preferences over sets of lotteries. We interpret the preferences as arising from a choice situation where the individual initially chooses from among sets (or menus) of lotteries and subsequently chooses a lottery from that set. The only primitive of the model is the preference over sets of lotteries, which corresponds to the individual's choice behavior in the first period; we do not explicitly model the second-period choice out of the sets. The key axioms in our analysis are aversion to contingent planning (ACP) and *independence of degenerate decisions* (IDD). These axioms allow for costly contemplation, but impose enough structure to rule out the possibility that the individual's beliefs themselves are changing.

Before stating the ACP axiom formally, note that in our restaurant example, Decision Problem 1 corresponds to a choice out of $A \times B = \{(t, f), (t, r), (s, f), (s, r)\}$, where, for instance, (s, f) is the plan where the individual indicates that she will have the salmon dish from the seafood restaurant if the coin comes up heads and she will have the filet mignon from the steak restaurant if the coin comes up tails. Also, note that each choice of a contingent plan eventually yields a lottery over meals. For example, if the individual chooses (s, f), then she will face the lottery $\frac{1}{2}s + \frac{1}{2}f$ that yields either salmon or filet mignon, each with one-half probability. Hence, Decision Problem 1 is identical to a choice out of the set of lotteries $\frac{1}{2}A + \frac{1}{2}B = \{\frac{1}{2}t + \frac{1}{2}f, \frac{1}{2}t + \frac{1}{2}r, \frac{1}{2}s + \frac{1}{2}f, \frac{1}{2}s + \frac{1}{2}r\}$. In general, we can represent the set of contingent plans between any two menus as a convex combination of these menus, with the weight on each menu corresponding to the probability that it will be the relevant menu. The individual's preference of Decision Problem 2 to Decision Problem 1 is thus equivalent to preferring the half–half lottery over A and B (resolving prior to her choice from the menus) to the convex combination of the two menus, $\frac{1}{2}A + \frac{1}{2}B$. Although we do not analyze preferences over lotteries over menus explicitly, it is intuitive that the individual would prefer the better menu, say A, to any lottery over the two menus. Under this assumption, aversion to contingent planning implies that the individual will prefer choosing from the better of the two menus to making a contingent plan from the two menus. Our ACP axiom is precisely the formalization of this statement: If $A \succsim B$, then $A \succsim \alpha A + (1-\alpha)B$ for any $\alpha \in [0, 1]$.

To motivate our IDD axiom, consider the situation in which the individual makes a contingent choice from a menu A and with probability α her contingent choice is carried out; with probability $1 - \alpha$ she is instead given a fixed lottery p. As argued above, this choice problem corresponds to the menu $\alpha A + (1-\alpha)\{p\}$. If the probability α that her contingent choice from A will be implemented decreases, then her benefit from contemplation decreases. However, if α is held fixed, then replacing the lottery p in the convex combination $\alpha A + (1 - \alpha)\{p\}$ with another lottery q does not change the probability that the individual's contingent choice from A will be implemented. Therefore, although replacing p with q could affect the individual's utility through its effect on the final composition of lotteries, it will not affect the individual's optimal level of contemplation. This observation motivates our IDD axiom, which states that for any fixed α, if $\alpha A + (1 - \alpha)\{p\}$ is preferred to $\alpha B + (1 - \alpha)\{p\}$, then $\alpha A + (1 - \alpha)\{q\}$ is also preferred to $\alpha B + (1 - \alpha)\{q\}$.

We present our model in detail in Section 2. Section 2.1 contains a detailed description of our axioms. Along with ACP and IDD, we consider three standard axioms in the setting of preferences over menus: (i) weak order, which states that the preference is complete and transitive, (ii) continuity, and (iii) monotonicity, which states that adding alternatives to any menu is always (weakly) better for the individual.

Our representation theorem is contained in Section 2.2. Letting p denote a lottery over some set of alternatives Z and letting A denote a menu of such lotteries, Theorem 1 shows that any preference over menus satisfying our axioms can be represented by the costly contemplation (CC) representation

$$V(A) = \max_{\mathcal{G}\in\mathbf{G}}\Big(\mathbb{E}\Big[\max_{p\in A}\mathbb{E}[U|\mathcal{G}]\cdot p\Big] - c(\mathcal{G})\Big). \tag{1}$$

We interpret Equation (1) as follows. The individual is uncertain regarding her tastes over alternatives in Z. This uncertainty is modeled by a probability space $(\Omega, \mathcal{F}, P)$ and a state-dependent expected-utility function U over $\triangle(Z)$. Before making a choice out of a menu A, the individual is able to engage in contemplation so as to resolve some of this uncertainty. Contemplation strategies are modeled as a collection of signals about the state or, more compactly, as a collection $\mathbf{G}$ of σ-algebras generated by these signals. If the individual carries out the contemplation strategy $\mathcal{G} \in \mathbf{G}$, she is able to update her expected-utility function using her information $\mathcal{G}$ and chose a lottery p in A maximizing her conditional expected utility $\mathbb{E}[U|\mathcal{G}] \cdot p = \sum_{z\in Z} p_z\mathbb{E}[U_z|\mathcal{G}]$. Faced with the menu A, the individual chooses her contemplation strategy optimally by maximizing the ex ante value $\mathbb{E}[\max_{p\in A}\mathbb{E}[U|\mathcal{G}]\cdot p]$ minus the cost $c(\mathcal{G})$ of contemplation, giving Equation (1). Note that this representation closely resembles a standard costly information acquisition problem. The difference is that the parameters $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$ of the CC representation are subjective in the sense that they are not directly observable, but instead must be elicited from the individual's preferences.

In Section 3, we discuss the extent to which we are able identify contemplation strategies and their costs from the preference. Due to the subjectivity of the state space and information structure in the CC representation, it is not possible to pin down the parameters of the representation from the preference. After providing an example to illustrate the nonuniqueness of the CC representation, we show that contemplation strategies can be uniquely identified when they are put into a reduced form using measures over ex post utility functions. To motivate this reduced form, suppose the individual selects a signal (i.e., contemplation strategy), the realization of which gives her some information about her tastes for the different alternatives in a menu. The information contained in a realization of the signal results in some ex post utility function, and hence the distribution of the signal translates into a distribution over ex post utility functions. Following this approach of transforming contemplation strategies into measures, Theorem 2 shows that any CC representation is equivalent to the reduced-form costly contemplation (RFCC) representation[2]

$$V(A) = \max_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p)\mu(du) - c(\mu) \right). \tag{2}$$

The set $\mathcal{U}$ is a collection of ex post expected-utility functions, and $\mathcal{M}$ is a set of measures on $\mathcal{U}$.[3] Each measure $\mu \in \mathcal{M}$ determines a particular weighting of the ex post utility functions. It is important to note that we do not require the measures in $\mathcal{M}$ to be probability measures. Although such a requirement seems natural given our motivation for the representation, to identify the parameters in the RFCC representation, we will impose a normalization on the utility functions in $\mathcal{U}$. Under this normalization, the measures in $\mathcal{M}$ are used to capture both the likelihood of an ex post utility function and the "magnitude" of that utility function, which requires the use of measures that are not probabilities. When costly contemplation is modeled in this reduced form, parameters can be uniquely identified from the preference. Theorem 4 establishes the

[2]This representation bears some similarity to the representation for "variational preferences" considered by Maccheroni, Marinacci, and Rustichini (2006) in the Anscombe–Aumann setting. There is also a technical connection between the two representations since we apply similar results from convex analysis to establish our representation theorems, although the setting of our model requires us to develop a stronger version of these results in Section S.1 of the Supplemental Material (Ergin and Sarver (2010)).

[3]The RFCC representation also imposes the following consistency condition on the measures in $\mathcal{M}$: For every $\mu, \nu \in \mathcal{M}$ and every lottery p, $\int_{\mathcal{U}} u(p)\mu(du) = \int_{\mathcal{U}} u(p)\nu(du)$. This condition implies that even though the individual's tastes after contemplation can be very different from her tastes before contemplation, the contemplation process should not affect the individual's tendencies on average.

uniqueness of the set of measures $\mathcal{M}$ and cost function c in an RFCC representation.

The uniqueness of the parameters in the RFCC representation makes it possible to conduct meaningful comparisons of contemplation costs between two representations. In Section 4, we introduce a measure of comparative contemplation costs and show the implications for our RFCC representation.

In Section 5.1, we introduce a variation of our model in which the individual has limited resources to devote to contemplation. That is, the cost of contemplation does not directly affect the utility of the individual, but instead enters indirectly by being constrained to be below some fixed upper bound. We show that such a representation is in fact a special case of our model, and we introduce the additional axiom needed to obtain this representation for limited contemplation resources.

Our work relates to several other papers in the literature on preferences over menus. This literature originated with Kreps (1979), who considered preferences over menus taken from a finite set of alternatives. Dekel, Lipman, and Rustichini (2001) (henceforth DLR) extended Kreps' analysis to the current setting of preferences over menus of lotteries and used the additional structure of this domain to obtain an essentially unique representation. In Section 5.2, we discuss a version of the independence axiom for preferences over menus of lotteries which was used by DLR in one of their representation results. We illustrate how our axioms relax the independence axiom and why such a relaxation of independence is necessary to model costly contemplation.

A model of costly contemplation was also considered by Ergin (2003), whose primitive was the same as that of Kreps (1979)—a preference over menus taken from a finite set of alternatives. The costly contemplation representation in Equation (1) is similar to the functional form of his representation. However, the parameters in Ergin's (2003) representation are not pinned down by the preference. The richer domain of our preferences, menus of lotteries, combined with the reduced form of our RFCC representation enables us to uniquely identify the parameters of our representation. Moreover, our richer domain yields additional behavioral implications of costly contemplation, such as ACP and IDD.

We conclude in Section 6 with a brief overview of the so-called *infinite-regress issue* for models of costly decision-making. We discuss how our model relates to the issue and explain how our representation result provides an *as if* solution to the issue. Unless otherwise indicated, all proofs are contained in the Appendix.

2. A MODEL OF COSTLY CONTEMPLATION

Let Z be a finite set of alternatives, and let $\triangle(Z)$ denote the set of all probability distributions on Z, endowed with the Euclidean metric d.[4] Let $\mathcal{A}$ denote

[4]Since Z is finite, the topology generated by d is equivalent to the weak* topology on $\triangle(Z)$.

the set of all closed subsets of $\triangle(Z)$, endowed with the Hausdorff metric, which is defined by

$$d_h(A,B)=\max\Big\{\max_{p\in A}\min_{q\in B}d(p,q),\max_{q\in B}\min_{p\in A}d(p,q)\Big\}.$$

Elements of $\mathcal{A}$ are called menus or option sets. The primitive of our model is a binary relation $\succsim$ on $\mathcal{A}$, representing the individual's preferences over menus. We maintain the interpretation that after committing to a particular menu A, the individual chooses a lottery out of A in an unmodeled second stage.

For any $A, B\in\mathcal{A}$ and $\alpha\in[0,1]$, define the convex combination of these two menus by $\alpha A+(1-\alpha)B\equiv\{\alpha p+(1-\alpha)q: p\in A \text{ and } q\in B\}$. Let $\mathrm{co}(A)$ denote the convex hull of the set A.

2.1. *Axioms*

We impose the following order and continuity axioms.

AXIOM 1—Weak Order: *$\succsim$ is complete and transitive.*

AXIOM 2—Strong Continuity:
(i) Continuity. *For all $A\in\mathcal{A}$, the sets $\{B\in\mathcal{A}: B\succsim A\}$ and $\{B\in\mathcal{A}: B\precsim A\}$ are closed.*
(ii) L-continuity. *There exist $p^*, p_*\in\triangle(Z)$ and $M>0$ such that for every $A,B\in\mathcal{A}$ and $\alpha\in(0,1)$ with $\alpha>Md_h(A,B)$,*

$$(1-\alpha)A+\alpha\{p^*\}\succ(1-\alpha)B+\alpha\{p_*\}.$$

The weak order axiom is entirely standard, as is the first part of the strong continuity axiom. The added assumption of L-continuity is used to obtain Lipschitz continuity of our representation in much the same way that the continuity axiom is used to obtain continuity.[5] To interpret L-continuity, first note that $\{p^*\}\succ\{p_*\}$.[6] For any $A,B\in\mathcal{A}$, continuity therefore implies that there exists $\alpha\in(0,1)$ such that $(1-\alpha)A+\alpha\{p^*\}\succ(1-\alpha)B+\alpha\{p_*\}$. L-continuity implies that such a preference holds for any $\alpha>Md_h(A,B)$, so as A and B get closer, the minimum required weight on p^* and p_* converges to 0 at a smooth rate. The constant M can be thought of as the sensitivity of this minimum α to the distance between A and B.

The next axiom captures an important aspect of our model of costly contemplation:

[5]Similar L-continuity axioms are used in Dekel, Lipman, Rustichini, and Sarver (2007) (henceforth DLRS) and Sarver (2008). There is also a connection between our L-continuity axiom and the properness condition proposed by Mas-Colell (1986).

[6]Let $\alpha=\frac{1}{2}$. Applying L-continuity with $A=B=\{p^*\}$ implies $\{p^*\}\succ\{\frac{1}{2}p^*+\frac{1}{2}p_*\}$, and applying L-continuity with $A=B=\{p_*\}$ implies $\{\frac{1}{2}p^*+\frac{1}{2}p_*\}\succ\{p_*\}$.

AXIOM 3—Aversion to Contingent Planning: *For any* $\alpha \in [0, 1]$,

$$A \succsim B \implies A \succsim \alpha A + (1-\alpha)B.$$

To interpret ACP, suppose we were to extend the individual's preferences to lotteries over menus. Let $\alpha \circ A \oplus (1-\alpha) \circ B$ denote the lottery that yields the menu A with probability α and the menu B with probability $1-\alpha$. We interpret this lottery as resolving prior to the individual making her choice of alternative from the menus. If instead the individual is asked to make her decision prior to the resolution of the lottery, then she must make a contingent choice. The situation in which the individual makes a contingent choice, p if A and q if B, prior to the resolution of the lottery over menus is equivalent to choosing the alternative $\alpha p + (1-\alpha)q \in \alpha A + (1-\alpha)B$. Thus, any contingent choice from A and B corresponds to a unique lottery in $\alpha A + (1-\alpha)B$. As discussed in the Introduction, if contemplation is costly for the individual, then she will prefer that a lottery over menus is resolved prior to her choosing an alternative so that she can avoid contingent planning. Hence,

$$\alpha \circ A \oplus (1-\alpha) \circ B \succsim \alpha A + (1-\alpha)B. \tag{3}$$

If in addition this extended preference satisfies stochastic dominance, then $A \succsim B$ implies $A \succsim \alpha \circ A \oplus (1-\alpha) \circ B$. Together with Equation (3), this implies ACP.[7]

The following axiom allows for the possibility that the individual contemplates to obtain information about her ex post utility, but it rules out the possibility that she changes her beliefs by becoming more optimistic about the utility she will obtain from a given lottery.

AXIOM 4—Independence of Degenerate Decisions: *For any* $A, B \in \mathcal{A}$, $p, q \in \triangle(Z)$, *and* $\alpha \in [0, 1]$,

$$\alpha A + (1-\alpha)\{p\} \succsim \alpha B + (1-\alpha)\{p\}$$
$$\implies \alpha A + (1-\alpha)\{q\} \succsim \alpha B + (1-\alpha)\{q\}.$$

Suppose the individual is asked to make a contingent plan, and she is told that she will be choosing from the menu A with probability α and from the menu $\{p\}$ with probability $1-\alpha$. We refer to a choice from the singleton menu $\{p\}$ as a *degenerate decision*. When faced with a degenerate decision, there is no benefit or loss to the individual from contemplating. Therefore, if the probability α that her contingent choice from A will be implemented decreases, then her benefit from contemplation decreases. Hence, we should expect the

[7]All the results in the text continue to hold if one replaces ACP with the following weaker condition: $A \sim B \implies A \succsim \alpha A + (1-\alpha)B \; \forall \alpha \in [0, 1]$.

individual to choose a less costly level of contemplation as α decreases. However, if α is held fixed, then replacing the degenerate decision $\{p\}$ in the convex combination $\alpha A + (1-\alpha)\{p\}$ with another degenerate decision $\{q\}$ does not change the probability that the individual's contingent choice from A will be implemented. Therefore, although replacing p with q could affect the individual's utility through its effect on the final composition of lotteries, it will not affect the individual's optimal level of contemplation.[8]

Our last axiom is a monotonicity axiom introduced by Kreps (1979).

AXIOM 5—Monotonicity: *If* $A \subset B$, *then* $B \succsim A$.

If additional alternatives are added to a menu A, the individual can always "ignore" these new alternatives and engage in the same contemplation as with the menu A.[9] Therefore, the utility from a menu $B \supset A$ must be at least as great as the utility from the menu A. Although at first glance it may seem that costly contemplation alone could lead to a preference for smaller menus to avoid "overanalyzing" the decision, this argument overlooks the fact that the individual chooses her contemplation strategy optimally and, in particular, can ignore any options.

The possibility of overanalysis could arise if the individual experiences some disutility from not selecting the ex post optimal choice from a menu, for example, because of regret. Therefore, regret could lead the individual to sometimes prefer a smaller menu, which we refer to as a *preference for commitment*. Other factors, such as temptation, could also lead to a preference for commitment. Regret is studied in a related framework by Sarver (2008), and temptation is studied by Gul and Pesendorfer (2001) and Dekel, Lipman, and Rustichini (2008). We leave the study of how to incorporate regret or temptation into our model of costly contemplation as an open question for future research.

2.2. *Representation Result*

We now define our costly contemplation representation.

DEFINITION 1: A *costly contemplation* (CC) *representation* is a tuple $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$, where $(\Omega, \mathcal{F}, P)$ is a probability space, $\mathbf{G}$ is a collection of sub-σ-algebras of $\mathcal{F}$, U is a Z-dimensional, $\mathcal{F}$-measurable, and integrable random vector, and $c : \mathbf{G} \to \mathbb{R}$ is a cost function such that $V : \mathcal{A} \to \mathbb{R}$ defined by

$$V(A) = \max_{\mathcal{G} \in \mathbf{G}} \left(\mathbb{E}\left[\max_{p \in A} \mathbb{E}[U|\mathcal{G}] \cdot p \right] - c(\mathcal{G}) \right) \tag{4}$$

[8]Our IDD axiom is similar in spirit to the weak certainty independence axiom used by Maccheroni, Marinacci, and Rustichini (2006) in the Anscombe–Aumann setting. In their axiom, arbitrary acts play the role of the menus A and B, and constant acts play the role of the singleton menus $\{p\}$ and $\{q\}$.

[9]Note that we are assuming it is costless for the individual to "read" the alternatives on the menu. What is costly for the individual is analyzing her tastes for these alternatives.

represents $\succsim$, where the outer maximization in Equation (4) has a solution for every $A \in \mathcal{A}$ and there exist $p, q \in \triangle(Z)$ such that $\mathbb{E}[U] \cdot p > \mathbb{E}[U] \cdot q$.[10]

The costly contemplation representation above is a generalized version of the costly contemplation representation in Ergin (2003).[11] The interpretation of Equation (4) is as follows: The individual has a subjective state space Ω representing her tastes over alternatives, endowed with a σ-algebra $\mathcal{F}$. She does not know the realization of the subjective state $\omega \in \Omega$ but has a prior P on $(\Omega, \mathcal{F})$. Her tastes over lotteries in $\triangle(Z)$ are summarized by the random vector U representing her state-dependent expected-utility function. Her utility from a lottery $p \in \triangle(Z)$ conditional on the subjective state $\omega \in \Omega$ is therefore given by $U(\omega) \cdot p = \sum_{z \in Z} p_z U_z(\omega)$.

Before making a choice out of a menu $A \in \mathcal{A}$, the individual may engage in contemplation. A *contemplation strategy* is modeled as a signal about the subjective state, which corresponds to a sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$. The contemplation strategies available to the individual are given by the collection of σ-algebras $\mathbf{G}$. If the individual carries out the contemplation strategy $\mathcal{G}$, she incurs a psychological cost of contemplation $c(\mathcal{G})$. However, she can then condition her choice out of A on $\mathcal{G}$ and pick an alternative that yields the highest expected utility conditional on the signal realization. Faced with the menu A, the individual chooses an optimal level of contemplation by maximizing the value minus the cost of contemplation. This yields $V(A)$ in Equation (4) as the ex ante value of the option set A. The CC formulation is similar to an optimal information acquisition formula. The difference from a standard information acquisition problem is that the parameters $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$ are subjective. Therefore, they are not directly observable, but need to be derived from the individual's preference.

[10]Two notes are in order regarding this definition: (i) We show in Appendix A that the integrability of U implies that the term $\mathbb{E}[\max_{p \in A} \mathbb{E}[U|\mathcal{G}] \cdot p]$ is well defined and finite for every $A \in \mathcal{A}$ and $\mathcal{G} \in \mathbf{G}$. (ii) For simplicity, we directly assume that the outer maximization in Equation (4) has a solution instead of making topological assumptions on $\mathbf{G}$ to guarantee the existence of a maximum. An alternative approach that does not require this indirect assumption on the parameters of the representation would be to replace the outer maximization in Equation (4) with a supremum, in which case all of our results would carry over.

[11]Ergin (2003) works in the framework introduced by Kreps (1979), where the primitive of the model is a preference over subsets of Z rather than subsets of $\triangle(Z)$. He shows that a preference $\succsim$ over sets of alternatives is monotone ($A \subset B \Longrightarrow B \succsim A$) if and only if there exists a costly contemplation representation with finite Ω such that $\succsim$ is represented by the ex ante utility function V in Equation (4). The formulation of costly contemplation in this paper allows for infinite subjective state space Ω and extends the formulation to menus of lotteries assuming that the state-dependent utility is von Neumann–Morgenstern.

Finally, note that $V(\{p\}) = \mathbb{E}[U] \cdot p - \min_{\mathcal{G}\in\mathbf{G}} c(\mathcal{G})$ for any $p \in \triangle(Z)$. Therefore, the requirement in Definition 1 that $\mathbb{E}[U] \cdot p > \mathbb{E}[U] \cdot q$ for some $p, q \in \triangle(Z)$ is equivalent to assuming $V(\{p\}) > V(\{q\})$.[12]

We now present our first representation theorem.

THEOREM 1: *The preference* $\succsim$ *has a CC representation if and only if it satisfies weak order*, *strong continuity*, *ACP*, *IDD*, *and monotonicity*.

We will not provide a direct proof for this result since it follows from two results presented in the next section (Theorems 2 and 3).

3. IDENTIFYING CONTEMPLATION STRATEGIES AND COSTS

The following example shows that two different CC representations can lead to the same value function V for menus and, hence, represent the same preference.

EXAMPLE 1: Let $Z = \{z_1, z_2, z_3\}$ and $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Let $\mathcal{F}$ be the discrete algebra and let P be the uniform distribution on Ω. For each $i \in \{1, 2, 3\}$, let $\mathcal{G}_i$ be the algebra generated by the partition $\{\{\omega_i\}, \{\omega_j, \omega_k\}\}$, and let the collection of contemplation strategies be $\mathbf{G} = \{\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3\}$. Let $c : \mathbf{G} \to \mathbb{R}$ be any cost function, and define $U : \Omega \to \mathbb{R}^3$ and $\hat{U} : \Omega \to \mathbb{R}^3$ by[13]

$$U(\omega_1) = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}, \quad U(\omega_2) = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}, \quad U(\omega_3) = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix},$$
$$\hat{U}(\omega_1) = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \quad \hat{U}(\omega_2) = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad \hat{U}(\omega_3) = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

Then

$$\mathbb{E}\Big[\max_{p\in A} \mathbb{E}[U|\mathcal{G}_1] \cdot p\Big] = \frac{1}{3} \max_{p\in A} U(\omega_1) \cdot p + \frac{2}{3} \max_{p\in A} \Big[\frac{1}{2} U(\omega_2) + \frac{1}{2} U(\omega_3)\Big] \cdot p \tag{5}$$

[12]If we take p^* and p_* as in the definition of L-continuity, then $\{p^*\} \succ \{p_*\}$, which gives rise to this condition. This "singleton-nontriviality" implication of L-continuity is not accidental, as it plays an important role in the proof of our representation theorem.

[13]Under this specification of the random vectors, we have $\mathbb{E}[U] = 0$ and $\mathbb{E}[\hat{U}] = 0$, and hence the singleton-nontriviality condition in Definition 1 is not satisfied. However, we allow for this violation purely for expositional simplicity. The representations can be modified to satisfy singleton-nontriviality as follows: Add a fourth state ω_4 to Ω, let $U(\omega_4) = \hat{U}(\omega_4) = (1, 0, 0)$, and let $\mathcal{G}_i$ be the algebra generated by the partition $\{\{\omega_i\}, \{\omega_j, \omega_k\}, \{\omega_4\}\}$.

$$
\begin{aligned}
&= \max_{p \in A} \begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \end{pmatrix} \cdot p + \max_{p \in A} \begin{pmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} \cdot p \\
&= \frac{2}{3} \max_{p \in A} \left[\frac{1}{2} \hat{U}(\omega_2) + \frac{1}{2} \hat{U}(\omega_3) \right] \cdot p \\
&\quad + \frac{1}{3} \max_{p \in A} \hat{U}(\omega_1) \cdot p \\
&= \mathbb{E}\Big[\max_{p \in A} \mathbb{E}[\hat{U} | \mathcal{G}_1] \cdot p \Big].
\end{aligned}
$$

Similar arguments can be made for each algebra in $\mathcal{G}_i \in \mathbf{G}$. Therefore, defining V and $\hat{V}$ as in Equation (4) for each of the representations $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$ and $((\Omega, \mathcal{F}, P), \mathbf{G}, \hat{U}, c)$, respectively, we have $V(A) = \hat{V}(A)$ for any menu $A \in \mathcal{A}$.

Although the CC representation is not unique, we will show that the contemplation strategies in the representation can be put into a "reduced form" which will allow them to be uniquely identified by the preference. Equation (5) in Example 1 illustrates the motivation for this reduced form. First, note that for any $\mathcal{G} \in \mathbf{G}$, the distributions over ex post utility functions induced by these two representations are not the same. For instance, for $\mathcal{G}_1$, the distribution over ex post utility functions induced by the first representation puts weight $\frac{1}{3}$ on $(2, -1, -1)$ and weight $\frac{2}{3}$ on $(-1, \frac{1}{2}, \frac{1}{2})$, whereas the second representation puts weight $\frac{2}{3}$ on $(1, -\frac{1}{2}, -\frac{1}{2})$ and weight $\frac{1}{3}$ on $(-2, 1, 1)$. However, the product of these ex post utility functions with their probabilities is the same for both representations, yielding the vectors $(\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3})$ and $(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3})$.

We now generalize these observations to show that contemplation strategies can be uniquely identified when they are represented using measures over ex post expected-utility functions, where each measure captures the combination of the likelihood and the magnitude of ex post utilities for the corresponding contemplation strategy. Since expected-utility functions on $\triangle(Z)$ are equivalent to vectors in $\mathbb{R}^Z$, we will use the notation $u(p)$ and $u \cdot p$ interchangeably. Define the set of *normalized (nonconstant) expected-utility functions* on $\triangle(Z)$ to be

$$
(6) \qquad \mathcal{U} = \left\{ u \in \mathbb{R}^Z : \sum_{z \in Z} u_z = 0, \sum_{z \in Z} u_z^2 = 1 \right\}.
$$

For any $v \in \mathbb{R}^Z$ (i.e., any expected-utility function), there exist $\alpha \geq 0$, $\beta \in \mathbb{R}$, and $u \in \mathcal{U}$ such that $v(p) = \alpha u(p) + \beta$ for all $p \in \triangle(Z)$. Therefore, modulo an affine transformation, $\mathcal{U}$ contains all possible ex post expected-utility functions.

The following lemma shows that in any CC representation, each contemplation strategy corresponds to a unique measure over $\mathcal{U}$.

LEMMA 1: *Let $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$ be any costly contemplation representation. For each $\mathcal{G} \in \mathbf{G}$, there exists a unique finite Borel measure $\mu_{\mathcal{G}}$ on $\mathcal{U}$ and scalar $\beta_{\mathcal{G}} \in \mathbb{R}$ such that for all $A \in \mathcal{A}$,*

$$\mathbb{E}\Big[\max_{p\in A} \mathbb{E}[U|\mathcal{G}] \cdot p\Big] = \int_{\mathcal{U}} \max_{p\in A} u(p)\mu_{\mathcal{G}}(du) + \beta_{\mathcal{G}}.$$

In particular, it must be that $\beta_{\mathcal{G}} = \frac{1}{|Z|}\sum_{z\in Z} \mathbb{E}[U_z]$ for all $\mathcal{G} \in \mathbf{G}$.

Note that the normalization of the ex post utility functions in $\mathcal{U}$ is necessary for obtaining the uniqueness of the measure $\mu_{\mathcal{G}}$ in this result. For instance, as we illustrated in the context of Example 1, different distributions over nonnormalized ex post utility functions may correspond to the same measure μ on the set of normalized utility functions $\mathcal{U}$.[14] Note also that while the normalization of the utility functions in $\mathcal{U}$ necessitates the use of measures that may not be probabilities, we can interpret any positive measure μ on $\mathcal{U}$ as a normalized version of a distribution over ex post utility functions. Specifically, let $\lambda = \mu(\mathcal{U}) > 0$, and consider the probability measure π on $\mathcal{V} = \lambda\mathcal{U}$ which (heuristically) puts $\mu(u)/\lambda$ weight on each $v = \lambda u \in \mathcal{V}$. Then, by a simple change of variables, $\int_{\mathcal{U}} \max_{p\in A} u(p)\mu(du) = \int_{\mathcal{V}} \max_{p\in A} v(p)\pi(dv)$.

We now sketch the proof of Lemma 1 for the case of a costly contemplation representation $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$ where the state space Ω is finite (and $\mathcal{F}$ is the discrete algebra). For each event $E \subset \Omega$, one can think of $\sum_{\omega\in E} P(\omega)U(\omega)$ as an expected-utility function over $\triangle(Z)$. By the definition of $\mathcal{U}$, there exist $\alpha_E \geq 0$, $\beta_E \in \mathbb{R}$, and $u_E \in \mathcal{U}$ such that

$$\alpha_E u_E(p) + \beta_E = \left[\sum_{\omega\in E} P(\omega)U(\omega)\right] \cdot p \quad \forall p \in \triangle(Z).$$

For simplicity, suppose that $\beta_E = 0$ for each event $E \subset \Omega$.

[14]The impossibility of uniquely identifying distributions over (nonnormalized) ex post utility functions in our model is similar to the issue common to most models of state-dependent utility that the probability distribution over states cannot be identified separately from the utility function (see Karni (1993) for a more detailed discussion of state-dependent utility within the context of the Anscombe–Aumann model). For example, this observation motivated Kreps (1979) to impose the implicit normalization that the expectation of the state-dependent utility function be taken with respect to the uniform distribution. We adopt the alternative approach of normalizing utilities and using measures to represent the product of the probability and ex post utility.

Let $\mathcal{G} \in \mathbf{G}$. Finiteness of Ω implies that there is a partition $\pi_{\mathcal{G}}$ of Ω that generates $\mathcal{G}$. We define a measure $\mu_{\mathcal{G}}$ over $\mathcal{U}$ which has finite support by $\mu_{\mathcal{G}}(u) = \sum_{E\in\pi_{\mathcal{G}}:u_E=u} \alpha_E$ for each $u \in \mathcal{U}$. Note that we sum over all E for which $u_E = u$ since it is possible to have multiple elements of the partition that lead to the same ex post expected-utility preference. Then

$$\begin{aligned}
\mathbb{E}\Big[\max_{p\in A} \mathbb{E}[U|\mathcal{G}]\cdot p\Big] &= \sum_{E\in\pi_{\mathcal{G}}} P(E)\Big[\max_{p\in A}\Big[\sum_{\omega\in E} P(\omega|E)U(\omega)\Big]\cdot p\Big] \\
&= \sum_{E\in\pi_{\mathcal{G}}}\Big[\max_{p\in A}\Big[\sum_{\omega\in E} P(\omega)U(\omega)\Big]\cdot p\Big] \\
&= \sum_{E\in\pi_{\mathcal{G}}} \alpha_E \max_{p\in A} u_E(p) \\
&= \int_{\mathcal{U}} \max_{p\in A} u(p)\mu_{\mathcal{G}}(du).
\end{aligned}$$

We show in Appendix B that without the assumption that $\beta_E = 0$ for all $E \subset \Omega$, the term $\beta_{\mathcal{G}} = \sum_{E\in\pi_{\mathcal{G}}} \beta_E = \frac{1}{|Z|}\sum_{z\in Z} \mathbb{E}[U_z]$ would be added to the above expression. We also show that the uniqueness of $\mu_{\mathcal{G}}$ can be established using the uniqueness results for the additive expected-utility (EU) representation of DLR.[15]

Going back to Example 1, define $u^1, u^2, u^3 \in \mathcal{U}$ by

$$u^1 = \frac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}, \quad u^2 = \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}, \quad u^3 = \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

Then, by the same arguments as those given in Equation (5), we see that the measures induced by the partition $\{\{\omega_i\}, \{\omega_j, \omega_k\}\}$ in the two representations are identical, giving $\frac{\sqrt{6}}{3}$ weight to u^i, $\frac{\sqrt{6}}{3}$ weight to $-u^i$, and 0 weight to $\mathcal{U} \setminus \{u^i, -u^i\}$.

Motivated by the equivalence obtained in Lemma 1, we now define our reduced-form representation.[16]

DEFINITION 2: A *reduced-form costly contemplation* (RFCC) *representation* is a pair $(\mathcal{M}, c)$ consisting of a compact set of finite Borel measures $\mathcal{M}$ on $\mathcal{U}$

[15] We discuss the relationship between our model and the additive EU representation of DLR in more detail in Section 5.2.

[16] Note that we endow the set of all finite Borel measures on $\mathcal{U}$ with the weak* topology, that is, the topology where a net $\{\mu_d\}_{d\in D}$ converges to μ if and only if $\int_{\mathcal{U}} f\mu_d(du) \to \int_{\mathcal{U}} f\mu(du)$ for every continuous function $f:\mathcal{U}\to\mathbb{R}$.

and a lower semicontinuous function $c : \mathcal{M} \to \mathbb{R}$ such that $V : \mathcal{A} \to \mathbb{R}$ defined by

$$V(A) = \max_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du) - c(\mu) \right) \tag{7}$$

represents $\succsim$ and the following conditions hold:

(i) The set $\mathcal{M}$ is *consistent*: For each $\mu, \nu \in \mathcal{M}$ and $p \in \triangle(Z)$,

$$\int_{\mathcal{U}} u(p) \mu(du) = \int_{\mathcal{U}} u(p) \nu(du).$$

(ii) The set $\mathcal{M}$ is *minimal*: For any compact proper subset $\mathcal{M}'$ of $\mathcal{M}$, the function V' obtained by replacing $\mathcal{M}$ with $\mathcal{M}'$ in Equation (7) no longer represents $\succsim$.

(iii) There exist $p, q \in \triangle(Z)$ such that $V(\{p\}) > V(\{q\})$.

The following result shows that the RFCC representation can be interpreted as a reduced form of the CC representation.

THEOREM 2: *Let $V : \mathcal{A} \to \mathbb{R}$. Then there exists a CC representation such that V is given by Equation* (4) *if and only if there exists an RFCC representation such that V is given by Equation* (7).

We now sketch the construction of an equivalent RFCC representation for a given CC representation $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$. Letting $\beta = \frac{1}{|Z|} \sum_{z \in Z} \mathbb{E}[U_z]$, we showed in Lemma 1 that for any $\mathcal{G} \in \mathbf{G}$, there exists a unique finite Borel measure $\mu_{\mathcal{G}}$ on $\mathcal{U}$ such that for all $A \in \mathcal{A}$,

$$\mathbb{E}\Big[\max_{p \in A} \mathbb{E}[U|\mathcal{G}] \cdot p\Big] = \int_{\mathcal{U}} \max_{p \in A} u(p) \mu_{\mathcal{G}}(du) + \beta.$$

Let $\mathcal{M} = \{\mu_{\mathcal{G}} : \mathcal{G} \in \mathbf{G}\}$ and, for each $\mu \in \mathcal{M}$, let

$$\tilde{c}(\mu) = \inf\{c(\mathcal{G}) : \mathcal{G} \in \mathbf{G} \text{ and } \mu = \mu_{\mathcal{G}}\} - \beta.$$

By the construction of $\mathcal{M}$ and $\tilde{c}$, for any $A \in \mathcal{A}$,

$$\max_{\mathcal{G} \in \mathbf{G}} \Big(\mathbb{E}\Big[\max_{p \in A} \mathbb{E}[U|\mathcal{G}] \cdot p\Big] - c(\mathcal{G}) \Big) = \max_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du) - \tilde{c}(\mu) \right). \tag{8}$$

Also, for any $\mathcal{G} \in \mathbf{G}$ and $p \in \triangle(Z)$,

$$\int_{\mathcal{U}} u(p) \mu_{\mathcal{G}}(du) = \mathbb{E}[\mathbb{E}[U|\mathcal{G}] \cdot p] - \beta = \mathbb{E}[U] \cdot p - \beta$$

by the law of iterated expectations. This implies that the measures in $\mathcal{M}$ must satisfy the consistency condition in Definition 2. Also, condition (iii) in Definition 2 corresponds to the requirement in the definition of the CC representation that $\mathbb{E}[U] \cdot p > \mathbb{E}[U] \cdot q$ for some $p, q \in \triangle(Z)$.

The minimality condition in the definition of the RFCC representation is needed to uniquely identify the parameters in the representation. To see this, note that it is always possible to add a measure $\mu \notin \mathcal{M}$ to the set $\mathcal{M}$ and assign it a cost $c(\mu)$ high enough to guarantee that this measure is never a maximizer in Equation (7). The minimality condition requires that all such unnecessary measures be dropped from the representation. In contrast, a CC representation may include contemplation strategies that are never optimal. In the construction of an equivalent RFCC representation from a CC representation, it is therefore necessary to remove measures from $\mathcal{M}$ that are not strictly optimal in Equation (8) for some $A \in \mathcal{A}$.[17] Thus, the minimal set $\mathcal{M}$ in an RFCC representation may not include all possible contemplation strategies available to the individual, but it identifies all of the "relevant" ones.

Theorem 2 also asserts that for any RFCC representation, there exists a CC representation giving rise to the same value function V for menus. The construction used to prove this part of the theorem is more involved, so we refer the reader to Appendix D for the details.

Using Theorem 2, we establish our CC representation result (Theorem 1) by proving the following RFCC representation theorem.

THEOREM 3: *The preference $\succsim$ has an RFCC representation if and only if it satisfies weak order, strong continuity, ACP, IDD, and monotonicity.*

The proof of Theorem 3 is contained in Appendix C and is divided into two parts[18]: In Appendix C.1, we construct a function V that represents $\succsim$ and satisfies certain desirable properties: Lipschitz continuity, convexity, and a type of "translation linearity" which is closely related to the consistency condition for the measures in our representation. Then, in Appendix C.2, we apply duality results from convex analysis to establish that this function V satisfies Equation (7) for some pair $(\mathcal{M}, c)$.

We claimed that contemplation strategies can be uniquely identified from the preference once they are put into the reduced form of measures over utility functions. The following uniqueness result for the RFCC representation formalizes this claim.

[17]If $\mathbf{G}$ is finite, then the set $\mathcal{M}$ obtained by the construction above is also finite. In this case, it can be shown that sequentially removing measures from $\mathcal{M}$ that are not strictly optimal in Equation (8) for some $A \in \mathcal{A}$ leads to a minimal set of measures $\tilde{\mathcal{M}} \subset \mathcal{M}$. Although this approach can be generalized to the case of infinite $\mathbf{G}$, in Appendix D we instead give a simpler indirect proof of this direction that does not use Lemma 1.

[18]In Appendix C, we also prove a related representation result for nonmonotone preferences, which may be useful in future applications.

THEOREM 4: *If* $(\mathcal{M}, c)$ *and* $(\mathcal{M}', c')$ *are two RFCC representations for* $\succsim$, *then there exist* $\alpha > 0$ *and* $\beta \in \mathbb{R}$ *such that* $\mathcal{M}' = \alpha\mathcal{M}$ *and* $c'(\alpha\mu) = \alpha c(\mu) + \beta$ *for all* $\mu \in \mathcal{M}$.

An RFCC representation $(\mathcal{M}, c)$ in which $\mathcal{M}$ is a singleton corresponds to a monotone additive EU representation of DLR. Since DLR did not impose a normalization on the ex post expected-utility functions in their representation, their uniqueness result appears weaker than the implication of our Theorem 4 for singleton $\mathcal{M}$. However, our uniqueness result for singleton $\mathcal{M}$ is not actually stronger than theirs since the same normalization also gives a unique belief in DLR.

For the intuition behind this theorem, note that the V defined by Equation (7) for an RFCC representation is a convex function. Although the nonlinearity of this function prevents the use of standard arguments from expected-utility theory, it can still be shown that V is unique up to a positive affine transformation (see Proposition 1 in Appendix C.1). From this it can then be shown that the parameters of an RFCC representation $(\mathcal{M}, c)$ are themselves unique up to the positive affine transformation described in Theorem 4.

For a simple example of how the type of transformation described in Theorem 4 could arise, consider a CC representation $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$, and let $(\mathcal{M}, \tilde{c})$ be the corresponding RFCC representation as described in Theorem 2. If we replace the state-dependent utility function U in this CC representation with the utility function $U' = \alpha U - \beta$ and replace the cost function c with $c' = \alpha c$, where $\alpha > 0$ and $\beta \in \mathbb{R}$, then the underlying preference is the same. This new representation corresponds to the RFCC representation $(\mathcal{M}', \tilde{c}')$, where $\mathcal{M}' = \alpha\mathcal{M}$ and $\tilde{c}'(\mu) = \alpha\tilde{c}(\frac{1}{\alpha}\mu) + \beta$ for all $\mu \in \mathcal{M}'$. However, due to the nonuniqueness of the CC representation, there are many other changes to the parameters of a CC representation that could also result in such a transformation of the corresponding RFCC representation (e.g., changes to the probability distribution or information structure, or other types of changes to the utility or cost function). In particular, as illustrated in Example 1, two sets of CC parameters can correspond to precisely the same RFCC representation. Given the sharp uniqueness result that is obtained for the RFCC representation (Theorem 4), the equivalence result established above (Theorem 2) allows the nonuniqueness issue associated with CC representations to be overcome by working with equivalent RFCC representations.[19,20]

[19]Note that in the model of Ergin (2003), the preference being over finitely many menus presents an additional, more basic source of nonuniqueness. In his framework, even a reduced-form representation would not be uniquely identified.

[20]In an alternative approach to modeling costly information acquisition, Hyogo (2007) studied preferences over pairs consisting of an action and a menu of Anscombe–Aumann acts. In his representation, each action yields a distribution over posteriors over the objective state space, and the individual anticipates that she will choose an ex post optimal act from the menu. Since in his framework the state space is objective and utility is not state dependent, he is able to uniquely identify the prior over the state space and the distribution of posteriors induced by each action.

4. COMPARING CONTEMPLATION COSTS

By identifying contemplation strategies with measures over ex post utility functions as described in the previous section, it is possible to conduct meaningful comparisons of contemplation costs between two representations. In this section, we consider one such comparative measure of the cost of contemplation. Our measure will apply to preferences $\succsim$ that are *bounded above by singleton menus* in the sense that there exists an alternative $z \in Z$ such that $\{\delta_z\} \succsim A$ for all $A \in \mathcal{A}$, where δ_z denotes the lottery that puts full probability on z. For example, such an alternative z could be a very large monetary prize that is known with certainty to be better than any other alternative $z' \in Z$.

DEFINITION 3: Suppose that the preferences $\succsim_1$ and $\succsim_2$ satisfy Axioms 1–5 and are bounded above by singleton menus. We say that $\succsim_1$ *has lower cost of contemplation than* $\succsim_2$ if for every $A \in \mathcal{A}$ and $p \in \triangle(Z)$,

$$A \succsim_2 \{p\} \quad \Longrightarrow \quad A \succsim_1 \{p\}.$$

In this comparative measure, individuals face a trade-off between a menu A that may offer some flexibility and a lottery p that may be better on average. For example, consider any menu A and lottery p. If there is some $q \in A$ such that $\{q\} \succsim_i \{p\}$ for $i = 1, 2$, then $A \succsim_i \{p\}$ for $i = 1, 2$ by the monotonicity of the preferences. In this case, the condition in Definition 3 holds vacuously. Alternatively, suppose $\{p\} \succ_i \{q\}$ for $i = 1, 2$ for all $q \in A$. Then the menu A may offer flexibility if it contains more than one alternative, while p is better than the alternatives in A on average. Definition 3 formalizes the intuition that an individual is more likely to favor A over $\{p\}$ as her cost of contemplation becomes smaller since flexibility is more valuable when information about the alternatives is available at a lower cost.

Assume that the preference $\succsim_i$ has the RFCC representation $(\mathcal{M}_i, c_i)$ for $i = 1, 2$. If the sets of measures $\mathcal{M}_1$ and $\mathcal{M}_2$ are different, then it is not clear what the statement "the cost function c_1 is lower than the cost function c_2" means. In this case, there are measures $\mu \in \mathcal{M}_1 \cup \mathcal{M}_2$ for which either $c_1(\mu)$ or $c_2(\mu)$ is not defined. To overcome this problem, we will extend the cost function in an RFCC representation to the set of all measures.

DEFINITION 4: Let $\mathbf{M}$ denote the set of all finite Borel measures on $\mathcal{U}$ and let $V : \mathcal{A} \to \mathbb{R}$ be continuous. The *minimum rationalizable cost of contemplation* for V is the function $c^* : \mathbf{M} \to \mathbb{R}$ defined by

$$c^*(\mu) = \max_{A \in \mathcal{A}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p)\mu(du) - V(A) \right). \tag{9}$$

Suppose V is defined by Equation (7) for some RFCC representation $(\mathcal{M}, c)$. Then the function c^* defined by Equation (9) agrees with the cost function c on $\mathcal{M}$.[21] Moreover, for any $A \in \mathcal{A}$ and $\mu \in \mathbf{M}$, we have $V(A) \geq \int_{\mathcal{U}} \max_{p\in A} u(p)\mu(du) - c^*(\mu)$, with equality for some $A \in \mathcal{A}$. Thus, c^* is the minimal extension of c to $\mathbf{M}$ that does not alter the function V. Recall from the discussion in Section 3 that if an individual has an RFCC representation $(\mathcal{M}, c)$, then $\mu \notin \mathcal{M}$ is not a statement that the contemplation strategy corresponding to μ is not available to the individual. Rather, the exclusion of this measure from $\mathcal{M}$ implies that it is never strictly optimal and, hence, is not needed to represent the individual's preference. In this sense, it is natural to consider what contemplation costs could be attributed to measures not contained in $\mathcal{M}$. The function c^* indicates the minimum rationalizable cost of contemplation for all measures in $\mathbf{M}$, which makes it possible to compare contemplation costs between different RFCC representations.

For the following result, we use $\mathcal{S} \equiv \{\{p\} : p \in \triangle(Z)\}$ to denote the set all of singleton menus, and we write $V_2|_{\mathcal{S}} \approx V_1|_{\mathcal{S}}$ to indicate that the restriction of V_2 to $\mathcal{S}$ is a positive affine transformation of the restriction of V_1 to $\mathcal{S}$ (i.e., there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V_2(\{p\}) = \alpha V_1(\{p\}) + \beta$ for all $p \in \triangle(Z)$).

THEOREM 5: *Assume that for $i = 1, 2$, the preference $\succsim_i$ has an RFCC representation $(\mathcal{M}_i, c_i)$ and is bounded above by singleton menus. Define V_i by Equation (7) and c_i^* by Equation (9) for $i = 1, 2$. Then the following statements are equivalent*:

(i) *$\succsim_1$ has lower cost of contemplation than $\succsim_2$.*
(ii) *$V_2|_{\mathcal{S}} \approx V_1|_{\mathcal{S}}$ and $V_2 \leq V_1$ (provided $V_2|_{\mathcal{S}} = V_1|_{\mathcal{S}}$).*
(iii) *$V_2|_{\mathcal{S}} \approx V_1|_{\mathcal{S}}$ and $c_2^* \geq c_1^*$ (provided $V_2|_{\mathcal{S}} = V_1|_{\mathcal{S}}$).*[22]

To interpret condition (iii) in this theorem, first note that if $V_2|_{\mathcal{S}} \approx V_1|_{\mathcal{S}}$, then by Theorem 4 it is without loss of generality to assume that $V_2|_{\mathcal{S}} = V_1|_{\mathcal{S}}$. In this case, we have $\int_{\mathcal{U}} u(p)\mu(du) = \int_{\mathcal{U}} u(p)\nu(du)$ for all $\mu \in \mathcal{M}_2$, $\nu \in \mathcal{M}_1$, and $p \in \triangle(Z)$. In other words, the average utility of any lottery is the same for both representations. However, $c_1^*(\mu) \leq c_2^*(\mu)$ for all $\mu \in \mathbf{M}$ implies that information is less costly for the first individual. In particular, consider any $\mu \in \mathcal{M}_2$. If $\mu \in \mathcal{M}_1 \cap \mathcal{M}_2$, then $c_1(\mu) = c_1^*(\mu) \leq c_2^*(\mu) = c_2(\mu)$, that is, the contemplation strategy corresponding to μ is less costly for the first individual. Alternatively, if $\mu \in \mathcal{M}_2 \setminus \mathcal{M}_1$, then $c_1^*(\mu) \leq c_2^*(\mu) = c_2(\mu)$. Thus, if the measure μ were added to the representation $(\mathcal{M}_1, c_1)$ at a cost $c_1^*(\mu)$, where $c_1^*(\mu) \leq c_2(\mu)$, the value function for menus V_1 would not be altered. Although we cannot infer from the

[21]This result is obtained as part of the proof of Theorem 4 in Appendix E. Note that although it is immediate that $c^*(\mu) \leq c(\mu)$ for all $\mu \in \mathcal{M}$, the minimality requirement on $\mathcal{M}$ is important for obtaining the opposite inequality. For example, consider a measure $\mu \in \mathcal{M}$ that is strictly suboptimal for every menu in the sense that $V(A) > \int_{\mathcal{U}} \max_{p\in A} u(p)\mu(du) - c(\mu)$ for all $A \in \mathcal{A}$. The minimality requirement rules out the possibility of having such a measure in $\mathcal{M}$, but if it were permitted, we would obtain $c(\mu) > c^*(\mu)$.

[22]This theorem continues to hold if that assumption that $c_2^* \geq c_1^*$ in condition (iii) is replaced with the weaker assumption that $c_2^*(\mu) \geq c_1^*(\mu)$ for all $\mu \in \mathcal{M}_2$.

preference $\succsim_1$ whether the contemplation strategy corresponding to $\mu \notin \mathcal{M}_1$ is available to the first individual or not, this contemplation strategy can be rationalized by the preference at a cost $c_1^*(\mu) \leq c_2(\mu)$. In this sense, all of the contemplation strategies available to the second individual can be thought of as being available to the first individual at a lower cost.

Using the mapping from contemplation strategies $\mathcal{G}$ to measures $\mu_{\mathcal{G}}$ described in Lemma 1, we obtain the following corollary for CC representations.

COROLLARY 1: *Assume that for $i = 1, 2$, the preference $\succsim_i$ has a CC representation $((\Omega_i, \mathcal{F}_i, P_i), \mathbf{G}_i, U_i, c_i)$ and is bounded above by singleton menus. Define V_i by Equation (4) and c_i^* by Equation (9) for $i = 1, 2$. Then the following statements are equivalent*:

(i) *$\succsim_1$ has lower cost of contemplation than $\succsim_2$.*

(ii) *$V_2|_{\mathcal{S}} \approx V_1|_{\mathcal{S}}$ and $V_2 \leq V_1$ (provided $V_2|_{\mathcal{S}} = V_1|_{\mathcal{S}}$).*

(iii) *$V_2|_{\mathcal{S}} \approx V_1|_{\mathcal{S}}$ and $c_2^*(\mu_{\mathcal{G}_2}) \geq c_1^*(\mu_{\mathcal{G}_2})$ for all $\mathcal{G}_2 \in \mathbf{G}_2$ (provided $V_2|_{\mathcal{S}} = V_1|_{\mathcal{S}}$).*

The interpretation of this corollary is similar to that of Theorem 5, with the following caveat: In a CC representation $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$, the interpretation of the function c^* as an extension of the cost function c is a little more subtle than in the case of the RFCC representation. Aside from the obvious distinction that the domain of c is not actually a subset of the domain of c^*, it is possible to have $c^*(\mu_{\mathcal{G}}) < c(\mathcal{G})$ for some $\mathcal{G} \in \mathbf{G}$. In particular, Lemma 14 in Appendix F.2 shows that $c^*(\mu_{\mathcal{G}}) \leq c(\mathcal{G})$, with equality if and only if $\mathcal{G}$ solves Equation (4) for some $A \in \mathcal{A}$.[23] Therefore, if we let $\hat{\mathbf{G}}$ denote the subset of contemplation strategies that solve Equation (4) for some menu, then $c^*(\mu_{\mathcal{G}}) = c(\mathcal{G})$ for any $\mathcal{G} \in \hat{\mathbf{G}}$. Thus, c^* can be thought of as the minimal extension of $c|_{\hat{\mathbf{G}}}$ to $\mathbf{M}$ that does not alter the function V.

5. SPECIAL CASES

5.1. *Limited Contemplation Resources*

In this section, we consider an alternative model of costly contemplation in which the cost of contemplation does not directly affect the utility of the individual. Instead, the cost of contemplation enters indirectly when it is constrained to be below some bound k. Such a model may be appropriate in instances where the only cost of contemplation is time and the individual has a limited amount of time to devote to her decision.

Formally, we continue to model contemplation in the reduced form of a compact set of finite Borel measures $\mathcal{M}$ over the set of ex post utility functions $\mathcal{U}$, with the requirement that $\mathcal{M}$ be consistent and minimal. Let $c : \mathcal{M} \to \mathbb{R}$ be a

[23]Throughout this discussion, we assume for expositional simplicity that $\beta \equiv \frac{1}{|Z|} \sum_{z \in Z} \mathbb{E}[U_z] = 0$. As we show in Lemma 14, without this assumption, the inequality would be $c^*(\mu_{\mathcal{G}}) \leq c(\mathcal{G}) - \beta$.

lower semicontinuous cost function and let $k \in \mathbb{R}$ be the maximum allowable contemplation cost. A representation for limited contemplation resources then takes the form of a function $V : \mathcal{A} \to \mathbb{R}$ defined by

$$V(A) = \max_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du) \right) \quad \text{subject to } c(\mu) \leq k. \tag{10}$$

If we let $\mathcal{M}' = \{\mu \in \mathcal{M} : c(\mu) \leq k\}$, then this representation is equivalent to

$$V(A) = \max_{\mu \in \mathcal{M}'} \left(\int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du) \right).$$

Moreover, since c is lower semicontinuous, $\mathcal{M}'$ is also compact. Thus, the limited contemplation resources representation in Equation (10) is equivalent to an RFCC representation with a zero cost function, $(\mathcal{M}', 0)$. Since the cost function in an RFCC representation is only unique up to an affine transformation, we see that a preference has a representation as in Equation (10) if and only if it has an RFCC representation $(\mathcal{M}', c')$, where c' is constant.

We now introduce an axiom that characterizes a constant cost of contemplation for all available contemplation strategies.

AXIOM 6—Strong IDD: *For any* $A, B \in \mathcal{A}$, $p \in \triangle(Z)$, *and* $\alpha \in (0, 1)$,

$$A \succsim B \quad \Longleftrightarrow \quad \alpha A + (1 - \alpha)\{p\} \succsim \alpha B + (1 - \alpha)\{p\}.$$

As the name suggests, strong IDD is a strengthening of IDD. Suppose

$$\alpha A + (1 - \alpha)\{p\} \succsim \alpha B + (1 - \alpha)\{p\}$$

for some $A, B \in \mathcal{A}$, $p \in \triangle(Z)$, and $\alpha \in (0, 1)$. Strong IDD then implies $A \succsim B$, and applying strong IDD again, we have

$$\beta A + (1 - \beta)\{q\} \succsim \beta B + (1 - \beta)\{q\}$$

for any $q \in \triangle(Z)$ and $\beta \in (0, 1)$. In contrast, IDD only guarantees that the above preference holds for $\beta = \alpha$. Thus, strong IDD implies an independence of degenerate decisions (IDD) and, in addition, independence of the weights on these degenerate decisions.[24]

[24]Strong IDD is similar in spirit to the certainty independence axiom used by Gilboa and Schmeidler (1989) in the Anscombe–Aumann setting. In their axiom, arbitrary acts play the role of the menus A and B, and a constant act plays the role of the singleton menu $\{p\}$. Our discussion of the relationship between strong IDD and IDD parallels the comparison of certainty independence and weak certainty independence found in Section 3.1 of Maccheroni, Marinacci, and Rustichini (2006).

For intuition, recall that the menu $\alpha A + (1-\alpha)\{p\}$ represents the decision problem in which the individual makes a contingent choice from A, this choice is implemented with probability α, and with probability $1-\alpha$ the individual instead receives p. We argued in Section 2.1 that as α decreases, the individual's benefit from contemplation decreases, causing her to choose a less costly contemplation strategy. However, if the cost of all available contemplation strategies is the same, then her optimal contemplation strategy when choosing from the menu A will be the same as her optimal contemplation strategy when choosing from $\alpha A + (1-\alpha)\{p\}$ for any $\alpha \in (0,1)$. Therefore, if $A \succsim B$, then taking the convex combination of these menus with some singleton menu $\{p\}$ could affect the individual's utility through its effect on the final composition of lotteries, but it will not affect her optimal contemplation strategy for each of the respective menus. Hence, her ranking of the menus will not change.

The following theorem formalizes the connection between strong IDD and constant contemplation costs.

THEOREM 6: *Suppose the preference $\succsim$ has an RFCC representation $(\mathcal{M}, c)$. Then $\succsim$ satisfies strong IDD if and only if c is constant.*

Given the relationship between the CC representation and the RFCC representation, we obtain the following corollary.

COROLLARY 2: *For a preference $\succsim$ on $\mathcal{A}$, the following statements are equivalent*:

(i) *The preference $\succsim$ satisfies weak order, strong continuity, ACP, strong IDD, and monotonicity.*

(ii) *There exists a probability space $(\Omega, \mathcal{F}, P)$, a collection $\mathbf{G}$ of sub-σ-algebras of $\mathcal{F}$, a Z-dimensional, $\mathcal{F}$-measurable, and integrable random vector U, a cost function $c: \mathbf{G} \to \mathbb{R}$, and a constant $k \in \mathbb{R}$ such that the preference $\succsim$ is represented by*

$$V(A) = \max_{\mathcal{G} \in \mathbf{G}} \mathbb{E}\Big[\max_{p \in A} \mathbb{E}[U|\mathcal{G}] \cdot p\Big] \quad \textit{subject to } c(\mathcal{G}) \leq k, \tag{11}$$

where the outer maximization in Equation (11) has a solution for every $A \in \mathcal{A}$ and there exist $p, q \in \triangle(Z)$ such that $\mathbb{E}[U] \cdot p > \mathbb{E}[U] \cdot q$.

5.2. *Connection to the Independence Axiom*

In this section, we discuss the special case of our model in which the full-information contemplation strategy is available and no more costly than any other (less informative) contemplation strategy. This special case will be closely related to the analysis of DLR, who introduced the following independence axiom for sets of lotteries.

AXIOM 7—Independence: *For any* $A, B, C \in \mathcal{A}$ *and* $\alpha \in (0, 1)$,

$$A \succ B \implies \alpha A + (1-\alpha)C \succ \alpha B + (1-\alpha)C.$$

It is easily verified that under weak order and continuity, independence implies ACP and strong IDD. Note also that under weak order and continuity, independence implies a form of indifference to contingent planning: For any $A, B \in \mathcal{A}$ and $\alpha \in [0, 1]$, $A \sim B$ implies $A \sim \alpha A + (1-\alpha)B$. Intuitively, this suggests that independence rules out the possibility of multiple contemplation strategies.

For a simple example of why multiple contemplation strategies are inconsistent with the independence axiom, let A and B be the restaurant menus described in the Introduction, that is, $A = \{t, s\}$ and $B = \{f, r\}$. Suppose the individual has two contemplation strategies, both of which have zero cost: (i) contemplate which seafood dish she would like and (ii) contemplate which steak dish she would like. In particular, it is not possible for the individual to contemplate both restaurant menus. This could occur if, as discussed in Section 5.1, the individual's contemplation is constrained due to limited time and there is not sufficient time to think about both restaurant menus. Then, when faced with either menu A or B, the individual can choose a contemplation strategy that allows her to pick the ex post optimal alternative with probability 1. However, since she cannot contemplate both menus simultaneously, it is not possible for her to choose the ex post optimal alternative with certainty when asked to make a contingent plan from $\alpha A + (1-\alpha)B$. Therefore, if the items on these menus are such that $A \sim B$, it follows that $A \succ \alpha A + (1-\alpha)B$, in violation of the independence axiom.[25] The following result generalizes these observations by showing that the independence axiom is equivalent to an RFCC representation with a single contemplation strategy.

THEOREM 7: *The preference* $\succsim$ *satisfies weak order, strong continuity, independence, and monotonicity if and only if it has an RFCC representation* $(\mathcal{M}, c)$ *in which* $\mathcal{M}$ *is a singleton.*

[25] It is well known that independence may be violated if the individual takes a payoff-relevant action prior to the resolution of uncertainty. In the context of our CC representation, the individual facing the complete contingent plan $\alpha A + (1-\alpha)B$ chooses her contemplation strategy before the uncertainty regarding the menu (A or B) is resolved. In the context of choices among lotteries, Mossin (1969) gave the example of an individual who has expected-utility preferences over two-period consumption vectors and makes a savings decision in period 1. Mossin argued that the individual's induced preferences over second-period income distributions may violate independence if the savings decision precedes the resolution of uncertainty regarding the second-period income. Such induced preferences over lotteries naturally satisfy a quasiconvexity property analogous to our ACP axiom: $p \succsim q \Rightarrow p \succsim \alpha p + (1-\alpha)q$. Quasiconvexity of preferences over monetary prizes has also been studied for entirely different purposes in economics. For instance, Green (1987) showed that an individual who has fixed, time-independent, continuous, and monotone preferences over lotteries over monetary prizes is prone to "money pumps" starting from nonrandom wealth levels if and only if her preferences are quasiconvex.

We will not provide a proof of this result since it is simply a restatement of the additive EU representation theorem of DLR and DLRS.[26] The following corollary shows the implications of independence for the CC representation.

COROLLARY 3: *The preference* $\succsim$ *satisfies weak order*, *strong continuity*, *independence*, *and monotonicity if and only if it has a CC representation* $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$ *such that* $\mathcal{F} \in \mathbf{G}$ *and* $c(\mathcal{F}) = \min_{\mathcal{G} \in \mathbf{G}} c(\mathcal{G})$.

Corollary 3 states that a preference $\succsim$ that satisfies independence (and our other axioms) can be represented with a CC representation in which the full-information contemplation strategy is available and no more costly than any other contemplation strategy. However, due to the nonuniqueness of the CC representation, there are also other CC representations for this preference in which the full-information contemplation strategy is not the least costly. Indeed, an individual's preference will satisfy independence whenever there is a single optimal contemplation strategy, even if it is not the most informative. Therefore, it is not possible to determine from the preference whether or not the full-information contemplation strategy is the least costly; the independence axiom simply indicates that the preference *can* be represented as if the full-information contemplation strategy is the least costly.

In the remainder of this section, we provide graphical intuition for our main axioms (ACP and IDD) and illustrate how these axioms relax the independence axiom. Consider preferences over menus of lotteries over two alternatives. That is, suppose $Z = \{z_1, z_2\}$. In this case, the set of lotteries over Z can be represented as the unit interval $[0, 1]$, with $p \in [0, 1]$ being the probability of alternative z_2. Under weak order and continuity, ACP and monotonicity imply that the individual is indifferent between any menu and its convex hull (see Lemma 2 in Appendix C.1). We can therefore restrict attention to convex menus. Closed and convex menus from $[0, 1]$ are simply closed intervals, and hence we are considering preferences over menus of the form $[p, q] \subset [0, 1]$ where $p, q \in [0, 1]$.

[26]Although DLR do not impose a normalization on the set of ex post expected-utility functions in the definition of their representation, the proof of their representation result uses a set of ex post utility functions that is precisely $\mathcal{U}$ as defined in Equation (6). In particular, it is shown in the Supplemental Material of DLRS that the preference $\succsim$ satisfies weak order, strong continuity, independence, and monotonicity if and only if there exists a finite Borel measure μ on $\mathcal{U}$ such that $\succsim$ is represented by the functional form

$$V(A) = \int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du).$$

Since DLRS used a slightly weaker L-continuity axiom, their representation need not satisfy singleton nontriviality (condition (iii) in Definition 2). However, under the strong continuity axiom of the current paper, singleton nontriviality will be satisfied. Hence, the pair $(\mathcal{M}, c)$, where $\mathcal{M} = \{\mu\}$ and $c = 0$, is an RFCC representation for $\succsim$.

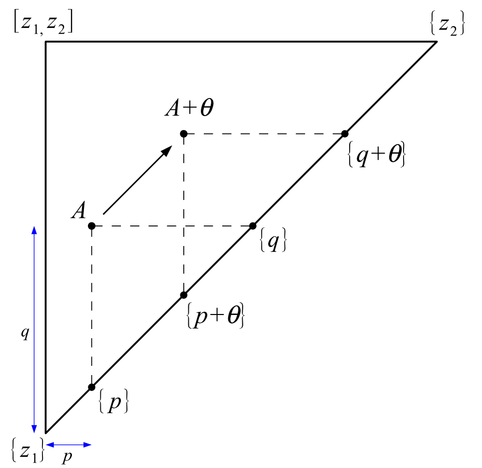

FIGURE 1.—Representing convex menus.

The set of all menus of this form is illustrated in Figure 1.[27] Consider any interval $A = [p, q]$. This interval corresponds to the point in the triangle whose first coordinate is p and whose second coordinate is q, that is, the point whose horizontal distance from the left side of the graph is p and whose vertical distance from the bottom of the graph is q. In particular, the set of all singleton menus (i.e., menus of the form $\{p\} = [p, p]$) is represented by the diagonal of the triangle in this figure. Note that we abuse notation slightly and let z_1 denote the lottery that gives z_1 with probability 1, and likewise for z_2. Thus, the corners of the triangle labeled $\{z_1\}$, $\{z_2\}$, and $[z_1, z_2]$ correspond to the menus $\{0\}$, $\{1\}$, and $[0, 1]$, respectively.

When the set of closed and convex menus is represented as in Figure 1, a convex combination of two menus corresponds to the convex combination of the points representing these menus. Therefore, the implication of ACP is simply that the lower contour sets for the preference are convex sets. Before illustrating the implications of IDD, we make a few observations about "translations" of menus. Consider the menu $A = [p, q]$ indicated in Figure 1, and take some real number θ. Adding the translation θ to the menu A yields a new menu $A + \theta = [p + \theta, q + \theta]$. Figure 1 illustrates that translating a menu results in a movement in a direction parallel to the diagonal of the triangle.

Figure 2 builds on these observations to show that IDD implies a type of translation invariance.[28] That is, we will show that if the individual is indifferent between two menus, then she is also indifferent between the new menus

[27]A similar depiction of menus of lotteries can be found in Olszewski (2007).

[28]This property is defined formally in Appendix C.1 and plays an important role in the proof of Theorem 3.

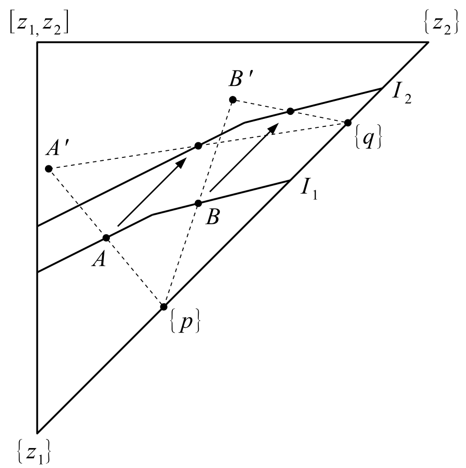

FIGURE 2.—Translation invariance.

obtained by translating them both the same distance in a direction parallel to the diagonal of the triangle. Consider any two menus A and B such that $A \sim B$. Therefore, as illustrated in Figure 2, A and B both lie on the same indifference curve I_1. Note that for this preference to satisfy ACP, the lower contour sets of the preference must be convex, and hence the points above I_1 must be preferred to the points below I_1. Figure 2 illustrates that the menus A and B can be written as convex combinations of the singleton menu $\{p\}$ with the menus A' and B', respectively. That is, there exists $\alpha \in (0, 1)$ such that $A = \alpha A' + (1 - \alpha)\{p\}$ and $B = \alpha B' + (1 - \alpha)\{p\}$. Fix any lottery q. Then by IDD, we have

$$\begin{aligned} &\alpha A' + (1 - \alpha)\{p\} \sim \alpha B' + (1 - \alpha)\{p\} \\ &\quad \Longrightarrow \quad \alpha A' + (1 - \alpha)\{q\} \sim \alpha B' + (1 - \alpha)\{q\}. \end{aligned}$$

Thus, the menus $\alpha A' + (1 - \alpha)\{q\}$ and $\alpha B' + (1 - \alpha)\{q\}$ must also be on the same indifference curve, which is indicated by I_2 in Figure 2. However, letting $\theta = (1 - \alpha)(q - p)$, it is easily seen that $A + \theta = \alpha A' + (1 - \alpha)\{q\}$ and $B + \theta = \alpha B' + (1 - \alpha)\{q\}$. In other words, if the menus A and B are both translated by θ, then the individual remains indifferent between them. More generally, it can be shown that IDD implies that when the same translation is applied to any two menus, the individual's ranking of these menus is not altered (see Lemma 3).

These figures show that although ACP and IDD allow for "kinks" in indifference curves, these axioms require that lower contour sets be convex and that indifference curves be translations of each other. Note that the kinks in the indifference curves in Figure 2 indicate a change in the optimal contemplation strategy, and our model allows for a possibly infinite number of kinks. In contrast, the independence axiom requires that indifference curves be linear and does not allow for such kinks. These observations illustrate why it is necessary to relax independence so as to allow for nondegenerate costly contemplation, that is, costly contemplation with more than one contemplation strategy.[29]

6. INFINITE REGRESS

We conclude by discussing the infinite-regress problem of bounded rationality (see Lipman (1991) and Conlisk (1996)) within the context of our main representation theorem. Let $\mathcal{D}$ stand for some collection of abstract decision problems. In theoretical economic analysis, standard rational agents are assumed to solve any decision problem $D \in \mathcal{D}$ optimally without any constraints. One may think that this is not a realistic assumption when the decision problem D is difficult in some sense, and be tempted to make the model more realistic by explicitly taking into account the costs of solving D. Let F be a correspondence that associates with every decision problem $D \in \mathcal{D}$ a set of new decision problems $F(D) \subset \mathcal{D}$ obtained by incorporating into D the costs of solving D.

Typically, the decision problems in $F(D)$ are even more "difficult" than D, in the same sense in which D is difficult to start with. Therefore, assuming that the individual solves the problems in $F(D)$ optimally is no more reasonable than assuming that she solves D optimally. Explicitly including the costs of solving the decision problems in $F(D)$ leads to a new class of decision problems $F^2(D) = F(F(D)) = \bigcup_{D' \in F(D)} F(D')$. This argument can be iterated ad infinitum. Since the classes of problems $D, F(D), F^2(D), \ldots, F^n(D), \ldots$ become progressively more complicated, assuming that the individual solves any one of them optimally defeats the initial purpose of building a more realistic model. This is the infinite-regress problem.

To state the infinite-regress problem within the context of our model, assume that each decision problem in $D \in \mathcal{D}$ specifies a set of actions and payoffs from these actions. The augmented decision problems in $F(D)$ introduce uncertainty about the individual's payoffs from the actions in D, but allow her to acquire costly information about this uncertainty and condition her choice of action from D on her information. As a result, the augmented actions in a decision problem $D' \in F(D)$ are pairs consisting of (i) the choice of information and (ii) the choice of action from D contingent on the realized information. The payoff function corresponding to D' is obtained by taking the expected payoff from the augmented action minus the cost of acquired information.

[29]A relaxation of the independence axiom in the setting of preferences over menus of lotteries was also considered by Epstein, Marinacci, and Seo (2007), who interpreted their axiom as the behavior of an individual with an incomplete (or coarse) conception of the future. This coarse conception entails a degree of pessimism on the part of the individual, and their resulting representations are intuitively similar to the maxmin representation of Gilboa and Schmeidler (1989).

The most basic type of decision problem D_0 we consider specifies a menu A of lotteries interpreted as actions and an expected-utility function $u : \triangle(Z) \to \mathbb{R}$. The optimization problem corresponding to D_0 is to find the utility-maximizing lottery out of the given menu. Let $\mathcal{D}_0$ denote the set of such basic decision problems. The (once) augmented decision problems $D_1 \in F(D_0)$ consist of all maximization problems of type

$$(\#) \qquad \max_{(\mathcal{G},f)} \mathbb{E}[U \cdot f] - c(\mathcal{G}),$$

where, as in the CC formulation, $(\Omega, \mathcal{F}, P)$ is a probability space and U is a state-dependent utility function representing the individual's uncertainty about her payoff from actions in A, $\mathbf{G}$ is a collection of sub-σ-algebras of $\mathcal{F}$ specifying the information that the individual can acquire, and $c(\mathcal{G})$ denotes the cost of information $\mathcal{G}$. The maximization is done over all augmented actions $(\mathcal{G}, f)$, where $\mathcal{G} \in \mathbf{G}$ and $f : \Omega \to A$ determines a plan of actions measurable with respect to the acquired information $\mathcal{G}$.[30]

We can define $F(D_1)$ by introducing further uncertainty about the augmented decision problem (#) above. More specifically, we can introduce uncertainty about the probability P, the state dependent utility function U, and the cost function c, and allow the individual to acquire costly information about this additional uncertainty and condition her choice of action $(\mathcal{G}, f)$ in Equation (#) on this information. It is straightforward to see how this construction can be iterated an arbitrary number of times to construct $F^n(D_0)$ for an arbitrary $n \geq 1$. We let $\mathcal{D} = \bigcup_{n=0}^{\infty} F^n(\mathcal{D}_0)$.

Although one can also argue within the context of our model that the classes of decision problems $D_0, F(D_0), F^2(D_0), \dots, F^n(D_0), \dots$ become progressively more complicated because they involve solving higher-order information acquisition problems, our representation result is immune to this criticism. To see this, consider an excerpt from Lipman (1995, p. 59), who explained why axiomatic approaches to bounded rationality are not susceptible to the infinite-regress criticism:

> Roughly, the axiomatic approach begins with a description of the agent and then translates this into a model of information processing. Clearly, it then makes no sense to ask whether the agent can carry out this information processing accurately. If the processing is simply a representation of what the agent is doing, the question boils down to asking whether an agent is able to do whatever it is that he does!

[30]The one-shot maximization problem corresponding to (#) is equivalent to the two-stage maximization problem in the CC formulation where the individual first chooses her information $\mathcal{G}$ and then chooses a lottery maximizing her ex post expected utility $\mathbb{E}[U|\mathcal{G}]$ conditional on the realized information. We are using the one-shot maximization formulation in Equation (#) because it is more explicit about the action space for D_1.

In particular, to the extent that one finds ACP and IDD to be convincing behavioral aspects of bounded rationality arising from contemplation costs, there is no loss of generality from restricting attention to the case where the decision maker optimally solves the problem of learning her preferences subject to costs, that is, to the case where she optimally solves $F(D_0)$. Therefore, our representation result may be seen as giving an *as if* solution to the infinite-regress problem.

On a final note, it is standard that every nth-level problem $D_n \in F^n(D_0)$ where $n \geq 1$ can be collapsed to a first-level problem $D_1 \in F(D_0)$ by rewriting the dynamic information acquisition problem in D_n as a one-shot augmented costly information acquisition problem. In particular, if we do not observe the individual's sequence of information acquisition choices, then we cannot distinguish between first-level and nth-level decision problems. Since the only observable in our model is the individual's preferences over menus, the representations where the individual solves a higher-order subjective information acquisition problem in $F^n(D_0)$ for $n \geq 2$ are behaviorally indistinguishable from those where she solves a first-order problem in $F(D_0)$.

APPENDIX A: SHOWING THE CC REPRESENTATION IS WELL DEFINED

Consider any CC representation $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$. In this section, we show that the term $\mathbb{E}[\max_{p\in A} \mathbb{E}[U|\mathcal{G}] \cdot p]$ is well defined and finite for every $A \in \mathcal{A}$ and $\mathcal{G} \in \mathbf{G}$. This in particular implies that $V(A)$ is finite whenever the outer maximization in Equation (4) has a solution. Let $\tilde{U}$ be an arbitrary version of $\mathbb{E}[U|\mathcal{G}]$. The existence and integrability of $\tilde{U}_z$ follow from integrability of U_z for each $z \in Z$ (see Billingsley (1995, p. 445)). Let B be a countable dense subset of A. At each $\omega \in \Omega$, $\max_{p\in A} \tilde{U}(\omega) \cdot p$ exists and is equal to $\sup_{p\in B} \tilde{U}(\omega) \cdot p$. For each $p \in B$, $\tilde{U} \cdot p$ is $\mathcal{F}$-measurable as a convex combination of $\mathcal{F}$-measurable random variables. Hence, $\max_{p\in A} \tilde{U} \cdot p = \sup_{p\in B} \tilde{U} \cdot p$ is an $\mathcal{F}$-measurable random variable as the pointwise supremum of countably many $\mathcal{F}$-measurable random variables (see Billingsley (1995, p. 184, Theorem 13.4(i))). Note also that for any $p \in \triangle(Z)$, $|\tilde{U} \cdot p| \leq \sum_{z\in Z} |\tilde{U}_z|$, and hence $|\max_{p\in A} \tilde{U} \cdot p| \leq \sum_{z\in Z} |\tilde{U}_z|$. Therefore, integrability of $\max_{p\in A} \tilde{U} \cdot p$ follows from integrability of $\tilde{U}$.

APPENDIX B: PROOF OF LEMMA 1

Fix a CC representation $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$ and fix $\mathcal{G} \in \mathbf{G}$. Let $\mathbf{1} \in \mathbb{R}^Z$ denote the vector whose coordinates are all equal to 1. It is easy to show that there exist $\mathcal{G}$-measurable and integrable functions $\alpha : \Omega \to \mathbb{R}_+$, $\beta : \Omega \to \mathbb{R}$, and

$u : \Omega \to \mathcal{U}$ such that

$$\mathbb{E}[U|\mathcal{G}] = \alpha u + \beta \mathbf{1}, \quad P\text{-almost surely.}$$[31]

Define a positive finite measure m on $(\Omega, \mathcal{G})$ via its Radon–Nikodym derivative $\frac{dm}{dP}(\omega) = \alpha(\omega)$, and define a finite Borel measure $\mu_{\mathcal{G}}$ on $\mathcal{U}$ via $\mu_{\mathcal{G}} = m \circ u^{-1}$.[32] Let $\beta_{\mathcal{G}} = \mathbb{E}[\beta]$. Then, for any menu $A \in \mathcal{A}$,

$$\begin{aligned}
\mathbb{E}\Big[\max_{p\in A} \mathbb{E}[U|\mathcal{G}] \cdot p\Big] &= \int_{\Omega} \Big[\alpha(\omega) \max_{p\in A} u(\omega) \cdot p\Big] P(d\omega) + \mathbb{E}[\beta] \\
&= \int_{\Omega} \Big[\max_{p\in A} u(\omega) \cdot p\Big] m(d\omega) + \beta_{\mathcal{G}} \\
&= \int_{\mathcal{U}} \Big[\max_{p\in A} u \cdot p\Big] \mu_{\mathcal{G}}(du) + \beta_{\mathcal{G}},
\end{aligned}$$

where the final equality follows from the change of variables formula. In addition, taking $p = (1/|Z|, \dots, 1/|Z|)$, we have $u \cdot p = 0$ for all $u \in \mathcal{U}$. Thus, letting $A = \{p\}$ in the above equation, we have

$$\beta_{\mathcal{G}} = \mathbb{E}[\mathbb{E}[U|\mathcal{G}] \cdot p] = \mathbb{E}\left[\frac{1}{|Z|} \sum_{z\in Z} \mathbb{E}[U_z|\mathcal{G}]\right] = \frac{1}{|Z|} \sum_{z\in Z} \mathbb{E}[U_z].$$

To show that the $\mu_{\mathcal{G}}$ and $\beta_{\mathcal{G}}$ defined above are unique, consider any other $\mu'_{\mathcal{G}}$ and $\beta'_{\mathcal{G}}$ such that for all $A \in \mathcal{A}$,

$$\int_{\mathcal{U}} \max_{p\in A} (u \cdot p) \mu_{\mathcal{G}}(du) + \beta_{\mathcal{G}} = \int_{\mathcal{U}} \max_{p\in A} (u \cdot p) \mu'_{\mathcal{G}}(du) + \beta'_{\mathcal{G}}.$$

Taking $p = (1/|Z|, \dots, 1/|Z|)$ and letting $A = \{p\}$, the above equation implies $\beta_{\mathcal{G}} = \beta'_{\mathcal{G}}$. This in turn implies that for any $A \in \mathcal{A}$,

$$\int_{\mathcal{U}} \max_{p\in A} (u \cdot p) \mu_{\mathcal{G}}(du) = \int_{\mathcal{U}} \max_{p\in A} (u \cdot p) \mu'_{\mathcal{G}}(du).$$

[31]For example, fix any version $\tilde{U}$ of $\mathbb{E}[U|\mathcal{G}]$. Letting $\bar{u}$ be any element of $\mathcal{U}$ and letting $\|\cdot\|$ denote the standard Euclidean norm on $\mathbb{R}^Z$, take $\beta(\omega) = \frac{1}{|Z|} \sum_{z\in Z} \tilde{U}_z(\omega)$, $\alpha(\omega) = \|\tilde{U}(\omega) - \beta(\omega)\mathbf{1}\|$, and

$$u(\omega) = \begin{cases} \dfrac{\tilde{U}(\omega) - \beta(\omega)\mathbf{1}}{\alpha(\omega)}, & \text{if } \alpha(\omega) \neq 0, \\ \bar{u} & \text{if } \alpha(\omega) = 0. \end{cases}$$

It is a standard exercise to check that α, β, and u so defined are $\mathcal{G}$-measurable and integrable.

[32]That is, $m(E) = \int_E \alpha(\omega) P(d\omega)$ for any $E \in \mathcal{G}$, and $\mu_{\mathcal{G}}(F) = m \circ u^{-1}(F) = \int_{u^{-1}(F)} \alpha(\omega) P(d\omega)$ for any Borel measurable set $F \subset \mathcal{U}$.

Since each side of this equality is what DLR referred to as an additive EU representation, we can apply their uniqueness result to conclude that $\mu_{\mathcal{G}} = \mu'_{\mathcal{G}}$.[33]

APPENDIX C: PROOF OF THEOREM 3

In this section, we prove two results. We first prove a general representation theorem for preferences that may violate monotonicity and subsequently establish Theorem 3 as a special case. The following definition is a generalization of the RFCC representation to allow for signed measures.

DEFINITION 5: A *signed RFCC representation* is a pair $(\mathcal{M}, c)$ consisting of a compact set of finite signed Borel measures $\mathcal{M}$ on $\mathcal{U}$ and a lower semicontinuous function $c : \mathcal{M} \to \mathbb{R}$ such that $V : \mathcal{A} \to \mathbb{R}$ defined by Equation (7) represents $\succsim$ and (i)–(iii) in Definition 2 are satisfied.

The signed RFCC representation is of interest since it can be used to model a preference for commitment in conjunction with costly contemplation. A preference for commitment could arise if an individual experiences regret or temptation. See, for example, Sarver (2008) for a model of regret and Gul and Pesendorfer (2001) or Dekel, Lipman, and Rustichini (2008) for models of temptation and self-control. The representations considered in those papers are special cases of the singleton signed RFCC representation (i.e., the signed RFCC representation with a single measure). We conjecture that models that combine regret or temptation with costly contemplation could be represented in reduced form as special cases of the general signed RFCC representation. We leave the investigation of such models as an open question for future research.

To allow for signed measures, we replace the monotonicity axiom with the following axiom introduced by DLR.

AXIOM 8—Indifference to Randomization: *For every* $A \in \mathcal{A}$, $A \sim \mathrm{co}(A)$.

Indifference to randomization (IR) is justified if the individual choosing from the menu A can also randomly select an alternative from the menu, for example, by flipping a coin. In that case, the menus A and $\mathrm{co}(A)$ offer the same set of options, and hence they are identical from the perspective of the individual. In this section, we prove the following theorem.

THEOREM 8: (A) *The preference* $\succsim$ *has a signed RFCC representation if and only if it satisfies weak order, strong continuity, ACP, IDD, and IR.*

(B) *The preference* $\succsim$ *has an RFCC representation if and only if it satisfies weak order, strong continuity, ACP, IDD, and monotonicity.*

[33] This particular version of the uniqueness result for the additive EU representation can be found in Sarver (2008, Lemma 18) for the case where $\mu_{\mathcal{G}}$ and $\mu'_{\mathcal{G}}$ are Borel probability measures. Extending the result to arbitrary finite Borel measures is trivial.

Theorem 8(B) is simply a restatement of Theorem 3, and Theorem 8(A) characterizes the signed RFCC representation. Note also that the IR axiom is not included in Theorem 8(B) because it is implied by the other axioms (see Lemma 2 in Appendix C.1).

The remainder of this section is devoted to the proof of Theorem 8. With the exception of L-continuity, the necessity of the axioms in Theorem 8 is straightforward and left to the reader. The proof of the necessity of L-continuity is contained in Section S.2 of the Supplemental Material. For the sufficiency direction, let $\mathcal{A}^c \subset \mathcal{A}$ denote the collection of all convex menus. In both parts (A) and (B) of Theorem 8, $\succsim$ satisfies IR. In part (A), IR is directly assumed, whereas in part (B) it is implied by the other axioms. Therefore, for all $A \in \mathcal{A}$, $A \sim \mathrm{co}(A) \in \mathcal{A}^c$. Note that for any $u \in \mathcal{U}$, we have

$$\max_{p \in A} u \cdot p = \max_{p \in \mathrm{co}(A)} u \cdot p.$$

Thus, if we establish the representations in Theorem 8 for convex menus and then apply the same functional form to all of $\mathcal{A}$, then by IR the resulting function represents $\succsim$ on $\mathcal{A}$.

Note also that $\mathcal{A}$ is a compact metric space since $\triangle(Z)$ is a compact metric space (see, e.g., Munkres (2000, pp. 280–281) or Theorem 1.8.3 in Schneider (1993, p. 49)). It is a standard exercise to show that $\mathcal{A}^c$ is a closed subset of $\mathcal{A}$, and hence $\mathcal{A}^c$ is also compact (see Theorem 1.8.5 in Schneider (1993, p. 50)).

In Section C.1, we construct a function V with certain desirable properties. In Section C.2, we apply the duality results from Section S.1 of the Supplemental Material to the function V, which completes the proof of the sufficiency part of Theorem 8.

C.1. *Construction of V*

We start by establishing a simple implication of the axioms introduced in the text.

LEMMA 2: *If $\succsim$ satisfies weak order, continuity, ACP, and monotonicity, then it also satisfies IR.*

PROOF: Let $A \in \mathcal{A}$. Monotonicity implies that $\mathrm{co}(A) \succsim A$, and hence we only need to prove that $A \succsim \mathrm{co}(A)$. Let us inductively define a sequence of sets via $A_0 = A$ and $A_k = \frac{1}{2}A_{k-1} + \frac{1}{2}A_{k-1}$ for $k \geq 1$. ACP implies that $A_{k-1} \succsim A_k$ and therefore, by transitivity, $A \succsim A_k$ for any k. It is straightforward to verify that $d_h(A_k, \mathrm{co}(A)) \to 0$, so we have $A \succsim \mathrm{co}(A)$ by continuity. *Q.E.D.*

For proving our representation theorem, it will be useful to derive an alternative formulation of our IDD axiom. Before introducing this new axiom, we define the set of *translations* to be

$$\Theta \equiv \left\{\theta \in \mathbb{R}^Z : \sum_{z\in Z} \theta_z = 0\right\}. \tag{12}$$

Any $\theta \in \Theta$ can be thought of as a signed measure on Z such that $\theta(Z) = 0$. For $A \in \mathcal{A}$ and $\theta \in \Theta$, define $A + \theta \equiv \{p + \theta : p \in A\}$. Intuitively, adding θ to A in this sense simply "shifts" A. Also, note that for any $p, q \in \triangle(Z)$, we have $p - q \in \Theta$. We now give a formulation of IDD in terms of translations.

AXIOM 9—Translation Invariance: *For any $A, B \in \mathcal{A}$ and $\theta \in \Theta$ such that $A + \theta, B + \theta \in \mathcal{A}$,*

$$A \succsim B \quad \Longrightarrow \quad A + \theta \succsim B + \theta.$$[34]

LEMMA 3: *The preference $\succsim$ satisfies IDD if and only if it satisfies translation invariance* (*TI*).

PROOF: To see that TI implies IDD, assume that $A, B \in \mathcal{A}$, $p, q \in \triangle(Z)$ are such that $\lambda A + (1-\lambda)\{q\} \succsim \lambda B + (1-\lambda)\{q\}$. Let $A' = \lambda A + (1-\lambda)\{q\}$, $B' = \lambda B + (1-\lambda)\{q\}$, and $\theta = (1-\lambda)(p-q)$. Note that $\theta \in \Theta$, $A' + \theta = \lambda A + (1-\lambda)\{p\} \in \mathcal{A}$, and $B' + \theta = \lambda A + (1-\lambda)\{p\} \in \mathcal{A}$. Hence, by TI, $\lambda A + (1-\lambda)\{p\} \succsim \lambda B + (1-\lambda)\{p\}$.

To see that IDD implies TI, assume that $A, B \in \mathcal{A}$ and $\theta \in \Theta$ are such that $A + \theta, B + \theta \in \mathcal{A}$ and $A \succsim B$. If $\theta = 0$, the conclusion of TI holds trivially, so assume that $\theta \neq 0$. Let $Z^- = \{z \in Z : \theta_z < 0\}$. Define θ^+, $\theta^- \in R^Z$ by $\theta_z^+ = \max\{0, \theta_z\}$ and $\theta_z^- = \max\{0, -\theta_z\}$ for any $z \in Z$. Then let $\kappa \equiv \sum_{z\in Z} \theta_z^+ = \sum_{z\in Z} \theta_z^- > 0$.

We will first show that for any $r \in A \cup B$,

$$0 \leq r_z - \theta_z^- \leq 1 - \kappa \quad \text{for all } z \in Z. \tag{13}$$

Note that for any $z \in Z^-$, $r_z - \theta_z^- = r_z + \theta_z \geq 0$ since $r + \theta \in \triangle(Z)$. Note also that if $z \notin Z^-$, then $r_z - \theta_z^- = r_z \geq 0$ since $\theta_z^- = 0$. So for any $z \in Z$,

$$\begin{aligned} 0 \leq r_z - \theta_z^- &\leq \left(1 - \sum_{z'\in Z^-\setminus\{z\}} r_{z'}\right) - \theta_z^- \leq \left(1 - \sum_{z'\in Z^-\setminus\{z\}} \theta_{z'}^-\right) - \theta_z^- \\ &= 1 - \kappa, \end{aligned}$$

[34]Note that TI implies its converse. Suppose $A + \theta \succsim B + \theta$. Then by TI, $A = (A + \theta) + (-\theta) \succsim (B + \theta) + (-\theta) = B$.

establishing Equation (13). Therefore, since $\theta \neq 0$, we have $0 < \kappa \leq 1$. Then $p \equiv \frac{1}{\kappa}\theta^+$, $q \equiv \frac{1}{\kappa}\theta^-$ are in $\triangle(Z)$, and $\theta = \kappa(p - q)$. There are two cases to consider:

First, consider the case of $\kappa < 1$. Define subsets A' and B' of $\mathbb{R}^Z$ by

$$A' \equiv \left\{ r' \in \mathbb{R}^Z : r' = \frac{1}{1-\kappa}(r - \theta^-) \text{ for some } r \in A \right\},$$

$$B' \equiv \left\{ r' \in \mathbb{R}^Z : r' = \frac{1}{1-\kappa}(r - \theta^-) \text{ for some } r \in B \right\}.$$

By Equation (13) and the definition of κ, we have that $A', B' \in \mathcal{A}$ and

$$(1-\kappa)A' + \kappa\{q\} = A \succsim B = (1-\kappa)B' + \kappa\{q\}. \tag{14}$$

Next, consider the $\kappa = 1$ case. By Equation (13) we have $r = \theta^- = q$ for any $r \in A \cup B$. Therefore, $A = B = \{q\}$, and hence Equation (14) holds for any choice of $A', B' \in \mathcal{A}$.

Since Equation (14) holds in each of the two cases above, we conclude by IDD that

$$A + \theta = (1-\kappa)A' + \kappa\{p\} \succsim (1-\kappa)B' + \kappa\{p\} = B + \theta.$$

Therefore, TI is satisfied. *Q.E.D.*

In light of Lemma 3, we will use IDD and TI interchangeably. Before proceeding, we define the following important subset of $\mathcal{A}^c$:

$$\mathcal{A}^\circ \equiv \{A \in \mathcal{A}^c : \forall \theta \in \Theta\ \exists \alpha > 0 \text{ such that } A + \alpha\theta \in \mathcal{A}^c\}. \tag{15}$$

Thus $\mathcal{A}^\circ$ contains menus that can be translated at least a little bit in the direction of any vector in Θ. It is easily verified that $\mathcal{A}^\circ$ is convex. In addition, the following result gives an alternative characterization of $\mathcal{A}^\circ$ along with some other important properties.

LEMMA 4: *The set $\mathcal{A}^\circ$ has the following properties*:

(i) $\mathcal{A}^\circ = \{A \in \mathcal{A}^c : \exists \varepsilon > 0$ *such that* $\forall p \in A, \forall z \in Z, p_z \geq \varepsilon\}$.

(ii) *Suppose* $p \in \triangle(Z)$ *is such that* $p_z > 0$ *for all* $z \in Z$. *Then for any* $A \in \mathcal{A}^c$ *and* $\lambda \in [0, 1)$, $\lambda A + (1-\lambda)\{p\} \in \mathcal{A}^\circ$.

(iii) $\mathcal{A}^\circ$ *is dense in* $\mathcal{A}^c$.

PROOF: (i) Let $\hat{\mathcal{A}}^\circ \equiv \{A \in \mathcal{A}^c : \exists \varepsilon > 0$ such that $\forall p \in A, \forall z \in Z, p_z \geq \varepsilon\}$. To see that $\hat{\mathcal{A}}^\circ \subset \mathcal{A}^\circ$, take any $A \in \hat{\mathcal{A}}^\circ$ and $\theta \in \Theta$. Let $\varepsilon > 0$ be such that $p_z \geq \varepsilon$ for all $p \in A$ and $z \in Z$. Choose $\alpha > 0$ sufficiently small to ensure that $\alpha \cdot \max_{z \in Z} |\theta_z| \leq \varepsilon$. Then $p_z + \alpha\theta_z \geq p_z - \varepsilon \geq 0$ for all $p \in A$ and $z \in Z$, so $A + \alpha\theta \in \mathcal{A}^c$. Thus $A \in \mathcal{A}^\circ$.

To see that $\mathcal{A}^\circ \subset \hat{\mathcal{A}}^\circ$, take any $A \in \mathcal{A}^\circ$. Fix any $z \in Z$ and take any $\theta \in \Theta$ such that $\theta_z = -1$. Then let $\alpha_z > 0$ be such that $A + \alpha_z \theta \in \mathcal{A}^c$, so for any $p \in A$, $p_z + \alpha_z \theta = p_z - \alpha_z \geq 0$. We obtain such an $\alpha_z > 0$ for every $z \in Z$, so let $\varepsilon \equiv \min_{z\in Z} \alpha_z > 0$. Then for any $p \in A$ and $z \in Z$, $p_z \geq \alpha_z \geq \varepsilon$, so $A \in \hat{\mathcal{A}}^\circ$.

(ii) Let $\varepsilon \equiv (1-\lambda)(\min_{z\in Z} p_z) > 0$. Then for any $q \in A$ and $z \in Z$, $\lambda q_z + (1-\lambda)p_z \geq \varepsilon$. Thus $\lambda A + (1-\lambda)\{p\} \in \mathcal{A}^\circ$ by part (i).

(iii) It is easily verified that for any $A \in \mathcal{A}^c$, $(1-1/n)A + (1/n)\{p\} \to A$ as $n \to \infty$. Hence $\mathcal{A}^\circ$ is dense in $\mathcal{A}^c$ by part (ii). *Q.E.D.*

We next define Lipschitz continuity.

DEFINITION 6: Given a metric space (X, d), a function $f : X \to \mathbb{R}$ is *Lipschitz continuous* if there is some real number K such that for every $x, y \in X$, $|f(x) - f(y)| \leq K d(x, y)$. The number K is called a *Lipschitz constant* of f.

We will construct a function $V : \mathcal{A}^c \to \mathbb{R}$ that represents $\succsim$ on $\mathcal{A}^c$ and has certain desirable properties. We next define the notion of translation linearity so as to present the main result of this section. Recall that the set of translations, denoted by Θ, is defined in Equation (12).

DEFINITION 7: Suppose that $V : \mathcal{A}^c \to \mathbb{R}$. Then V is *translation linear* if there exists $v \in \mathbb{R}^Z$ such that for all $A \in \mathcal{A}^c$ and $\theta \in \Theta$ with $A + \theta \in \mathcal{A}^c$, we have $V(A + \theta) = V(A) + v \cdot \theta$.

PROPOSITION 1: *If the preference $\succsim$ satisfies weak order, strong continuity, ACP, and IDD, then there exists a function $V : \mathcal{A}^c \to \mathbb{R}$ with the following properties:*

(i) *For any $A, B \in \mathcal{A}^c$, $A \succsim B \iff V(A) \geq V(B)$.*

(ii) *V is Lipschitz continuous, convex, and translation linear.*

(iii) *There exist $p, q \in \triangle(Z)$ such that $V(\{p\}) > V(\{q\})$.*

Moreover, if V and V' are two functions that satisfy (ii)–(iii) *and are ordinally equivalent in the sense that for any $A, B \in \mathcal{A}^c$, $V(A) \geq V(B) \iff V'(A) \geq V'(B)$, then there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V' = \alpha V + \beta$.*

First note that by taking the p^* and p_* from the L-continuity axiom, it follows that $\{p^*\} \succ \{p_*\}$. Thus part (iii) of Proposition 1 follows from part (i). The proof of the rest of the proposition is in Section S.3 of the Supplemental Material. In the remainder of the current section, we present an outline of the proof. Intuitively, the assumptions of strong continuity, ACP, and IDD (equivalently TI) on $\succsim$ play key roles in establishing Lipschitz continuity, convexity, and translation linearity of V, respectively.

Let $\mathcal{S} \equiv \{\{p\} : p \in \triangle(Z)\}$ be the set all of singleton sets in $\mathcal{A}^c$. Lemma S.5 in the Supplemental Material shows that given the assumptions of Proposition 1, $\succsim$ satisfies the von Neumann–Morgenstern axioms on $\mathcal{S}$. Therefore, there exists $v \in \mathbb{R}^Z$ such that for all $p, q \in \triangle(Z)$, $\{p\} \succsim \{q\}$ if and only if $v \cdot p \geq v \cdot q$. We will abuse notation and also treat v as a function $v : \mathcal{S} \to \mathbb{R}$ naturally defined by $v(\{p\}) = v \cdot p$. Note that v is translation linear since $v(\{p\} + \theta) = v(\{p\}) + v \cdot \theta$ whenever $p \in \triangle(Z)$, $\theta \in \Theta$, and $p + \theta \in \triangle(Z)$.

We want to extend v to a function V on $\mathcal{A}^c$ that represents $\succsim$ and is translation linear. The outline of the construction of the desired extension is the following: We first restrict attention to menus in $\mathcal{A}^\circ$, as defined in Equation (15). This restriction allows us to make extensive use of the translation invariance (TI) property. We construct a sequence of subsets of $\mathcal{A}^\circ$, starting with $\mathcal{A}^\circ \cap \mathcal{S}$, such that each set is contained in its successor set. We then extend v sequentially to each of these domains, while still representing $\succsim$ and preserving translation linearity (with respect to the vector v). The domain will grow to eventually contain all of the sets in $\mathcal{A}^\circ$, and we show how to extend it to all of $\mathcal{A}^c$ by continuity. Then we prove that the resulting function is translation linear, Lipschitz continuous, and convex.

As above, take p^* and p_* from the L-continuity axiom, and let $\theta^* \equiv p^* - p_*$. Define a sequence $\mathcal{A}_0, \mathcal{A}'_0, \mathcal{A}_1, \mathcal{A}'_1, \ldots$ of subsets of $\mathcal{A}^\circ$ inductively as follows: Let $\mathcal{A}_0 \equiv \mathcal{A}^\circ \cap \mathcal{S}$. By part (i) of Lemma 4, we have that $\mathcal{A}_0 = \{\{p\} : p \in \triangle(Z) \text{ and } \forall z \in Z, p_z > 0\}$. Define $\mathcal{A}'_i$ for all $i \geq 0$ by

$$\mathcal{A}'_i \equiv \{A \in \mathcal{A}^\circ : A \sim B \text{ for some } B \in \mathcal{A}_i\},$$

and define $\mathcal{A}_i$ for all $i \geq 1$ by

$$\mathcal{A}_i \equiv \{A \in \mathcal{A}^\circ : A = B + \alpha\theta^* \text{ for some } \alpha \in \mathbb{R}, B \in \mathcal{A}'_{i-1}\}.$$

Intuitively, we first extend $\mathcal{A}_0$ by including all $A \in \mathcal{A}^\circ$ that are viewed with indifference to some $B \in \mathcal{A}_0$. Then we extend to all translations by multiples of θ^*. We repeat the process, alternating between extension by indifference and extension by translation. Note that $\mathcal{A}_0 \subset \mathcal{A}'_0 \subset \mathcal{A}_1 \subset \mathcal{A}'_1 \subset \cdots$.

Figure 3 illustrates this construction for the special case of $Z = \{z_1, z_2\}$. In this case, closed and convex menus of lotteries over Z can be represented as ordered pairs in the triangle in Figure 3 (see the discussion in Section 5.2). In this figure, the set $\mathcal{A}_0$ is the diagonal of the triangle, and the set $\mathcal{A}'_0$ is the region labeled I. The combination of regions I and II is the set $\mathcal{A}_1$, and the combination of regions I, II, and III is the set $\mathcal{A}'_1$. One could continue in this fashion to obtain the remaining sets $\mathcal{A}_2, \mathcal{A}'_2, \ldots$.

We also define a sequence of functions $V_0, V'_0, V_1, V'_1, \ldots$ from these domains. That is, for all $i \geq 0$, $V_i : \mathcal{A}_i \to \mathbb{R}$ and $V'_i : \mathcal{A}'_i \to \mathbb{R}$. Define these functions recursively as follows:

(i) Let $V_0 \equiv v|_{\mathcal{A}_0}$.

(ii) For $i \geq 0$, if $A \in \mathcal{A}'_i$, then $A \sim B$ for some $B \in \mathcal{A}_i$, so define V'_i by $V'_i(A) \equiv V_i(B)$.

(iii) For $i \geq 1$, if $A \in \mathcal{A}_i$, then $A = B + \alpha\theta^*$ for some $\alpha \in \mathbb{R}$ and $B \in \mathcal{A}'_{i-1}$, so define V_i by $V_i(A) \equiv V'_{i-1}(B) + \alpha(v \cdot \theta^*)$.

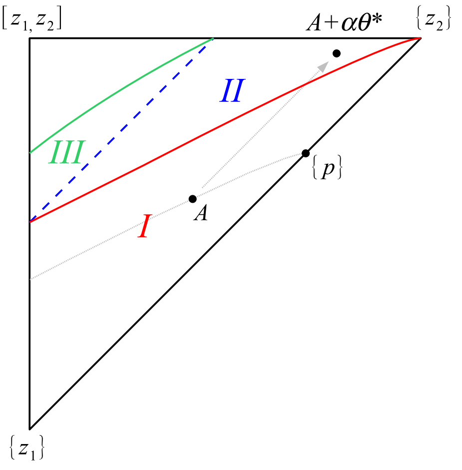

FIGURE 3.—Construction of $\mathcal{A}_i$ and $\mathcal{A}'_i$.

In a series of lemmas in Section S.3 in the Supplemental Material, we show that these are well defined functions which represent $\succsim$ on their domains and are translation linear.

C.2. *Application of Duality Results*

In this section, we apply the duality results from Section S.1 of the Supplemental Material to the function V constructed in Section C.1 to obtain the desired signed RFCC representation. Thus, in the remainder of this section assume that V satisfies (i)–(iii) from Proposition 1. Note that if $\succsim$ also satisfies monotonicity, then V is *monotone* in the sense that for all $A, B \in \mathcal{A}^c$ such that $A \subset B$, we have $V(A) \leq V(B)$. We explicitly assume monotonicity of V at the end of this section to prove the stronger representation of Theorem 8(B).

We follow a construction similar to that in DLR to obtain from V a function W whose domain is the set of support functions. Let $\mathcal{U}$ be defined as in Equation (6). For any $A \in \mathcal{A}^c$, the support function $\sigma_A : \mathcal{U} \to \mathbb{R}$ of A is defined by $\sigma_A(u) = \max_{p \in A} u \cdot p$. For a more complete introduction to support functions, see Rockafellar (1970) or Schneider (1993). Let $C(\mathcal{U})$ denote the set of continuous real-valued functions on $\mathcal{U}$. When endowed with the supremum norm $\|\cdot\|_\infty$, $C(\mathcal{U})$ is a Banach space. Define an order $\geq$ on $C(\mathcal{U})$ by $f \geq g$ if $f(u) \geq g(u)$ for all $u \in \mathcal{U}$. Let $\Sigma = \{\sigma_A \in C(\mathcal{U}) : A \in \mathcal{A}^c\}$. For any $\sigma \in \Sigma$, let

$$A_\sigma = \bigcap_{u \in \mathcal{U}} \left\{ p \in \triangle(Z) : u \cdot p = \sum_{z \in Z} u_z p_z \leq \sigma(u) \right\}.$$

LEMMA 5: (i) *For all $A \in \mathcal{A}^c$ and $\sigma \in \Sigma$, $A_{(\sigma_A)} = A$ and $\sigma_{(A_\sigma)} = \sigma$. Hence, σ is a bijection from $\mathcal{A}^c$ to Σ.*

(ii) *For all* $A, B \in \mathcal{A}^c$, $\sigma_{\lambda A+(1-\lambda)B} = \lambda\sigma_A + (1-\lambda)\sigma_B$.
(iii) *For all* $A, B \in \mathcal{A}^c$, $d_h(A, B) = \|\sigma_A - \sigma_B\|_\infty$.

PROOF: These are standard results that can be found in Rockafellar (1970) or Schneider (1993).[35] For instance, in Schneider (1993), part (i) follows from Theorem 1.7.1, part (ii) follows from Theorem 1.7.5, and part (iii) follows from Theorem 1.8.11. *Q.E.D.*

LEMMA 6: *Σ is convex and compact, and $0 \in \Sigma$.*

PROOF: The set Σ is convex by the convexity of $\mathcal{A}^c$ and part (ii) of Lemma 5. As discussed above, the set $\mathcal{A}^c$ is compact, and hence by parts (i) and (iii) of Lemma 5, Σ is a compact subset of the Banach space $C(\mathcal{U})$. Also, if we take $q = (1/|Z|, \dots, 1/|Z|) \in \triangle(Z)$, then $q \cdot u = 0$ for all $u \in \mathcal{U}$. Thus $\sigma_{\{q\}} = 0$, and hence $0 \in \Sigma$. *Q.E.D.*

Define the function $W : \Sigma \to \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$. Then, by part (i) of Lemma 5, $V(A) = W(\sigma_A)$ for all $A \in \mathcal{A}^c$. We say the function W is *monotone* if for all $\sigma, \sigma' \in \Sigma$ such that $\sigma \le \sigma'$, we have $W(\sigma) \le W(\sigma')$.

LEMMA 7: *W is convex and Lipschitz continuous with the same Lipschitz constant as V. If V is monotone, then W is monotone.*

PROOF: To see that W is convex, let $A, B \in \mathcal{A}^c$. Then

$$\begin{aligned} W(\lambda\sigma_A + (1-\lambda)\sigma_B) &= W\big(\sigma_{\lambda A+(1-\lambda)B}\big) = V(\lambda A + (1-\lambda)B) \\ &\le \lambda V(A) + (1-\lambda)V(B) \\ &= \lambda W(\sigma_A) + (1-\lambda)W(\sigma_B) \end{aligned}$$

by parts (i) and (ii) of Lemma 5 and convexity of V. The function W is Lipschitz continuous with the same Lipschitz constant as V by parts (i) and (iii) of Lemma 5. The function W inherits monotonicity from V because of the following fact which is easy to see from part (i) of Lemma 5: For all $A, B \in \mathcal{A}^c$, $A \subset B$ if and only if $\sigma_A \le \sigma_B$. *Q.E.D.*

We denote the set of continuous linear functionals on $C(\mathcal{U})$ (the dual space of $C(\mathcal{U})$) by $C(\mathcal{U})^*$. It is well known that $C(\mathcal{U})^*$ is the set of finite signed Borel measures on $\mathcal{U}$, where the duality is given by

$$\langle f, \mu\rangle = \int_{\mathcal{U}} f(u)\mu(du)$$

[35]The standard setting for support functions is the set of nonempty closed and convex subsets of $\mathbb{R}^n$. However, by imposing our normalizations on the domain of the support functions $\mathcal{U}$, the standard results are easily adapted to our setting of nonempty closed and convex subsets of $\triangle(Z)$.

for any $f \in C(\mathcal{U})$ and $\mu \in C(\mathcal{U})^*$.[36] For $\sigma \in \Sigma$, the *subdifferential* of W at σ is defined to be

$$\partial W(\sigma) = \{\mu \in C(\mathcal{U})^* : \langle \sigma' - \sigma, \mu \rangle \leq W(\sigma') - W(\sigma) \text{ for all } \sigma \in \Sigma\}.$$

The *conjugate* (or *Fenchel conjugate*) of W is the function $W^* : C(\mathcal{U})^* \to \mathbb{R} \cup \{+\infty\}$ defined by

$$W^*(\mu) = \sup_{\sigma \in \Sigma} [\langle \sigma, \mu \rangle - W(\sigma)].$$

There is an important duality between a convex function and its conjugate. We discuss this duality in detail in Section S.1 of the Supplemental Material. Lemma 8 summarizes certain properties of W^* that will be used in the sequel. Lemma S.1 provides a proof of these properties for general convex functions.

LEMMA 8: (i) *W^* is lower semicontinuous in the weak* topology*.
(ii) *$W(\sigma) \geq \langle \sigma, \mu \rangle - W^*(\mu)$ for all $\sigma \in \Sigma$ and $\mu \in C(\mathcal{U})^*$*.
(iii) *$W(\sigma) = \langle \sigma, \mu \rangle - W^*(\mu)$ if and only if $\mu \in \partial W(\sigma)$*.

We next define Σ_W, $\mathcal{N}_W$, and $\mathcal{M}_W$ as in Equations (S.3), (S.4), and (S.5), respectively, from Section S.1 of the Supplemental Material:

$$\Sigma_W = \{\sigma \in \Sigma : \partial W(\sigma) \text{ is a singleton}\},$$
$$\mathcal{N}_W = \{\mu \in C(\mathcal{U})^* : \mu \in \partial W(\sigma) \text{ for some } \sigma \in \Sigma_W\},$$
$$\mathcal{M}_W = \overline{\mathcal{N}_W},$$

where the closure is taken with respect to the weak* topology. We now apply Theorem S.1 in the Supplemental Material to the current setting.

LEMMA 9: *$\mathcal{M}_W$ is weak* compact, and for any weak* compact $\mathcal{M} \subset C(\mathcal{U})^*$,*

$$\mathcal{M}_W \subset \mathcal{M} \quad \iff \quad W(\sigma) = \max_{\mu \in \mathcal{M}} [\langle \sigma, \mu \rangle - W^*(\mu)] \quad \forall \sigma \in \Sigma.$$

PROOF: We simply need to verify that $C(\mathcal{U})$, Σ, and W satisfy the assumptions of Theorem S.1, that is, (i) $C(\mathcal{U})$ is a separable Banach space, (ii) Σ is a closed and convex subset of $C(\mathcal{U})$ containing the origin such that $\text{span}(\Sigma)$ is dense in $C(\mathcal{U})$, and (iii) $W : \Sigma \to \mathbb{R}$ is Lipschitz continuous and convex. Since $\mathcal{U}$

[36]Since $\mathcal{U}$ is a compact metric space, by the Riesz representation theorem (see Royden (1988, p. 357)), each continuous linear functional on $C(\mathcal{U})$ corresponds uniquely to a finite signed Baire measure on $\mathcal{U}$. Since $\mathcal{U}$ is a locally compact separable metric space, the Baire sets and the Borel sets of $\mathcal{U}$ coincide (see Royden (1988, p. 332)). Hence the sets of Baire and Borel finite signed measures also coincide.

is a compact metric space, $C(\mathcal{U})$ is separable (see Theorem 8.48 of Aliprantis and Border (1999)). By Lemma 6, Σ is a closed and convex subset of $C(\mathcal{U})$ containing the origin. Although the result is stated slightly differently, it is shown in Hörmander (1954) that $\text{span}(\Sigma)$ is dense in $C(\mathcal{U})$. This result is also proved in DLR. Finally, W is Lipschitz continuous and convex by Lemma 7. *Q.E.D.*

One consequence of Lemma 9 is that for all $\sigma \in \Sigma$,

$$W(\sigma) = \max_{\mu \in \mathcal{M}_W} [\langle \sigma, \mu \rangle - W^*(\mu)].$$

Therefore, for all $A \in \mathcal{A}^c$,

$$V(A) = \max_{\mu \in \mathcal{M}_W} \left(\int_{\mathcal{U}} \max_{p \in A} (u \cdot p) \mu(du) - W^*(\mu) \right).$$

The function W^* is lower semicontinuous by part (i) of Lemma 8, and $\mathcal{M}_W$ is compact by Lemma 9. It remains only to show that $\mathcal{M}_W$ is consistent and minimal, and that monotonicity of W implies each $\mu \in \mathcal{M}_W$ is positive.

Since V is translation linear, there exists $v \in \mathbb{R}^Z$ such that for all $A \in \mathcal{A}^c$ and $\theta \in \Theta$ with $A + \theta \in \mathcal{A}^c$, we have $V(A + \theta) = V(A) + v \cdot \theta$. The following result shows that a certain subset of $\mathcal{M}_W$ must agree with v in a way that will imply the consistency of this subset. In what follows, let $q = (1/|Z|, \dots, 1/|Z|) \in \triangle(Z)$ and let $\mathcal{A}^\circ \subset \mathcal{A}^c$ be defined as in Equation (15).

LEMMA 10: *If $A \in \mathcal{A}^\circ$ and $\mu \in \partial W(\sigma_A)$, then $\langle \sigma_{\{p\}}, \mu \rangle = v \cdot (p - q)$ for all $p \in \triangle(Z)$.*

PROOF: Fix any $A \in \mathcal{A}^\circ$ and $\mu \in \partial W(\sigma_A)$. We can apply the definition of the support function to $\theta \in \Theta$, so that $\sigma_{\{\theta\}}(u) = u \cdot \theta$ for $u \in \mathcal{U}$. It is easily verified that for any $A \in \mathcal{A}^c$ and $\theta \in \Theta$, $\sigma_{A+\theta} = \sigma_A + \sigma_{\{\theta\}}$.

We first prove that $\langle \sigma_{\{\theta\}}, \mu \rangle = v \cdot \theta$ for all $\theta \in \Theta$. Fix any $\theta \in \Theta$. Since $A \in \mathcal{A}^\circ$, there exists $\alpha > 0$ such that $A + \alpha\theta, A - \alpha\theta \in \mathcal{A}^c$. By the translation linearity of V, we have

$$\alpha(v \cdot \theta) = V(A + \alpha\theta) - V(A) = W(\sigma_{A+\alpha\theta}) - W(\sigma_A).$$

Since $\mu \in \partial W(\sigma_A)$, by part (iii) of Lemma 8, $W(\sigma_A) = \langle \sigma_A, \mu \rangle - W^*(\mu)$. Also, by part (ii) of the same lemma, $W(\sigma_{A+\alpha\theta}) \geq \langle \sigma_{A+\alpha\theta}, \mu \rangle - W^*(\mu)$. Therefore, we have

$$\alpha(v \cdot \theta) \geq \langle \sigma_{A+\alpha\theta}, \mu \rangle - \langle \sigma_A, \mu \rangle = \langle \sigma_{\{\alpha\theta\}}, \mu \rangle = \alpha \langle \sigma_{\{\theta\}}, \mu \rangle.$$

A similar argument can be used to show that

$$-\alpha(v \cdot \theta) = W(\sigma_{A-\alpha\theta}) - W(\sigma_A) \geq -\alpha \langle \sigma_{\{\theta\}}, \mu \rangle.$$

Hence, we have $\alpha(v \cdot \theta) = \alpha\langle\sigma_{\{\theta\}}, \mu\rangle$ or, equivalently, $v \cdot \theta = \langle\sigma_{\{\theta\}}, \mu\rangle$.

We now prove that $\langle\sigma_{\{p\}}, \mu\rangle = v \cdot (p - q)$ for all $p \in \triangle(Z)$. Since $\sum_z u_z = 0$ for $u \in \mathcal{U}$, we have $u \cdot q = 0$ for all $u \in \mathcal{U}$. Clearly, this implies that $\sigma_{\{q\}} = 0$, so that $\langle\sigma_{\{q\}}, \mu\rangle = 0$. For any $p \in \triangle(Z)$, $p - q \in \Theta$, so the above results imply

$$\langle\sigma_{\{p\}}, \mu\rangle = \langle\sigma_{\{p-q\}}, \mu\rangle + \langle\sigma_{\{q\}}, \mu\rangle = \langle\sigma_{\{p-q\}}, \mu\rangle = v \cdot (p - q),$$

which completes the proof. *Q.E.D.*

By part (ii) of Lemma 4, if $q = (1/|Z|, \dots, 1/|Z|)$, then $\lambda A + (1 - \lambda)\{q\} \in \mathcal{A}^\circ$ for any $A \in \mathcal{A}^c$ and $\lambda \in (0, 1)$. Therefore, we can use Lemma 10 and the continuity of W to prove the consistency of $\mathcal{M}_W$.

LEMMA 11: *If $\mu \in \mathcal{M}_W$, then $\langle\sigma_{\{p\}}, \mu\rangle = v \cdot (p - q)$ for all $p \in \triangle(Z)$.*

PROOF: Define $\mathcal{M} \subset \mathcal{M}_W$ by

$$\mathcal{M} \equiv \{\mu \in \mathcal{M}_W : \langle\sigma_{\{p\}}, \mu\rangle = v \cdot (p - q) \text{ for all } p \in \triangle(Z)\}.$$

It is easily verified that $\mathcal{M}$ is a closed subset of $\mathcal{M}_W$ and is therefore compact. We want to show $\mathcal{M}_W \subset \mathcal{M}$, which would imply $\mathcal{M} = \mathcal{M}_W$. By Lemma 9, we only need to verify that $W(\sigma) = \max_{\mu\in\mathcal{M}}[\langle\sigma, \mu\rangle - W^*(\mu)]$ for all $\sigma \in \Sigma$.

Let $\sigma \in \Sigma$ be arbitrary. For all $\lambda \in (0, 1)$, we have $\lambda A_\sigma + (1 - \lambda)\{q\} \in \mathcal{A}^\circ$. Note that $\sigma_{\lambda A_\sigma + (1-\lambda)\{q\}} = \lambda\sigma_{(A_\sigma)} + (1 - \lambda)\sigma_{\{q\}} = \lambda\sigma$. Therefore, Lemma 10 implies that for all $\lambda \in (0, 1)$, $\mathcal{M}_W \cap \partial W(\lambda\sigma) \subset \mathcal{M}$. By Lemma 9, there exists $\mu \in \mathcal{M}_W$ such that $W(\lambda\sigma) = \langle\lambda\sigma, \mu\rangle - W^*(\mu)$, which implies $\mu \in \partial W(\lambda\sigma)$ by part (iii) of Lemma 8. Thus, $\mathcal{M}_W \cap \partial W(\lambda\sigma) \neq \emptyset$.

Take any net $\{\lambda_d\}_{d\in D}$ such that $\lambda_d \to 1$, and let $\sigma_d \equiv \lambda_d\sigma$, so that $\sigma_d \to \sigma$. From the above arguments, for all $d \in D$ there exists $\mu_d \in \mathcal{M}_W \cap \partial W(\sigma_d) \subset \mathcal{M}$. Since $\mathcal{M}$ is weak* compact, every net in $\mathcal{M}$ has a convergent subnet. Without loss of generality, suppose the net itself converges, so that $\mu_d \xrightarrow{w^*} \mu$ for some $\mu \in \mathcal{M}$. By the definition of the subdifferential and the continuity of W, for any $\sigma' \in \Sigma$,

$$\begin{aligned}
\langle\sigma' - \sigma, \mu\rangle &= \lim_d \langle\sigma' - \sigma_d, \mu_d\rangle \\
&\leq \lim_d [W(\sigma') - W(\sigma_d)] \\
&= W(\sigma') - W(\sigma),
\end{aligned}$$

which implies $\mu \in \partial W(\sigma)$.[37] Hence, $W(\sigma) = \langle\sigma, \mu\rangle - W^*(\mu)$ by part (iii) of Lemma 8. Since $\sigma \in \Sigma$ was arbitrary, this completes the proof. *Q.E.D.*

[37]To establish the first equality in this equation, note that $\{\mu_d\}_{d\in D}$ is norm bounded by the compactness of $\mathcal{M}$ and Alaoglu's theorem (see Theorem 6.25 in Aliprantis and Border (1999)).

The consistency of $\mathcal{M}_W$ follows immediately from Lemma 11 since for any $\mu, \mu' \in \mathcal{M}_W$ and $p \in \triangle(Z)$, we have

$$\int_{\mathcal{U}} (u \cdot p)\mu(du) = \langle \sigma_{\{p\}}, \mu \rangle = v \cdot (p - q)$$
$$= \langle \sigma_{\{p\}}, \mu' \rangle = \int_{\mathcal{U}} (u \cdot p)\mu'(du).$$

We now prove the minimality of $\mathcal{M}_W$.

LEMMA 12: *$\mathcal{M}_W$ is minimal.*

PROOF: Suppose $\mathcal{M}' \subset \mathcal{M}_W$ is compact and $(\mathcal{M}', W^*|_{\mathcal{M}'})$ still represents $\succsim$. We will show that this implies $\mathcal{M}' = \mathcal{M}_W$.

Define $V' : \mathcal{A}^c \to \mathbb{R}$ as in Equation (7) for the representation $(\mathcal{M}', W^*|_{\mathcal{M}'})$, and define $W' : \Sigma \to \mathbb{R}$ by $W'(\sigma) = V'(A_\sigma)$. Then

$$W'(\sigma) = \max_{\mu \in \mathcal{M}'} [\langle \sigma, \mu \rangle - W^*(\mu)]$$

for all $\sigma \in \Sigma$. Note that V' satisfies (i)–(iii) from Proposition 1. Lipschitz continuity and translation linearity follow from Lemma S.2 in the Supplemental Material, and the other properties are immediate. Therefore, by the uniqueness part of Proposition 1, there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V' = \alpha V + \beta$, which implies $W' = \alpha W + \beta$. By singleton nontriviality, there exist $p_*, p^* \in \triangle(Z)$ such that $\{p^*\} \succ \{p_*\}$. Therefore, by Lemma 11, for any $\mu \in \mathcal{M}_W$,

$$\langle \sigma_{\{p^*\}} - \sigma_{\{p_*\}}, \mu \rangle = \langle \sigma_{\{p^*\}}, \mu \rangle - \langle \sigma_{\{p_*\}}, \mu \rangle = v \cdot (p^* - p_*) > 0.$$

We can therefore apply Proposition S.1 from the Supplemental Material with $\bar{x} = \sigma_{\{p^*\}} - \sigma_{\{p_*\}}$ to conclude that $\mathcal{M}' = \mathcal{M}_W$. Thus, $\mathcal{M}_W$ is minimal. *Q.E.D.*

We have now completed the proof of Theorem 8(A). To complete the proof of Theorem 8(B), note that $C(\mathcal{U})$ is a Banach lattice (see Aliprantis and Border (1999, p. 302)) and Σ has the property that $\sigma \vee \sigma' \in \Sigma$ for all $\sigma, \sigma' \in \Sigma$. Therefore, by Theorem S.2 from the Supplemental Material, if W is monotone, then each $\mu \in \mathcal{M}_W$ is positive.

Thus, there exists $K > 0$ such that $\|\mu_d\| \le K$ for all $d \in D$. Therefore,

$$\begin{aligned}
&|\langle \sigma' - \sigma, \mu \rangle - \langle \sigma' - \sigma_d, \mu_d \rangle| \\
&\quad \le |\langle \sigma' - \sigma, \mu \rangle - \langle \sigma' - \sigma, \mu_d \rangle| + |\langle \sigma' - \sigma, \mu_d \rangle - \langle \sigma' - \sigma_d, \mu_d \rangle| \\
&\quad = |\langle \sigma' - \sigma, \mu - \mu_d \rangle| + |\langle \sigma_d - \sigma, \mu_d \rangle| \\
&\quad \le |\langle \sigma' - \sigma, \mu - \mu_d \rangle| + \|\sigma_d - \sigma\| \|\mu_d\| \\
&\quad \le |\langle \sigma' - \sigma, \mu - \mu_d \rangle| + \|\sigma_d - \sigma\| K.
\end{aligned}$$

The right side of this inequality converges to zero since $\mu_d \xrightarrow{w^*} \mu$ and $\sigma_d \to \sigma$.

APPENDIX D: PROOF OF THEOREM 2

D.1. *CC ⇒ RFCC*

Assume there exists a CC representation $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$ such that V is given by Equation (4). Then the restriction of V to $\mathcal{A}^c$ is monotone and satisfies (ii) and (iii) in Proposition 1 in Appendix C.1. It is easy to see that V is monotone, convex, and translation linear, and that there exist $p, q \in \triangle(Z)$ such that $V(\{p\}) > V(\{q\})$. It remains only to show that V is Lipschitz continuous. Note that $K = \sum_{z\in Z} \mathbb{E}[|U_z|] > 0$ is finite since U is integrable. Let $\|\cdot\|$ denote the usual Euclidean norm in $\mathbb{R}^Z$. Let $\mathcal{G} \in \mathbf{G}$ and define $f_{\mathcal{G}} : \mathcal{A} \to \mathbb{R}$ by

$$f_{\mathcal{G}}(A) = \mathbb{E}\Big[\max_{p\in A} \mathbb{E}[U|\mathcal{G}] \cdot p\Big] - c(\mathcal{G}).$$

Let $A, B \in \mathcal{A}$. Given a state $\omega \in \Omega$, let p^* be a solution of $\max_{p\in A} \mathbb{E}[U|\mathcal{G}](\omega) \cdot p$. By definition of Hausdorff distance, there exists $q^* \in B$ such that $\|p^* - q^*\| \leq d_h(A, B)$. Then

$$\begin{aligned}
&\max_{p\in A} \mathbb{E}[U|\mathcal{G}](\omega) \cdot p - \max_{q\in B} \mathbb{E}[U|\mathcal{G}](\omega) \cdot q \\
&\quad = \mathbb{E}[U|\mathcal{G}](\omega) \cdot p^* - \max_{q\in B} \mathbb{E}[U|\mathcal{G}](\omega) \cdot q \\
&\quad \leq \mathbb{E}[U|\mathcal{G}](\omega) \cdot p^* - \mathbb{E}[U|\mathcal{G}](\omega) \cdot q^* \\
&\quad \leq \|\mathbb{E}[U|\mathcal{G}](\omega)\| \times \|p^* - q^*\| \\
&\quad \leq \|\mathbb{E}[U|\mathcal{G}](\omega)\| \times d_h(A, B).
\end{aligned}$$

Taking the expectation of the above inequality we obtain

$$f_{\mathcal{G}}(A) - f_{\mathcal{G}}(B) \leq \mathbb{E}\big[\|\mathbb{E}[U|\mathcal{G}]\|\big] d_h(A, B),$$

where

$$\begin{aligned}
\mathbb{E}\big[\|\mathbb{E}[U|\mathcal{G}]\|\big] &\leq \mathbb{E}\Big[\sum_{z\in Z} |\mathbb{E}[U_z|\mathcal{G}]|\Big] \leq \mathbb{E}\Big[\sum_{z\in Z} \mathbb{E}[|U_z||\mathcal{G}]\Big] \\
&= \sum_{z\in Z} \mathbb{E}[|U_z|] = K.
\end{aligned}$$

Hence $f_{\mathcal{G}}$ is Lipschitz continuous with a Lipschitz constant K that does not depend on $\mathcal{G}$. Since V is the pointwise maximum of $f_{\mathcal{G}}$ over $\mathcal{G} \in \mathbf{G}$, it is also Lipschitz continuous with the same Lipschitz constant K.

Since the restriction of V to $\mathcal{A}^c$ is monotone and satisfies (ii) and (iii) in Proposition 1 in Appendix C.1, the construction in Appendix C.2 implies that there exists an RFCC representation such that $V(A)$ is given by Equation (7) for all $A \in \mathcal{A}^c$. Since $V(A) = V(\text{co}(A))$ for all $A \in \mathcal{A}$ (which follows immediately from Equation (4)), this implies that $V(A)$ is given by Equation (7) for all $A \in \mathcal{A}$.

D.2. *RFCC* ⇒ *CC*

We begin by establishing a result in probability theory that will be useful later in the proof. Given a finite set $N = \{1, \dots, n\}$, let $\triangle(N) = \{\alpha \in [0,1]^N : \sum_{i\in N} \alpha_i = 1\}$ denote the simplex over N. In the following discussion, we will always assume without explicit mention that N is endowed with its discrete algebra consisting of all subsets of N and that $\triangle(N)$ is endowed with the Borel σ-algebra $\mathcal{B}$ induced by its Euclidean metric. The integral of an n-dimensional variable is used as a shorthand for the n-tuple of integrals of each dimension of the variable.

Suppose for a moment that the set N is a state space. Consider an individual who has uncertainty about the state $i \in N$ and observes a noisy signal that gives her additional information about i (a statistical experiment). Blackwell (1951, 1953) conveniently represented such a signal through the distribution over posterior beliefs over N that it induces.[38] The next result establishes the converse of this approach by representing a collection of probability measures over beliefs over N satisfying a certain consistency condition as conditional probabilities resulting from statistical experiments. More specifically, Lemma 13 shows that for any collection of probability measures $\{\pi_d\}_{d\in D}$ on $\triangle(N)$ having the same mean, there exists a probability space $(\Omega, \mathcal{F}, P)$ with the properties that (i) the state space is of the form $\Omega = N \times \Lambda$ and (ii) for each $d \in D$ there exists a sub-σ-algebra $\mathcal{G}_d$ of $\mathcal{F}$ such that the random vector $(P(\{i\} \times \Lambda | \mathcal{G}_d))_{i\in N}$, denoting the posterior over N conditional on $\mathcal{G}_d$, is distributed according to π_d.[39]

LEMMA 13: *Let $N = \{1, \dots, n\}$ and let $\Lambda = [\triangle(N)]^D$ for some an arbitrary index set D. Let $\mathcal{F}$ denote the product σ-algebra on $N \times \Lambda$ and let $\mathbf{G} = \{\mathcal{G}_d : d \in D\}$ where each $\mathcal{G}_d$ denotes the sub-σ-algebra of $\mathcal{F}$ consisting of events measurable with respect to the dth coordinate only, that is,*

$$\mathcal{G}_d = \{N \times E \times [\triangle(N)]^{D\setminus\{d\}} \in \mathcal{F} : E \in \mathcal{B}\}$$

for each $d \in D$. Let $\{\pi_d\}_{d\in D}$ be any collection of probability measures on $\triangle(N)$ that satisfies the following consistency condition for some $\alpha \in \triangle(N)$:

$$\int_{\triangle(N)} \beta \pi_d(d\beta) = \alpha \quad \forall d \in D.$$

[38] This approach is also used extensively in papers on mechanism design with information acquisition. For instance, see Bergemann and Välimäki (2002, 2006) and Persico (2000).

[39] Blackwell (1951) gave a proof of this result for the special case where there is only a single measure π and where $\alpha = (\frac{1}{n}, \dots, \frac{1}{n})$.

Then there exists a probability measure P on $(N \times \Lambda, \mathcal{F})$ such that the following statements hold:

(i) *The marginal of P on N agrees with α, that is, $P(\{i\} \times \Lambda) = \alpha_i$ for all $i \in N$.*

(ii) *The marginal of P on the dth coordinate of Λ agrees with π_d, that is,*

$$P\big(N \times E \times [\triangle(N)]^{D\setminus\{d\}}\big) = \pi_d(E) \quad \forall E \in \mathcal{B}.$$

(iii) *For any $d \in D$, the random vector $X^d : N \times \Lambda \to \triangle(N)$ defined by $X^d(j, \lambda) = \lambda(d)$ for all $(j, \lambda) \in N \times \Lambda$ satisfies*

$$P(\{i\} \times \Lambda | \mathcal{G}_d) = X_i^d$$

P-almost surely for all $i \in N$.[40]

PROOF: We first define a probability measure $P_d(\cdot|i)$ on $\triangle(N)$ for each $i \in N$ and $d \in D$. If $\alpha_i = 0$, fix the probability measure $P_d(\cdot|i)$ arbitrarily. If $\alpha_i > 0$, then let

$$P_d(E|i) = \frac{1}{\alpha_i} \int_E \beta_i \pi_d(d\beta)$$

for all $E \in \mathcal{B}$. The consistency condition on $\{\pi_d\}_{d\in D}$ and α implies that each $P_d(\cdot|i)$ is a probability measure. By Theorem 4.4.6 in Dudley (2002), for each $i \in N$ and nonempty finite subset $D' \subset D$, there exists a unique product probability measure $\prod_{d\in D'} P_d(\cdot|i)$ on $[\triangle(N)]^{D'}$ and its associated product σ-algebra. By the Kolmogorov extension theorem (see, e.g., Corollary 14.27 in Aliprantis and Border (1999)), there exists a unique extension $P(\cdot|i)$ of these finite product probability measures to $\Lambda = [\triangle(N)]^D$ and its associated product σ-algebra.

Define the probability measure P on $(N \times \Lambda, \mathcal{F})$ by

$$P(F) = \sum_{i\in N} \alpha_i P\big(\{\lambda \in \Lambda : (i, \lambda) \in F\}|i\big)$$

for all $F \in \mathcal{F}$. The marginal of P on N agrees with α by definition. Also, for any $d \in D$, $i \in N$, and $E \in \mathcal{B}$,

$$\begin{aligned} P\big(\{i\} \times E \times [\triangle(N)]^{D\setminus\{d\}}\big) &= \alpha_i P\big(E \times [\triangle(N)]^{D\setminus\{d\}}|i\big) \\ &= \alpha_i P_d(E|i) = \int_E \beta_i \pi_d(d\beta). \end{aligned} \tag{16}$$

[40]Note that $\mathcal{G}_d$ is the σ-algebra generated by the signal X^d. Thus, conditional on observing the signal X^d, the posterior over the first dimension of the state space is almost surely equal to the realization of the signal itself. For the case where $\alpha = (\frac{1}{n}, \ldots, \frac{1}{n})$, Blackwell (1951, 1953) refered to the distribution of such a signal as a *standard measure*.

Summing Equation (16) over $i \in N$ implies that the marginal of P on the dth coordinate of Λ agrees with π_d:

$$\begin{aligned} P\big(N \times E \times [\triangle(N)]^{D\setminus\{d\}}\big) &= \sum_{i\in N} P\big(\{i\} \times E \times [\triangle(N)]^{D\setminus\{d\}}\big) \\ &= \sum_{i\in N} \int_E \beta_i \pi_d(d\beta) \\ &= \int_E \Bigg[\sum_{i\in N} \beta_i\Bigg] \pi_d(d\beta) = \pi_d(E). \end{aligned}$$

To verify the final claim of the lemma, fix any $d \in D$ and $i \in N$. Then, for any $G = N \times E \times [\triangle(N)]^{D\setminus\{d\}} \in \mathcal{G}_d$,

$$\begin{aligned} \int_G X_i^d(j,\lambda) P(dj, d\lambda) &= \int_G \lambda_i(d) P(dj, d\lambda) = \int_E \beta_i \pi_d(d\beta) \\ &= P\big(\{i\} \times E \times [\triangle(N)]^{D\setminus\{d\}}\big) \\ &= P\big((\{i\} \times \Lambda) \cap G\big), \end{aligned}$$

where the second equality follows from the second claim of the lemma and the third equality follows from Equation (16). Hence, the claim holds by definition of conditional probability.[41] *Q.E.D.*

Using these results, we now complete the proof of the RFCC $\Rightarrow$ CC part of Theorem 2. Let $N = \{1, \dots, n\}$ for $n = |Z|$. Assume that there exists an RFCC representation $(\mathcal{M}, c)$ such that V is given by Equation (7). Since $\mathcal{M}$ is compact, there is $\kappa > 0$ such that $\mu(\mathcal{U}) \le \kappa$ for all $\mu \in \mathcal{M}$. The set $\kappa\mathcal{U}$ is compact and $(n-1)$-dimensional, which implies there exist affinely independent vectors $v^1, \dots, v^n \in \mathbb{R}^Z$ such that $\kappa\mathcal{U} \subset \mathrm{co}(\{v^1, \dots, v^n\})$. By affine independence of $v^1, \dots, v^n$, for all $u \in \mathrm{co}(\{v^1, \dots, v^n\})$, there exist unique coefficients (barycentric coordinates) $\gamma(u) = (\gamma_1(u), \dots, \gamma_n(u)) \in \triangle(N)$ such that $u = \gamma_1(u)v^1 + \cdots + \gamma_n(u)v^n$. The mapping $\gamma : \mathrm{co}(\{v^1, \dots, v^n\}) \to \triangle(N)$ is a continuous bijection.

In the first step of the proof, we transform each measure $\mu \in \mathcal{M}$ into a probability measure π_μ over $\triangle(N)$ such that the following statements hold:

(i) For every $\mu \in \mathcal{M}$ and $A \in \mathcal{A}$,

$$\int_{\mathcal{U}} \max_{p\in A} u(p)\mu(du) = \int_{\triangle(N)} \max_{p\in A} \Bigg[\sum_{i\in N} \beta_i v^i\Bigg] \cdot p\, \pi_\mu(d\beta). \tag{17}$$

[41] By the definition of conditional probability, a random variable Y is a version of $P(F|\mathcal{G}_d)$ for $F \in \mathcal{F}$ if (i) Y is $\mathcal{G}_d$-measurable and integrable and (ii) $\int_G Y(j,\lambda)P(dj, d\lambda) = P(F \cap G)$ for all $G \in \mathcal{G}_d$ (see, e.g., Billingsley (1995, p. 430)).

(ii) There exists $\alpha \in \triangle(N)$ such that for every $\mu \in \mathcal{M}$,

$$\int_{\triangle(N)} \beta \pi_\mu(d\beta) = \alpha. \tag{18}$$

To interpret Equation (17), suppose that v^i is the individual's expected-utility function over $\triangle(Z)$ conditional on state $i \in N$. In period 1, the individual is uncertain about her posterior belief $\beta = (\beta_1, \dots, \beta_n)$ over N. In period 2, she chooses $p \in A$, maximizing her ex post expected utility $\sum_{i\in N} \beta_i v^i$ determined by her posterior belief β. She believes that β is distributed according to π_μ, and hence the term on the right-hand side of the first condition is her expected utility before β is realized. Equation (18) corresponds to the consistency requirement that her prior belief, given by the expected value of the posterior belief, is the same for any probability measure in the collection $\{\pi_\mu\}_{\mu\in\mathcal{M}}$.

Take any $\mu \in \mathcal{M}$. To define π_μ, first consider the probability measure $\tilde{\mu}$ on $\mu(\mathcal{U})\mathcal{U}$ defined by $\tilde{\mu}(E) = \frac{1}{\mu(\mathcal{U})}\mu(\frac{1}{\mu(\mathcal{U})}E)$ for any measurable $E \subset \mu(\mathcal{U})\mathcal{U}$.[42] By a simple change of variables, we have

$$\int_{\mathcal{U}} \max_{p\in A}(u \cdot p)\mu(du) = \int_{\mu(\mathcal{U})\mathcal{U}} \max_{p\in A}(v \cdot p)\tilde{\mu}(dv). \tag{19}$$

The above equation reinterprets the integral expression in the RFCC representation in a probabilistic sense by rescaling the utility functions in $\mathcal{U}$, where the rescaling coefficient depends on the particular measure μ. Recall that $\mu(\mathcal{U})\mathcal{U} \subset \mathrm{co}(\kappa\mathcal{U}) \subset \mathrm{co}(\{v^1, \dots, v^n\})$. By affine independence of $v^1, \dots, v^n$, each point in $\mathrm{co}(\{v^1, \dots, v^n\})$ can be uniquely expressed as a convex combination of the vertices $v^1, \dots, v^n$. We can therefore interpret each such point as a probability measure on N where the probability of $i \in N$ is given by the coefficient of v^i in the unique convex combination. Hence, the probability measure $\tilde{\mu}$ can be identified with a probability measure π_μ over $\triangle(N)$ defined by $\pi_\mu = \tilde{\mu} \circ \gamma^{-1}$. Figure 4 illustrates this construction for the case where $n = 3$.

It is easy to see that Equation (17) is satisfied for every $A \in \mathcal{A}$.[43] In addition, letting $\alpha = \int_{\triangle(N)} \beta \pi_\mu(d\beta)$, we have[44]

$$\sum_{i\in N} \alpha_i v^i = \int_{\triangle(N)} \left[\sum_{i\in N} \beta_i v^i\right] \pi_\mu(d\beta) = \int_{\mathcal{U}} u\mu(du). \tag{20}$$

[42]Note that $\mu(\mathcal{U}) > 0$ since μ is positive and the RFCC representation satisfies consistency and singleton nontriviality.

[43]To see this, define a continuous function $g : \triangle(N) \to \mathbb{R}$ by $g(\beta) = \max_{p\in A}(\sum_{i\in N} \beta_i v^i) \cdot p$. Then $\int_{\triangle(N)} g(\beta)\pi_\mu(d\beta) = \int_{\mu(\mathcal{U})\mathcal{U}} (g \circ \gamma)(v)\tilde{\mu}(dv)$ by $\pi_\mu = \tilde{\mu} \circ \gamma^{-1}$ and the change of variables formula. This implies Equation (17) by $g(\gamma(v)) = \max_{p\in A} v \cdot p$ and Equation (19).

[44]To see the second equality, consider the continuous function $g : \triangle(N) \to \mathbb{R}^Z$ defined by $g(\beta) = \sum_{i\in N} \beta_i v^i$. Then $\int_{\triangle(N)} g(\beta)\pi_\mu(d\beta) = \int_{\mu(\mathcal{U})\mathcal{U}} (g \circ \gamma)(v)\tilde{\mu}(dv)$ by $\pi_\mu = \tilde{\mu} \circ \gamma^{-1}$ and the change of variables formula. This implies Equation (20) by $g(\gamma(v)) = v$ and Equation (19).

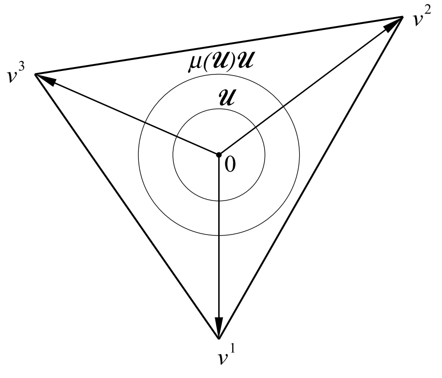

FIGURE 4.—Construction of the distribution over posteriors.

In particular, $\alpha = \gamma(\int_{\mathcal{U}} u\mu(du))$ is independent of the particular choice of μ by the consistency of the measures in $\mathcal{M}$. Therefore, Equation (18) is also satisfied for all $\mu \in \mathcal{M}$.

We next use Lemma 13 to express the probability measures $\{\pi_\mu\}_{\mu\in\mathcal{M}}$ over $\triangle(N)$ as distributions over posteriors resulting from statistical experiments. Let $\Omega = N \times \Lambda$, where Λ, $\mathcal{F}$, and $\mathbf{G}$ are as in Lemma 13 with $D = \mathcal{M}$ and $N = \{1, \dots, n\}$ for $n = |Z|$. That is, $\Lambda = [\triangle(N)]^{\mathcal{M}}$, $\mathcal{F}$ is the product σ-algebra on $\Omega = N \times [\triangle(N)]^{\mathcal{M}}$, and $\mathbf{G} = \{\mathcal{G}_\mu : \mu \in \mathcal{M}\}$, where

$$\mathcal{G}_\mu = \left\{N \times E \times [\triangle(N)]^{\mathcal{M}\setminus\{\mu\}} \in \mathcal{F} : E \in \mathcal{B}\right\}$$

for each $\mu \in \mathcal{M}$.

By Equation (18), the collection $\{\pi_\mu\}_{\mu\in\mathcal{M}}$ satisfies the consistency condition of Lemma 13, so there exists a probability measure P on $(\Omega, \mathcal{F})$ such that the following statements hold:

(i) $P(\{i\} \times \Lambda) = \alpha_i$ for all $i \in N$.

(ii) $P(N \times E \times [\triangle(N)]^{\mathcal{M}\setminus\{\mu\}}) = \pi_\mu(E)$ for all $E \in \mathcal{B}$ and $\mu \in \mathcal{M}$.

(iii) For any $\mu \in \mathcal{M}$, the random vector $X^\mu : \Omega \to \triangle(N)$ defined by $X^\mu(j, \lambda) = \lambda(\mu)$ for all $(j, \lambda) \in \Omega$ satisfies

$$P(\{i\} \times \Lambda | \mathcal{G}_\mu) = X_i^\mu$$

P-almost surely for all $i \in N$.

Let $U : \Omega \to \mathbb{R}^Z$ be defined by $U(i, \lambda) = v^i$ for every $i \in N$ and $\lambda \in \Lambda$. Fix any $\mu \in \mathcal{M}$. Defining $X^\mu : \Omega \to \triangle(N)$ by $X^\mu(j, \lambda) = \lambda(\mu)$ for all $(j, \lambda) \in \Omega$, condition (iii) on the measure P implies that

$$\mathbb{E}[U|\mathcal{G}_\mu] = \sum_{i\in N} P(\{i\} \times \Lambda | \mathcal{G}_\mu) v^i = \sum_{i\in N} X_i^\mu v^i \tag{21}$$

P-almost surely for $(j, \lambda) \in \Omega$.[45] Therefore, for any $A \in \mathcal{A}$,

$$\begin{aligned}
(22) \qquad \mathbb{E}\Big[\max_{p \in A} \mathbb{E}[U|\mathcal{G}_\mu] \cdot p\Big] &= \int_\Omega \bigg[\max_{p \in A}\bigg(\sum_{i \in N} \lambda_i(\mu) v^i\bigg) \cdot p\bigg] P(dj, d\lambda) \\
&= \int_{\triangle(N)} \bigg[\max_{p \in A}\bigg(\sum_{i \in N} \beta_i v^i\bigg) \cdot p\bigg] \pi_\mu(d\beta) \\
&= \int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du),
\end{aligned}$$

where the first equality follows from Equation (21), the second equality follows from condition (ii) on the measure P, and the third equality follows from Equation (17). By Equation (22) and defining $\tilde{c}(\mathcal{G}_\mu) = c(\mu)$, we have established that V can be expressed as

$$V(A) = \max_{\mathcal{G} \in \mathbf{G}} \Big\{\mathbb{E}\Big[\max_{p \in A} \mathbb{E}[U|\mathcal{G}] \cdot p\Big] - \tilde{c}(\mathcal{G})\Big\},$$

giving the desired CC representation.

APPENDIX E: PROOF OF THEOREM 4

In this section, we show that the uniqueness asserted in Theorem 4 applies not only to the RFCC representation, but to any signed RFCC representation (see Definition 5 in Appendix C). Throughout this section, we will continue to use the notation for support functions that was introduced in Appendix C.2. Suppose $(\mathcal{M}, c)$ and $(\mathcal{M}', c')$ are two signed RFCC representations for $\succsim$. Let $V : \mathcal{A}^c \to \mathbb{R}$ and $V' : \mathcal{A}^c \to \mathbb{R}$ be defined as in Equation (7) for these respective representations, and define $W : \Sigma \to \mathbb{R}$ and $W' : \Sigma \to \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$ and $W'(\sigma) = V'(A_\sigma)$.

We first show that $\mathcal{M} = \mathcal{M}_W$ and $c = W^*|_{\mathcal{M}_W}$, and likewise for $(\mathcal{M}', c')$. To see this, first note that by the definitions of V and W, we have

$$W(\sigma) = \max_{\mu \in \mathcal{M}} [\langle \sigma, \mu \rangle - c(\mu)] \quad \forall \sigma \in \Sigma.$$

Therefore, by Theorem S.3 in the Supplemental Material and the compactness of Σ, W is Lipschitz continuous and convex, $\mathcal{M}_W \subset \mathcal{M}$, and $W^*(\mu) = c(\mu)$ for all $\mu \in \mathcal{M}_W$. By Lemma 9 and the minimality of $\mathcal{M}$, this implies $\mathcal{M} = \mathcal{M}_W$ and $c = W^*|_{\mathcal{M}_W}$.

It is easily verified that both V and V' satisfy (i)–(iii) from Proposition 1. Therefore, by the uniqueness part of Proposition 1, there exist $\alpha > 0$ and $\beta \in \mathbb{R}$

[45] The first equality can be seen by applying Example 34.2 of Billingsley (1995, p. 446) to each coordinate of U, since U_z is the simple function $\sum_{i \in N} v_z^i I_{\{i\} \times \Lambda}$ for each $z \in Z$.

such that $V' = \alpha V - \beta$. This implies that $W' = \alpha W - \beta$. For any $\mu \in C(\mathcal{U})^*$ and $\sigma, \sigma' \in \Sigma$, note that

$$W(\sigma') - W(\sigma) \geq \langle \sigma' - \sigma, \mu \rangle$$
$$\iff \quad W'(\sigma') - W'(\sigma) \geq \langle \sigma' - \sigma, \alpha\mu \rangle$$

and hence $\partial W'(\sigma) = \alpha\, \partial W(\sigma)$. In particular, $\Sigma_{W'} = \Sigma_W$ and $\mathcal{N}_{W'} = \alpha \mathcal{N}_W$. Taking closures, we also have that $\mathcal{M}_{W'} = \alpha \mathcal{M}_W$. Since from our earlier arguments $\mathcal{M}' = \mathcal{M}_{W'}$ and $\mathcal{M} = \mathcal{M}_W$, we conclude that $\mathcal{M}' = \alpha \mathcal{M}$. Finally, let $\mu \in \mathcal{M}$. Then

$$\begin{aligned} c'(\alpha\mu) &= \sup_{\sigma \in \Sigma}[\langle \sigma, \alpha\mu \rangle - W'(\sigma)] = \alpha \sup_{\sigma \in \Sigma}[\langle \sigma, \mu \rangle - W(\sigma)] + \beta \\ &= \alpha c(\mu) + \beta, \end{aligned}$$

where the first and last equalities follow from our earlier findings that $c' = W'^*|_{\mathcal{M}_{W'}}$ and $c = W^*|_{\mathcal{M}_W}$. This concludes the proof of the theorem.

APPENDIX F: PROOFS OF RESULTS FROM SECTION 4

F.1. *Proof of Theorem 5*

(i) $\Rightarrow$ (ii) Suppose $\succsim_1$ has a lower cost of contemplation than $\succsim_2$. For any $p, q \in \triangle(Z)$, taking $A = \{q\}$ in Definition 3 yields $V_2(\{q\}) \geq V_2(\{p\}) \Longrightarrow V_1(\{q\}) \geq V_1(\{p\})$. Since the restrictions of V_1 and V_2 to singleton menus are nonconstant affine functions, it is a standard result that this condition implies there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V_2(\{p\}) = \alpha V_1(\{p\}) + \beta$ for all $p \in \triangle(Z)$ (see, e.g., Corollary B.3 of Ghirardato, Maccheroni, and Marinacci (2004)).

The preference $\succsim_2$ was assumed to be bounded above by singletons. Thus, there exists $z \in Z$ such that $\{\delta_z\} \succsim_2 A$ for all $A \in \mathcal{A}$. It is also easy to verify that V_2 being affine on singletons implies there exists some $z' \in Z$ such that $V_2(\{p\}) \geq V(\{\delta_{z'}\})$ for all $p \in \triangle(Z)$. Combined with monotonicity, this implies $A \succsim_2 \{\delta_{z'}\}$ for all $A \in \mathcal{A}$. Fix any $A \in \mathcal{A}$. Since $\{\delta_z\} \succsim_2 A \succsim_2 \{\delta_{z'}\}$, continuity implies there exists $\lambda \in [0, 1]$ such that $A \sim_2 \{\lambda\delta_z + (1-\lambda)\delta_{z'}\}$. Since $\succsim_1$ has a lower cost of contemplation than $\succsim_2$, this implies $A \succsim_1 \{\lambda\delta_z + (1-\lambda)\delta_{z'}\}$. Therefore,

$$\begin{aligned} V_2(A) &= V_2\big(\{\lambda\delta_z + (1-\lambda)\delta_{z'}\}\big) \\ &= \alpha V_1\big(\{\lambda\delta_z + (1-\lambda)\delta_{z'}\}\big) + \beta \leq \alpha V_1(A) + \beta. \end{aligned}$$

(ii) $\Rightarrow$ (i) Suppose there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V_2(\{p\}) = \alpha V_1(\{p\}) + \beta$ for all $p \in \triangle(Z)$, and $V_2 \leq \alpha V_1 + \beta$. Then $A \succsim_2 \{p\}$ implies

$$\alpha V_1(A) + \beta \geq V_2(A) \geq V_2(\{p\}) = \alpha V_1(\{p\}) + \beta,$$

which implies $A \succsim_1 \{p\}$. Thus, $\succsim_1$ has a lower cost of contemplation than $\succsim_2$.

(ii) $\Rightarrow$ (iii) Suppose there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V_2(\{p\}) = \alpha V_1(\{p\}) + \beta$ for all $p \in \triangle(Z)$, and $V_2 \leq \alpha V_1 + \beta$. Then, for any $\mu \in \mathbf{M}$,

$$\begin{aligned} c_2^*(\alpha\mu) &= \max_{A \in \mathcal{A}}(\alpha\langle\sigma_A, \mu\rangle - V_2(A)) \\ &\geq \max_{A \in \mathcal{A}}(\alpha\langle\sigma_A, \mu\rangle - \alpha V_1(A) - \beta) \\ &= \alpha \max_{A \in \mathcal{A}}(\langle\sigma_A, \mu\rangle - V_1(A)) - \beta = \alpha c_1^*(\mu) - \beta. \end{aligned}$$

(iii) $\Rightarrow$ (ii) Suppose there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V_2(\{p\}) = \alpha V_1(\{p\}) + \beta$ for all $p \in \triangle(Z)$, and $c_2^*(\alpha\mu) \geq \alpha c_1^*(\mu) - \beta$ for all $\mu \in \mathbf{M}$. Fix any $A \in \mathcal{A}$. Since $c_2^*(\mu) = c_2(\mu)$ for all $\mu \in \mathcal{M}_2$,[46] it follows from the definition of V_2 that there exists $\mu \in \mathcal{M}_2 \subset \mathbf{M}$ such that

$$\begin{aligned} V_2(A) &= \langle\sigma_A, \mu\rangle - c_2^*(\mu) \\ &\leq \langle\sigma_A, \mu\rangle - \alpha c_1^*\left(\frac{1}{\alpha}\mu\right) + \beta \\ &= \alpha\left[\left\langle\sigma_A, \frac{1}{\alpha}\mu\right\rangle - c_1^*\left(\frac{1}{\alpha}\mu\right)\right] + \beta \\ &\leq \alpha V_1(A) + \beta, \end{aligned}$$

where the last inequality follows from the definition of c_1^*.

F.2. *Proof of Corollary 1*

LEMMA 14: *Let $((\Omega, \mathcal{F}, P), \mathbf{G}, U, c)$ be any CC representation. Define V by Equation (4), define c^* by Equation (9), and let $\gamma = \frac{1}{|Z|}\sum_{z \in Z}\mathbb{E}[U_z]$. Then the following statements hold*:

(i) *$c^*(\mu_\mathcal{G}) \leq c(\mathcal{G}) - \gamma$ for any $\mathcal{G} \in \mathbf{G}$.*

(ii) *$c^*(\mu_\mathcal{G}) = c(\mathcal{G}) - \gamma$ if and only if $\mathcal{G}$ solves Equation (4) for some $A \in \mathcal{A}$.*

PROOF: (i) Note first that for any $A \in \mathcal{A}$ and $\mathcal{G} \in \mathbf{G}$,

$$V(A) \geq \mathbb{E}\Big[\max_{p \in A}\mathbb{E}[U|\mathcal{G}] \cdot p\Big] - c(\mathcal{G}) = \langle\sigma_A, \mu_\mathcal{G}\rangle + \gamma - c(\mathcal{G}),$$

where the inequality follows from Equation (4) and the equality follows from Lemma 1. Therefore, $\langle\sigma_A, \mu_\mathcal{G}\rangle - V(A) \leq c(\mathcal{G}) - \gamma$, implying by the definition of c^* that $c^*(\mu_\mathcal{G}) \leq c(\mathcal{G}) - \gamma$.

[46]That $c = c^*|_\mathcal{M}$ for any RFCC representation $(\mathcal{M}, c)$ follows from the observations made in Appendix E regarding conjugate convex functionals since c^* is precisely the restriction of the functional W^* to the set $\mathbf{M}$.

(ii) Suppose that $\mathcal{G} \in \mathbf{G}$ solves Equation (4) for some $A \in \mathcal{A}$. Then, by Equation (4) and Lemma 1, $V(A) = \langle \sigma_A, \mu_{\mathcal{G}} \rangle + \gamma - c(\mathcal{G})$. Along with the definition of c^*, this implies that $c^*(\mu_{\mathcal{G}}) \geq \langle \sigma_A, \mu_{\mathcal{G}} \rangle - V(A) = c(\mathcal{G}) - \gamma$. By part (i), we have $c^*(\mu_{\mathcal{G}}) = c(\mathcal{G}) - \gamma$.

Conversely, suppose that $c^*(\mu_{\mathcal{G}}) = c(\mathcal{G}) - \gamma$. Then taking $A \in \mathcal{A}$ such that $\langle \sigma_A, \mu_{\mathcal{G}} \rangle - V(A) = c^*(\mu_{\mathcal{G}})$, Lemma 1 implies

$$V(A) = \langle \sigma_A, \mu_{\mathcal{G}} \rangle + \gamma - c(\mathcal{G}) = \mathbb{E}\Big[\max_{p \in A} \mathbb{E}[U|\mathcal{G}] \cdot p\Big] - c(\mathcal{G}).$$

Thus, $\mathcal{G}$ solves Equation (4) for the menu A. *Q.E.D.*

We now complete the proof of Corollary 1. Note for each $i = 1, 2$, by Theorem 2, there exists an RFCC representation such that V_i is given by Equation (7). Therefore, the implications (i) $\Leftrightarrow$ (ii) and (ii) $\Rightarrow$ (iii) follow from Theorem 5.

To see (iii) $\Rightarrow$ (ii), suppose there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V_2(\{p\}) = \alpha V_1(\{p\}) + \beta$ for all $p \in \triangle(Z)$, and $c_2^*(\mu_{\mathcal{G}_2}) \geq \alpha c_1^*(\frac{1}{\alpha}\mu_{\mathcal{G}_2}) - \beta$ for all $\mathcal{G}_2 \in \mathbf{G}_2$. Let $A \in \mathcal{A}$ and define $\beta_2 = \frac{1}{|Z|}\sum_{z \in Z} \mathbb{E}^{P_2}[U_{2,z}]$. By Equation (4), there exists $\mathcal{G}_2 \in \mathbf{G}_2$ such that

$$\begin{aligned}
V_2(A) &= \mathbb{E}^{P_2}\Big[\max_{p \in A} \mathbb{E}^{P_2}[U_2|\mathcal{G}_2] \cdot p\Big] - c_2(\mathcal{G}_2) \\
&= \langle \sigma_A, \mu_{\mathcal{G}_2} \rangle + \beta_2 - c_2(\mathcal{G}_2) \\
&\leq \langle \sigma_A, \mu_{\mathcal{G}_2} \rangle - c_2^*(\mu_{\mathcal{G}_2}) \\
&\leq \langle \sigma_A, \mu_{\mathcal{G}_2} \rangle - \alpha c_1^*\left(\frac{1}{\alpha}\mu_{\mathcal{G}_2}\right) + \beta \\
&= \alpha\left[\left\langle \sigma_A, \frac{1}{\alpha}\mu_{\mathcal{G}_2} \right\rangle - c_1^*\left(\frac{1}{\alpha}\mu_{\mathcal{G}_2}\right)\right] + \beta \\
&\leq \alpha V_1(A) + \beta,
\end{aligned}$$

where the second equality follows from Lemma 1, the first inequality follows from part (i) of Lemma 14, the second inequality follows from our assumptions on c_1^* and c_2^*, and the last inequality follows from the definition of c_1^*.

APPENDIX G: PROOFS OF RESULTS FROM SECTION 5

G.1. *Proof of Theorem 6*

In this section, we show that for any signed RFCC representation (see Definition 5 in Appendix C)—in particular, for any RFCC representation—strong IDD is equivalent to a constant cost function. The necessity of strong IDD is straightforward and left to the reader. For sufficiency, suppose V is defined by

Equation (7) for a signed RFCC representation $(\mathcal{M}, c)$ for the preference $\succsim$ and that $\succsim$ satisfies strong IDD.

Lemma S.15 in the Supplemental Material shows that for any $A \in \mathcal{A}$, $p \in \triangle(Z)$, and $\alpha \in [0, 1]$,[47]

$$V(\alpha A + (1-\alpha)\{p\}) = \alpha V(A) + (1-\alpha)V(\{p\}). \tag{24}$$

As in Appendix C.2, define $W : \Sigma \to \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$. By Equation (24) and parts (i) and (ii) of Lemma 5, for any $A \in \mathcal{A}$, $p \in \triangle(Z)$, and $\alpha \in [0, 1]$,

$$W\big(\alpha\sigma_A + (1-\alpha)\sigma_{\{p\}}\big) = \alpha W(\sigma_A) + (1-\alpha)W\big(\sigma_{\{p\}}\big).$$

It was shown in Appendix E that for any signed RFCC representation $(\mathcal{M}, c)$, defining W as we have here gives $\mathcal{M} = \mathcal{M}_W$ and $c = W^*|_{\mathcal{M}_W}$. In particular, W satisfies

$$W(\sigma) = \max_{\mu \in \mathcal{M}_W} [\langle \sigma, \mu\rangle - W^*(\mu)]. \tag{25}$$

Therefore, it suffices to show that W^* is constant on $\mathcal{M}_W$. Let $\bar{w} = \min_{\mu\in\mathcal{M}_W} W^*(\mu)$. Note that this minimum is well defined since W^* is lower semicontinuous and $\mathcal{M}_W$ is compact. Let $\bar{\mu} \in \mathcal{M}_W$ be a minimizing measure, so that $W^*(\bar{\mu}) = \bar{w}$.

We first show that $W^*(\mu) = \bar{w}$ for all $\mu \in \mathcal{N}_W$. Let $\mu \in \mathcal{N}_W$ be arbitrary. By the definition of $\mathcal{N}_W$ and Lemma 8, there exists some $A \in \mathcal{A}$ such that μ is the unique maximizer of Equation (25) at σ_A. That is, $W(\sigma_A) = \langle \sigma_A, \mu\rangle - W^*(\mu) > \langle \sigma_A, \mu'\rangle - W^*(\mu')$ for any $\mu' \in \mathcal{M}_W$, $\mu' \neq \mu$. Now, for any $p \in \triangle(Z)$ and $\alpha \in (0, 1)$, choose $\mu' \in \mathcal{M}_W$ that maximizes Equation (25) at $\alpha\sigma_A + (1-\alpha)\sigma_{\{p\}}$. Then

$$\begin{aligned}
\alpha W(\sigma_A) + (1-\alpha)W\big(\sigma_{\{p\}}\big) &= W\big(\alpha\sigma_A + (1-\alpha)\sigma_{\{p\}}\big)\\
&= \big\langle \alpha\sigma_A + (1-\alpha)\sigma_{\{p\}}, \mu'\big\rangle - W^*(\mu')
\end{aligned}$$

[47]To have an intuition for Equation (24), suppose there exist alternatives $z, z' \in Z$ such that $\{\delta_z\} \succsim A \succsim \{\delta_{z'}\}$ for any $A \in \mathcal{A}$. It is easy to see that under this simplifying assumption, every menu is indifferent to a singleton menu. It is also easily verified that for any signed RFCC representation, the consistency of the measures implies that V is affine on singleton menus:

$$V(\alpha\{q\} + (1-\alpha)\{p\}) = \alpha V(\{q\}) + (1-\alpha)V(\{p\}) \quad \forall p, q \in \triangle(Z). \tag{23}$$

Let $A \in \mathcal{A}$, $p \in \triangle(Z)$, and $\alpha \in [0, 1]$. By the simplifying assumption, there exists $q \in \triangle(Z)$ such that $A \sim \{q\}$. Then

$$\begin{aligned}
V(\alpha A + (1-\alpha)\{p\}) &= V(\alpha\{q\} + (1-\alpha)\{p\}) \quad \text{(by strong IDD)}\\
&= \alpha V(\{q\}) + (1-\alpha)V(\{p\}) \quad \text{(by Equation (23))}\\
&= \alpha V(A) + (1-\alpha)V(\{p\}).
\end{aligned}$$

$$= \alpha[\langle \sigma_A, \mu' \rangle - W^*(\mu')] + (1-\alpha)\big[\big\langle \sigma_{\{p\}}, \mu' \big\rangle - W^*(\mu')\big].$$

Since $\langle \sigma_A, \mu' \rangle - W^*(\mu') \leq W(\sigma_A)$ and $\langle \sigma_{\{p\}}, \mu' \rangle - W^*(\mu') \leq W(\sigma_{\{p\}})$, the above equation implies that we must in fact have $\langle \sigma_A, \mu' \rangle - W^*(\mu') = W(\sigma_A)$ and $\langle \sigma_{\{p\}}, \mu' \rangle - W^*(\mu') = W(\sigma_{\{p\}})$. By the choice of A, the former implies $\mu' = \mu$. Therefore, the latter implies

$$\big\langle \sigma_{\{p\}}, \mu \big\rangle - W^*(\mu) = W\big(\sigma_{\{p\}}\big) \geq \big\langle \sigma_{\{p\}}, \bar{\mu} \big\rangle - \bar{w}.$$

Consistency implies $\langle \sigma_{\{p\}}, \mu \rangle = \langle \sigma_{\{p\}}, \bar{\mu} \rangle$ and therefore $W^*(\mu) \leq \bar{w}$. Since $\bar{w}$ is the minimum of W^* on $\mathcal{M}_W$, we have $W^*(\mu) = \bar{w}$.

The proof is completed by showing that $W^*(\mu) = \bar{w}$ for all $\mu \in \mathcal{M}_W$. If $\mu \in \mathcal{M}_W$, then there exists a net $\{\mu_d\}_{d \in D}$ in $\mathcal{N}_W$ such that $\mu_d \overset{w^*}{\to} \mu$. Since each μ_d is in $\mathcal{N}_W$, our previous arguments imply that $W^*(\mu_d) = \bar{w}$. Since W^* is lower semicontinuous, it follows that $W^*(\mu) \leq \liminf_d W^*(\mu_d) = \bar{w}$. Since $\bar{w}$ is minimal, we have $W^*(\mu) = \bar{w}$.

G.2. *Proof of Corollary 2*

(ii) $\Rightarrow$ (i) Let $\mathbf{G}' = \{\mathcal{G} \in \mathbf{G} : c(\mathcal{G}) \leq k\}$ and let $c' : \mathbf{G}' \to \mathbb{R}$ be any constant function. Then $((\Omega, \mathcal{F}, P), \mathbf{G}', U, c')$ is a CC representation for $\succsim$. Hence, Theorem 1 implies that $\succsim$ satisfies weak order, strong continuity, ACP, and monotonicity. It is easily verified that since c' is constant, $\succsim$ also satisfies strong IDD.

(i) $\Rightarrow$ (ii) First, apply Theorem 3 to conclude that $\succsim$ has an RFCC representation $(\mathcal{M}, c)$. Then, by Theorem 6, strong IDD implies that c is constant. Theorem 2 implies there is a CC representation $((\Omega, \mathcal{F}, P), \mathbf{G}, U, \tilde{c})$ that gives the same value function for menus V as the RFCC representation $(\mathcal{M}, c)$. Moreover, since c is constant, it is immediate from the construction in the proof of Theorem 2 that $\tilde{c}$ can be taken to be constant. If we choose $k \in \mathbb{R}$ larger than this constant value, then the function V defined by Equation (11) for these parameters represents $\succsim$.

G.3. *Proof of Corollary 3*

If: The necessity of weak order, strong continuity, and monotonicity follows from Theorem 1. Since $c(\mathcal{F}) = \min_{\mathcal{G} \in \mathbf{G}} c(\mathcal{G})$, $\mathcal{F}$ is an optimal contemplation strategy for any menu. Thus, for any $A \in \mathcal{A}$,

$$V(A) = \mathbb{E}\Big[\max_{p \in A} \mathbb{E}[U|\mathcal{F}] \cdot p\Big] - c(\mathcal{F}).$$

This implies that V is affine, and hence $\succsim$ satisfies independence.

Only if: By Theorem 7, $\succsim$ has an RFCC representation $(\mathcal{M}, c)$ in which $\mathcal{M} = \{\mu\}$ for some finite Borel measure μ. Since μ is positive, we can define a Borel probability measure on $\mathcal{U}$ by $P = \frac{\mu}{\mu(\mathcal{U})}$. Let $\Omega = \mathcal{U}$, let $\mathcal{F}$ be the Borel σ-algebra on $\mathcal{U}$, and let $\mathbf{G} = \{\mathcal{F}\}$. If we define $U : \mathcal{U} \to \mathbb{R}^Z$ by $U(u) = u$, then for any $A \in \mathcal{A}$,

$$\mathbb{E}\Big[\max_{p\in A} \mathbb{E}[U|\mathcal{F}] \cdot p\Big] = \mathbb{E}\Big[\max_{p\in A} U \cdot p\Big] = \frac{1}{\mu(\mathcal{U})} \int_{\mathcal{U}} \max_{p\in A} u(p)\mu(du).$$

Therefore, taking $c(\mathcal{F})$ to be any real number, we have the desired CC representation for $\succsim$.

REFERENCES

ALIPRANTIS, C., AND K. BORDER (1999): *Infinite Dimensional Analysis*. Berlin, Germany: Springer-Verlag. [1323-1325,1328]

BERGEMANN, D., AND J. VÄLIMÄKI (2002): "Information Acquisition and Efficient Mechanism Design," *Econometrica*, 70, 1007–1033. [1327]

——— (2006): "Information in Mechanism Design," in *Proceedings of the 9th World Congress of the Econometric Society*, ed. by R. Blundell, W. Newey, and T. Persson. Cambridge: Cambridge University Press, 186–221. [1327]

BILLINGSLEY, P. (1995): *Probability and Measure*. New York: Wiley. [1312,1329,1332]

BLACKWELL, D. (1951): "Comparisons of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley and Los Angeles: University of California Press, 93–102. [1327,1328]

——— (1953): "Equivalent Comparisons of Experiments," *The Annals of Mathematical Statistics*, 24, 265–272. [1327,1328]

CONLISK, J. (1996): "Why Bounded Rationality," *Journal of Economic Literature*, 34, 669–700. [1310]

DEKEL, E., B. LIPMAN, AND A. RUSTICHINI (2001): "Representing Preferences With a Unique Subjective State Space," *Econometrica*, 69, 891–934. [1289]

——— (2008): "Temptation-Driven Preferences," *Review of Economic Studies*, 76, 937–971. [1292,1314]

DEKEL, E., B. LIPMAN, A. RUSTICHINI, AND T. SARVER (2007): "Representing Preferences With a Unique Subjective State Space: Corrigendum," *Econometrica*, 75, 591–600. [1290]

DUDLEY, R. (2002): *Real Analysis and Probability*. Cambridge: Cambridge University Press. [1328]

EPSTEIN, L. G., M. MARINACCI, AND K. SEO (2007): "Coarse Contingencies and Ambiguity," *Theoretical Economics*, 2, 355–394. [1310]

ERGIN, H. (2003): "Costly Contemplation," Mimeo, Washington University in Saint Louis. [1289, 1293,1300]

ERGIN, H., AND T. SARVER (2010): "Supplement to 'A Unique Costly Contemplation Representation'," *Econometrica Supplemental Material*, 78, http://www.econometricsociety.org/ecta/Supmat/7801_proofs.pdf. [1288]

GHIRARDATO, P., F. MACCHERONI, AND M. MARINACCI (2004): "Differentiating Ambiguity and Ambiguity Attitude," *Journal of Economic Theory*, 118, 133–173. [1333]

GILBOA, I., AND D. SCHMEIDLER (1989): "Maxmin Expected Utility With Non-Unique Priors," *Journal of Mathematical Economics*, 18, 141–153. [1304,1310]

GREEN, J. (1987): "'Making Book Against Oneself,' the Independence Axiom and Nonlinear Utility Theory," *The Quarterly Journal of Economics*, 102, 785–796. [1306]

GUL, F., AND W. PESENDORFER (2001): "Temptation and Self-Control," *Econometrica*, 69, 1403–1435. [1292,1314]
HÖRMANDER, L. (1954): "Sur la Fonction d'Appui des Ensembles Convexes dans un Espace Localement Convexe," *Arkiv för Matematik*, 3, 181–186. [1323]
HYOGO, K. (2007): "A Subjective Model of Experimentation," *Journal of Economic Theory*, 133, 316–330. [1300]
KARNI, E. (1993): "A Definition of Subjective Probabilities With State-Dependent Preferences," *Econometrica*, 61, 187–198. [1296]
KREPS, D. (1979): "A Representation Theorem for Preference for Flexibility," *Econometrica*, 47, 565–578. [1289,1292,1293,1296]
LIPMAN, B. (1991): "How to Decide How to Decide How to ...: Modeling Limited Rationality," *Econometrica*, 59, 1105–1125. [1310]
——— (1995): "Information Processing and Bounded Rationality: A Survey," *The Canadian Journal of Economics*, 28, 42–67. [1311]
MACCHERONI, F., M. MARINACCI, AND A. RUSTICHINI (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 1447–1498. [1288,1292, 1304]
MAS-COLELL, A. (1986): "The Price Equilibrium Existence Problem in Topological Vector Lattices," *Econometrica*, 54, 1039–1054. [1290]
MOSSIN, J. (1969): "A Note on Uncertainty and Preferences in a Temporal Context," *The American Economic Review*, 59, 172–174. [1306]
MUNKRES, J. (2000): *Topology*. Upper Saddle River, NJ: Prentice Hall. [1315]
OLSZEWSKI, W. (2007): "Preferences Over Sets of Lotteries," *Review of Economic Studies*, 74, 567–595. [1308]
PERSICO, N. (2000): "Information Acquisition in Auctions," *Econometrica*, 68, 135–148. [1327]
ROCKAFELLAR, R. T. (1970): *Convex Analysis*. Princeton, NJ: Princeton University Press. [1320, 1321]
ROYDEN, H. L. (1988): *Real Analysis*. Englewood Cliffs, NJ: Prentice Hall. [1322]
SARVER, T. (2008): "Anticipating Regret: Why Fewer Options May Be Better," *Econometrica*, 76, 263–305. [1290,1292,1314]
SCHNEIDER, R. (1993): *Convex Bodies: The Brunn–Minkowski Theory*. Cambridge: Cambridge University Press. [1315,1320,1321]

Dept. of Economics, Washington University in Saint Louis, Campus Box 1208, Saint Louis, MO 63130, U.S.A.; hergin@artsci.wustl.edu

and

Dept. of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, U.S.A.; tsarver@northwestern.edu.

Manuscript received March, 2008; final revision received February, 2010.