A Class of Multi-Prior Preferences

Tan Wang*

First Version: July, 2001

This Version: January, 2003

Abstract

We axiomatize a new class of multi-prior preferences. The unique feature of this class of preferences is that it allows for the role of a reference probability measure in decision-making under uncertainty. The class of preferences has a tractable representation. It takes the form of minimization, over a set of priors, of an expected utility plus a penalty function that penalizes deviation from the reference probability measure. The preference reduces to the standard expected utility when there is no uncertainty. The paper also discusses the potential applications for the axiomatized preferences.

*Tan Wang is with the Faculty of Commerce and Business Administration, University of British Columbia, Vancouver, British Columbia, Canada, V6T 1Z2. tan.wang@commerce.ubc.ca. The author thank Larry Epstein, Marciano Siniscalchi and Raman Uppal for helpful comments. The paper was presented at the 2003 American Economic Association meetings. The author is grateful to the Social Sciences and Humanities Research Council of Canada for financial support. An earlier version of the paper was titled "Two Classes of Multi-Prior Preferences".

1 Introduction

It has long been argued that there is a qualitative difference between a risky situation where the objective/subjective probability law is known and an uncertainty situation where that law is not known (Knight (1921)). However, it is only recently that preferences with tractable analytic representation have been axiomatized that can distinguish between risk and uncertainty.[1] Adding to that literature, this paper axiomatizes a new class of multi-prior preferences. The unique feature of this class of preferences is that it allows for the role of a reference probability measure in decision-making under uncertainty.

To motivate the preferences to be axiomatized, it is useful to consider a variation of the well-known Ellersberg experiment (Ellsberg (1961)). The setting is the same as the original Ellsberg experiment: there are two urns. In the first urn there are 50 red balls and 50 white balls. In the second urn, there are 100 red and white balls but the proportion of red balls is unknown. In the original Ellsberg experiment, this is the information that the subject has. In the variation of the experiment the subjects are further told that the second urn has been sampled one hundred times in which 45 red balls were drawn. There are two sets of bets. The first set of bets is (a) winning $100 if a random draw from the first urn produces a red ball and zero otherwise; and (b) winning $100 if a random draw from the first urn produces a white ball and zero otherwise. The second set is similar, but the bets are based on the second urn. The subject is asked to give his preference ordering of the four bets.

In the original Ellsberg experiment, it has been found that most people are indifferent between the two bets in the first set and also between the two bets in the second set. However, they prefer any bet from the first set to that from the second set. In the variation of the Ellsberg experiment, it is quite possible that the individual prefers to bet on white balls from the second urn, but still prefers a bet on the first urn to any bet on the second urn.[2] The intuition is that even after 100 draws, the subject still cannot rule out the possibility that there is only one red ball in the urn, or any other possibilities for that matter. On the other hand, the 45/55 ratio does tilt the odds more in favor of drawing a white than a red. Such choice behavior, as well as that in the

[1] See Gilboa (1987), Schmeidler (1989), Gilboa and Schmeidler (1989), Nakamura (1990), Sarin and Wakker (1992), Chew and Karni (1994), Casadesus-Masanell, Klibanoff and Ozdenoren (2000), and Klibanoff, Marinacci and Mukerji (2002). For a comprehensive survey of this literature, the reader is referred to Camerer and Weber (1992) and Sarin and Wakker (1998). In multi-period settings, see Klibanoff (1995), Wang (2002) and Epstein and Schneider (2002).

[2] To the best of my knowledge, there is no formal experiment demonstrating such results. The statement here is based on informal discussion with my colleagues.

original Ellsberg, is inconsistent with Savage's expected utility preferences, and more generally probabilistically sophisticated preferences (Machina and Schmeidler (1992)).

It is easily seen that the difference between the two urns lies in the fact that the probability of drawing a red is known for urn one but not for urn two. Formally, the difference cab be described by a set of probability measures as follows. Let $\Omega_1 = \{R_1, W_1\}$ and $\Omega_2 = \{R_2, W_2\}$ denote the sets of outcomes of draws from urns one and two, respectively, and let $\Omega = \Omega_1 \times \Omega_2$ denote the set of all possible combinations of the outcomes. The set of probability measures describing the setting of Ellsberg experiment is then given by

$$\mathcal{P} = \{Q = m_1 \times m_2 : m_1 = (0.5, 0.5),\ m_2 = (p, 1-p), \text{where } p \in (0, 1)\}.$$

Here $m_1 = (0.5, 0.5)$ is the probability measure on Ω_1 that gives the probabilities of drawing a red ball or a white ball from the first urn, whereas $m_2 = (p, 1-p)$ is a probability measure on Ω_2. Since the number of red balls in the second urn is unknown, the true probability p of drawing a red ball can be anywhere between 0 to 1. Note that the subset $\{R_1\} \times \Omega_2$ can be interpreted as the event that a ball randomly drawn from the first urn is red. Thus the first bet in the Ellsberg experiment can be described by $c_1 = 1_{\{R_1\}\times\Omega_2}$. That is, $c_1 = 1$ if a red ball is drawn from the first urn and zero otherwise. Similarly, $c_2 = 1_{\{W_1\}\times\Omega_2}$, $c_3 = 1_{\Omega_1\times\{R_2\}}$ and $c_4 = 1_{\Omega_1\times\{W_2\}}$ describe the other three bets in the experiment. Clearly, for any $Q \in \mathcal{P}$, $Q(\{c_1 = 1\})$ is 0.5, but $Q(\{c_3 = 1\}) = p$, which can be any number between 0 and 1.

One may argue that in the variation of the Ellsberg experiment, due to the extra piece of information, the subject has learned about the true probability of drawing a red from the second urn. There is no question that the 45/55 ratio is informative about the unknown probability and, as the sample size gets larger, it can become even more informative. What is interesting in this experiment is that knowing that 45 red balls were drawn is still not enough to rule out the possibility that there is only one red ball in the urn. In other words, learning has not shrunk the set of probability measures $\mathcal{P}$. At the same time, however, the information about earlier draws does affect individuals' choice behavior.

While the experiment is in an idealized setting, it has natural counterpart in the real world. Consider the follow scenario. An econometrician tries to estimate a model of the economy for a decision maker. After a whole slew of econometric analysis, typically including specification analysis and parameter estimation, the econometrician comes up with a model described by a probability measure P. However, the econometrician is not completely confident that this is the true model, due either to not having enough data in the specification analysis and the parameter estimation, or to approximation for parsimony. At the end of the analysis, the decision maker is presented with a probability measure P, called the reference probability measure, and a set of probability measures $\mathcal{P}$ that summarizes the precision of the econometric analysis, e.g., a confidence region. Intuitively, both P and $\mathcal{P}$ should be relevant for the decision maker in his decision-making process.

The discussion above suggests that while the Ellsberg experiment highlights the distinction between risk and uncertainty advocated by Knight (1921), its setting does not fully capture what a decision maker faces in the real world. A more satisfactory description of the uncertainty faced by a decision maker in the real world requires that the modelling of uncertainty goes beyond using only a set of probability measures and allows for the role of a reference probability.

This paper is also motivated by the development in the application front. Multi-prior expected utility, axiomatized by Schmeidler (1989) and Gilboa and Schmeidler (1989) for example, has been applied in Dow and Werlang (1992) to study a portfolio choice problem and in Epstein and Wang (1994, 1995), Chen and Epstein (2001), and Epstein and Miao (2001) to develop intertemporal asset pricing models. Recently, Anderson, Hansen and Sargent (1999), in their study of the importance of preference for robustness in macro and asset pricing problems, introduced a preference that has not been axiomatized nor studied in the decision theory literature. This paper can be viewed as a step in the study of the decision-theoretic foundation for that preference.

The rest of this paper is organized as follows. Section 2 lays out the framework in which we axiomatize the preferences. Section 3 discusses a general structure of uncertainty averse preferences. Section 4 provides the axiomatization of the class of multi-prior preferences. Finally, the paper ends with a discussion of potential applications. The appendix contains the proofs.

2 Modelling Uncertainty

To formalize what is described in the last section, let Ω denote the state space. Ω is assumed to have n elements. Let X denote the set of prizes. For example, $X = [0, M] \subset R_+$. The consumption space is $\mathcal{C}$ which is the set of all functions from Ω to X. Let Δ_n denote the simplex in R^n,

$$\Delta_n = \left\{ p : p \in R^n_+, \sum_{i=1}^{n} p_i = 1 \right\},$$

and let $\mathbf{L}$ denote the family of all closed subsets $\mathcal{P}$ of Δ_n. We define an uncertain prospect as a triple $(c, \mathcal{P}, P)$ with $c \in \mathcal{C}$, $P \in \Delta_n$, and $P \in \mathcal{P} \in \mathbf{L}$. The interpretation is that the random variable c describes the state-contingent consumption, P is the reference probability measure. The set $\mathcal{P}$ and the reference probability measure P together provide the information on the likelihood of the states. The preference of the decision maker is an ordering over triples $(c, \mathcal{P}, P)$. Obviously, for each fixed $(\mathcal{P}, P)$, $\succeq$ reduces to a preference over consumption in the more standard Savage setting.

It should be emphasized that while simple, our modelling of an uncertain prospect deviates significantly from the Savage approach commonly adopted in the literature. The deviation is motivated by the objective of our axiomatization. In Savage's expected utility theory, uncertainty is assumed to be a subjective matter. In deriving his subjective expected utility theory, Savage's objective is to derive simultaneously both a probability measure P as the decision maker's belief and a utility function u such that the decision maker's preference is represented by

$$c \succeq c' \quad \text{if and only if} \quad \int u(c)dP \geq \int u(c')dP.$$

For Savage, the derivation of P as the subjective belief of the decision maker is as important a part of his derivation as the expected utility functional form. As such, it requires that random consumption c be the only input of the preference. Our objective, however, is not to extract the belief component of the decision maker's preference, but to give a characterization of how the information $(\mathcal{P}, P)$ is utilized in the preference ordering. For instance, a pertinent question for us is: is it true that the information is utilized according to a maxmin principle, i.e.,

$$\succeq \quad \text{is represented by} \quad \min_{p \in \mathcal{P}} \sum_i p_i u(c_i), \tag{1}$$

as in Gilboa and Schmeidler (1989) and Casadesus-Masanell, Klibanoff and Ozdenoren (2000), or by the minimum entropy principle as in Anderson, Hansen and Sargent (1999), i.e.,

$$\succeq \quad \text{is represented by} \quad \inf_{p \in \mathcal{P}} \left\{ \phi \sum_i p_i \ln\left(\frac{p_i}{P_i}\right) + \sum_i p_i u(c_i) \right\}, \tag{2}$$

where ϕ is a positive real number? Note that $\sum_i P_i \left(\frac{p_i}{P_i}\right) \ln\left(\frac{p_i}{P_i}\right)$ is the relative entropy of p with respect to P. The objective of our paper is to find the set of behavioral rules under which the decision maker behaves according to the principle in (2).

Our approach is in fact closer to that of von-Neumann-Morgenstern's objective expected utility theory. In that theory, the risk of a state contingent consumption is taken as objectively given and is described by a probability measure p. We treat $(c, \mathcal{P}, P)$ as an objective description of an uncertainty prospect. We view $(\mathcal{P}, P)$ as being provided by the econometrician based on analysis of the data. For instance, the data can provide a "point estimate", P, and a "confidence region", $\mathcal{P}$, of the underlying probability law. As long as there are data that are not perfectly informative, we are in the situation of objective uncertainty. The preferences we will axiomatize take that pair $(\mathcal{P}, P)$ as given.

3 A Minimization Principle

Having defined uncertain prospects $(c, \mathcal{P}, P)$, we are ready for the axiomatization of preferences. We begin with a general characterization of uncertainty averse preferences. We maintain throughout the assumption that $\succeq$ can be represented by some utility function $U(c, \mathcal{P}, P)$. The existence of such representation is readily obtained by applying Debreu (1954). For notational simplicity, if P is the reference probability measure and $\mathcal{P} = \{p, P\}$, we will write $U(c, \mathcal{P}, P)$ as $U(c, p, P)$.

Axiom 1 *(Intermediateness) For each $\mathcal{P} \in \mathbf{L}$, $P \in \Delta_n$ and $c \in \mathcal{C}$, there exist p and $q \in \mathcal{P}$ such that $(c, p, P) \succeq (c, \mathcal{P}, P) \succeq (c, q, P)$.*

Axiom 2 *(Uncertainty Aversion) For each $c \in \mathcal{C}$ and $P \in \Delta_n$, if $\mathcal{P}' \subset \mathcal{P}$, then $(c, \mathcal{P}', P) \succeq (c, \mathcal{P}, P)$.*

Axiom 3 *(Continuity) $\succeq$ is continuous in $(c, \mathcal{P}, P)$, i.e., for each $(c', \mathcal{P}', P)$, the set $\{(c, \mathcal{P}, P) \in \mathcal{C} \times \mathbf{L} \times \Delta_n : (c, \mathcal{P}, P) \succeq (c', \mathcal{P}', P')\}$ is closed.*[3]

The interpretation of these three axioms are straightforward. The probability measures p and q can be viewed as the optimistic and pessimistic prospects respectively. Axiom 1 says that in each $\mathcal{P}$, there is always an optimistic as well as a pessimistic prospect and, naturally, the ranking of $(c, \mathcal{P}, P)$ should be between these two extreme prospects. When $\mathcal{P}$ is a larger set than $\mathcal{P}'$, the

[3]Here $\mathcal{C}$ is endowed with pointwise convergence topology, $\mathbf{L}$ is endowed with the Hausdorff topology, and Δ_n is endowed with the weak convergence topology. Finally, $\mathcal{C} \times \mathbf{L} \times \Delta_n$ is endowed with the product topology.

prospect $(c, \mathcal{P}, P)$ is more uncertain than $(c, \mathcal{P}', P)$. Axiom 2 says that if $(c, \mathcal{P}, P)$ is more uncertain than $(c, \mathcal{P}', P)$, then it is less preferred. Axiom 3 is standard.

Theorem 3.1 *(Minimization Principle) The preference $\succeq$ satisfies Axioms 1, 2 and 3 if and only if $U(c, \mathcal{P}, P)$ is continuous and*

$$U(c, \mathcal{P}, P) = \min_{q \in \mathcal{P}} U(c, q, P). \tag{3}$$

All maxmin expected utility preferences, such as Gilboa and Schmeidler (1989) and Casadesus-Masanell, Klibanoff and Ozdenoren (2000), satisfy this minimization principle, although they are derived in different settings.

4 Multi-Prior Preference with A Reference Prior

In this section, we characterize a class of uncertainty averse multi-prior preferences that behave according to the minimum entropy rule in (2).

First we introduce a notation. For any set of indices $J = \{j_1, \ldots, j_k\}$, $1 \le j_1 < \ldots < j_k \le n$, and any vector $a \in R^n$, denote by a_J the vector in R^k defined by $a_J = (a_{j_1}, \ldots, a_{j_k})$.

Axiom 4 *(Separability) For any $J = \{j_1, \ldots, j_k\}$, c_J, c'_J in R^k_+, c_{J^c}, c'_{J^c} in R^{n-k}_+, p_J, p'_J in R^k_+, and p_{J^c}, p'_{J^c} in R^{n-k}_+ such that (p_J, p_{J^c}), (p'_J, p_{J^c}), (p_J, p'_{J^c}) and (p'_J, p'_{J^c}) are all in Δ_n, if*

$$((c_J, c_{J^c}), (p_J, p_{J^c}), P) \succeq ((c'_J, c_{J^c}), (p'_J, p_{J^c}), P)$$

then

$$((c_J, c'_{J^c}), (p_J, p'_{J^c}), P) \succeq ((c'_J, c'_{J^c}), (p'_J, p'_{J^c}), P).$$

Furthermore, if $p_J 1_J = 0$ and $P_J 1_J = 0$, then $((c_J, c_{J^c}), (p_J, p_{J^c}), P) \sim ((c'_J, c_{J^c}), (p_J, p_{J^c}), P)$.

This axiom says that if restricted to only prospects of the form (c, p, P), the preference ordering is separable in the sense that if (c, p, P) and (c', p', P) have identical payoffs and probabilities in event A, then the ranking of the two prospects depend only on their payoffs and probabilities in the event A^c. Moreover, if A is an impossible event, then what c assumes on the event is irrelevant.

Axiom 5 *(Smoothness) Let $\mathcal{M}$ be the collection of all sets of the form*

$$\left\{p \in \Delta_n : a^\top p = b\right\} \tag{4}$$

where a is a n-vector and b is a real number. (i) For each c and P, $U(c,p,P)$ is continuously differentiable in p. (ii) For all $c \in \mathcal{C}$, $P \in \Delta_n$, $\mathcal{P}$ and $\mathcal{P}'$ both in $\mathcal{M}$, if $p \in \mathcal{P}$ and $p' \in \mathcal{P}'$ are such that $p \neq P$, $p' \neq P$, $(c, \mathcal{P} \cup \{P\}, P) \succeq (c,p,P)$ and $(c, \mathcal{P}' \cup \{P\}, P) \succeq (c, p', P)$, then $p \neq p'$ if and only if $\mathcal{P} \neq \mathcal{P}'$ unless both $\mathcal{P}$ and $\mathcal{P}'$ contain p_0 where p_0 is such that $(c, \Delta_n, P) \sim (c, p_0, P)$.

The meaning of differentiability is clear. To understand (ii), note that each $\mathcal{P}$ in $\mathcal{M}$ is a hyperplane in Δ_n. By Theorem 3, if $U(c, \mathcal{P} \cup \{P\}, P) \succeq U(c,p,P)$, then p is the tangency point of $U(c,\cdot,P)$ to the hyperplane. Part (ii) of Axiom 5 says that, as one rotates the hyperplane, the tangency point p changes. Furthermore, as one moves from one p to another, the tangent hyperplane also changes. This axiom rules out preferences with kinky indifference curves, such as multi-prior expected utility preferences, and those with flat indifference curves.

Axiom 6 *(Consistency) There exists a partition $J_1, \ldots, J_m$ of $\{1, \ldots, n\}$ such that*

(a) if $i \neq j$ are both in J_k for some $k = 1, \ldots, m$, and $P_i + P_j = \alpha$, $0 < \alpha \leq 1$, then for all $c \in \mathcal{C}$ such that $c_i = c_j$ and $0 < t \leq 1$, there exists a $p \in \mathcal{P}_{ij}(t) = \{p \in \Delta_n : p_i + p_j = t\}$ with $p_i/p_j = P_i/P_j$ such that if $(c, \mathcal{P}_{ij}(t) \cup \{P\}, P) \not\preceq (c,P,P)$, then $(c, \mathcal{P}_{ij}(t) \cup \{P\}, P) \succeq (c,p,P)$. Moreover, if $p \in \mathcal{P}_{ij}(t)$ and $(c, \mathcal{P}_{ij}(t) \cup \{P\}, P) \succeq (c,p,P)$ then $p_i/p_j = P_i/P_j$;

(b) if $\mathcal{P}$ is convex and is such that for some $t_k \geq 0$, $k = 1, \ldots, m$ with $\sum_{k=1}^m t_k = 1$, all p in $\mathcal{P}$ satisfy $\sum_{i \in J_k} p_i = t_k$, $k = 1, \ldots, m$, then for all c and c' with the property that $c_i = c_j$, $c'_i = c'_j$ whenever i, j are both in J_k for some k, $(c, \mathcal{P} \cup \{P\}, P) \sim (c,p,P)$ where $p \in \mathcal{P} \cup \{P\}$ if and only if $(c', \mathcal{P} \cup \{P\}, P) \sim (c', p, P)$, provided that $p \neq P$;

(c) if $\mathcal{P}'$ and $\mathcal{P}$ are as in (b) and $\mathcal{P}' \subset \mathcal{P}$, then for any c of the form $((c_1, i \in J_1), \ldots, (c_m, i \in J_m))$, and $P \in \Delta_n$, if $(c, \mathcal{P}, P) \sim (c,q,P)$ and $(c, \mathcal{P}', q) \sim (c,p,q)$ with $q \in \mathcal{P}$, $p \in \mathcal{P}'$, $q \neq P$ and $p \neq q$, then $(c, \mathcal{P}', P) \sim (c,p,P)$.

While the formal statement of this axiom is a bit involved, the intuition behind it is rather straightforward. Consider (b) first. For simplicity assume that $P \in \mathcal{P}$. When c is of the form

$((c_1, i \in J_1), \ldots, (c_m, i \in J_m))$, its distribution is the same under any probability measure in $\mathcal{P}$. In fact, in this case, the probability of event J_k is equal to t_k for all $p \in \mathcal{P}$. Then, the pessimistic prospect p referred to in the discussion following Axioms 1-3 in Section 3 should be independent of c. This is what part (b) of the axiom says. An immediate consequence of this axiom is that when the probabilities of the partition, $J_1, \ldots, J_m$, of the state space Ω are known, the ranking of contingent consumption measurable to that partition is based entirely on the distribution of the consumptions.

How is the pessimistic prospect of $(c, \mathcal{P}, P)$ determined, in general? Normally, there should be two forces at work in its determination. One force tries to pick p closest to the reference probability measure. Recall that $(\mathcal{P}, P)$ is the best information the econometrician can provide regarding the true model with P being the best point estimate. So other things being equal, there is no reason for the decision maker not to pick p close to P. The other force is the uncertainty aversion of the decision maker which tends to pick the most pessimistic prospect. When c is of the form $((c_1, i \in J_1), \ldots, (c_m, i \in J_m))$, its distribution is the same under any probability measure in $\mathcal{P}$. Thus the second force becomes irrelevant and only the first force is at work. In this case how the decision maker picks the p reveals which p in $\mathcal{P}$ he considers to be the closest to the reference model P.[4] Part (c) of Axiom 6 says that to be the closest to P, the pessimistic q should be a good summary of the information in $(\mathcal{P}, P)$ so that if $\mathcal{P}$ is further narrowed down to $\mathcal{P}'$, then taking q as the reference probability measure leads consistently to the same $p \in \mathcal{P}'$ as the closest to the true model as taking P as the reference probability measure.

Part (a) of the axiom shares similar intuition as part (c), but applies to a more specific setting. To see the intuition, consider the following experiment. There is one urn that contains 90 balls of either red, white or black color. It is known that there are 30 black balls in the urn. There is no further information about the proportion of the red balls in the urn except that the urn has been sampled T times, in which red ball was drawn T_1 times and white ball T_2 times. With this information, it is natural to take $P(red) = \frac{2}{3}\frac{T_1}{T_1+T_2}$, $P(white) = \frac{2}{3}\frac{T_2}{T_1+T_2}$ and $P(black) = 1/3$ as the reference probability measure. Now suppose that the number of black balls is increased to 60. The question is which probability measure should be considered as closest to the reference measure P. In this case, it seems intuitive to require the probability measure p be such that $p(red) + p(white) = 1/2$ and $p(red)/p(white) = P_1/P_2$, which is what is required in the axiom.

[4]This is where we really need P not to be in $\mathcal{P}$, otherwise the closest p would be P itself.

Axiom 7 *(P-Independence) (i) For all* (c, P, P), (c', P', P') *in* $\mathcal{C} \times \Delta_n \times \Delta_n$, $F_{c,P} = F_{c',P'}$ *implies* $(c, P, P) \sim (c', P', P')$. *(ii) For any* $\alpha \in (0, 1)$, *let* c_1, c_2, c_3 *and* $c_4 \in \mathcal{C}$, $p_1, \ldots, p_4$ *in* Δ_n, *and* p, p', q *in* $\Delta(X)$ *be such that* $F_{c_1,p_1} = p$ *and* $F_{c_2,p_2} = p'$, $F_{c_3,p_3} = \alpha p + (1-\alpha)q$ *and* $F_{c_4,p_4} = \alpha p' + (1-\alpha)q$. *If* $(c_1, p_1, p_1) \succeq (c_2, p_2, p_2)$, *then* $(c_3, p_3, p_3) \succeq (c_4, p_4, p_4)$.

This axiom requires that the ranking of any two uncertain prospects (c, p, P) and (c', p', P') depend only on the distributions $F_{c,p}$ and $F_{c',p'}$, and the independence axiom of von Neumann-Morgenstern holds, regardless the prior. Axiom 7 requires so only when $\mathcal{P} = \{P\}$. In essence, it says that when the econometrician and the decision maker are perfectly confident about the prior, the preference conforms to expected utility.

Theorem 4.1 *Let* $J_1, \ldots, J_m$ *be the partition as described in Axiom 6. Suppose that* $m \geq 2$ *and each* J_k *contains more than three elements.*

(a) The preference $\succeq$ *satisfies Axioms 1-5, 6(a,b) and 7 if and only if there exist a continuous function* u, *unique upto a positive affine transform, and* m *continuously differentiable strictly convex functions* v_k *with* $v_k(0) = v_k(1) = 0$ *such that*

$$U(c, \mathcal{P}, P) = \min_{p \in \mathcal{P}} \left\{ \sum_{k=1}^{m} \sum_{i \in J_k} P_i v_k(p_i/P_i) + \sum_{i=1}^{n} p_i u(c) \right\} \tag{5}$$

(b) The preference $\succeq$ *satisfies Axioms 1-7 if and only if there exist a continuous function* u, *unique upto a positive affine transform, and* m *positive constants* ϕ_k *such that*

$$U(c, \mathcal{P}, P) = \min_{p \in \mathcal{P}} \left\{ \sum_{k=1}^{m} \phi_k \left(\sum_{i \in J_k} p_i \ln(p_i/P_i) \right) + \sum_{i=1}^{n} p_i u(c_i) \right\}. \tag{6}$$

5 Discussion

We conclude this paper with some discussion on the potential applications of the preferences characterized in Theorem 4.1 and several useful special cases.

5.1 Resolving Ellsberg Type Paradox

The most natural application of the preference in (6) is obviously in the experiment described in Section 1. Recall that $\Omega = \Omega_1 \times \Omega_2$, $\Omega_1 = \{R_1, W_1\}$, $\Omega_2 = \{R_2, W_2\}$, and

$$\mathcal{P} = \{m_1 \times m_2 : m_1 = (0.5, 0.5),\ m_2 = (p, 1-p), \text{where } p \in [0,1]\}.$$

Given that 45 red balls are drawn in the sample of 100 draws, it is natural to take $P = m_1 \times \hat{m}_2$ where $\hat{m}_2 = (0.45, 0.55)$ as the reference probability measure. Let $c_1 = 1_{\{R_1\}\times\Omega_2}$, $c_2 = 1_{\{W_1\}\times\Omega_2}$, $c_3 = 1_{\Omega_1\times\{R_2\}}$ and $c_4 = 1_{\Omega_1\times\{W_2\}}$ denote the four bets as before. Now, let the preference be given by (6) with $\phi_k = \phi = 1 = u(1)$,

$$U(c, \mathcal{P}, P) = \min_{p\in\mathcal{P}} \left\{ \phi \sum_{i=1}^{n} p_i \ln(p_i/P_i) + \sum_{i=1}^{n} p_i u(c_i) \right\}.$$

It is readily verified that for this utility function,

$$U(c_1, \mathcal{P}, P) = 0.5 = U(c_2, \mathcal{P}, P), \quad \text{and} \quad U(c_3, \mathcal{P}, P) = 0.3347 < U(c_4, \mathcal{P}, P) = 0.4272.$$

So the individual is still indifferent between the two bets on the first urn, but is no longer indifferent between the two bets on the second urn. Furthermore, the individual still prefers bets on the first urn (c_1 and c_2) over bets on the second urn (c_3 and c_4).

While the utility function in (6) finds natural applications in situations where there is some additional information that leads to a reference probability measure, it is less useful in setting like the original Ellsberg experiment where the information that the decision maker has is best described by just a set of probability measures as the $\mathcal{P}$ above. Notice, however, when $\phi_k \to 0$ for all k, the utility function in (6) approaches

$$U(c, \mathcal{P}, P) = \min_{p\in\mathcal{P}} \left\{ \sum_{i=1}^{n} p_i u(c_i) \right\}.$$

The reference probability measure does not play any role in this utility function. Thus multi-prior expected utility can be viewed as the limit of (6), which is readily applicable to situations like the Ellsberg experiment. For instance, it is readily verified that for this utility function,

$$U(c_1, \mathcal{P}, P) = 0.5u(1) = U(c_2, \mathcal{P}, P);$$

and

$$U(c_3, \mathcal{P}, P) = U(c_4, \mathcal{P}, P) = 0 < 0.5u(1).$$

Therefore, such preference will exhibit behavior consistent with that observed in the Ellsberg experiment.

5.2 Special Cases

There are several special cases of (6) which are particularly interesting from the point of view of applications. The first one is where $\phi_k = \phi$. In this case, (6) becomes

$$U(c, \mathcal{P}, P) = \min_{p \in \mathcal{P}} \left\{ \phi \sum_i p_i \ln(p_i/P_i) + \sum_{i=1}^{n} p_i u(c_i) \right\}.$$

When $\mathcal{P} = \Delta_n$, it is the static counterpart of the utility function used in Anderson, Hansen and Sargent (1999) and Maenhout (1999).

The second case is one where uncertainty arises from several sources. Let $S = \{s_1, \ldots, s_N\}$ be a state space on which M random variables, Y_m, $m = 1, \ldots, M$, are defined. The distributions of these Y_m random variables are not known to the decision maker. This uncertainty is described by $(\mathcal{P}, P)$. Introduce the auxiliary set $\Omega = \Omega_1 \times S$ where $\Omega_1 = \{1, \ldots, M\}$. The random variables Y_m have natural extensions to Ω. We extend $(\mathcal{P}, P)$ to Ω by

$$\hat{P} = \hat{m} \times P, \qquad \hat{\mathcal{P}} = \{\hat{m} \times p : p \in \mathcal{P}\}$$

where $\hat{m} = (1/M, \ldots, 1/M)$. Let $J_m = \{m\} \times S$, $m = 1, \ldots, M$. Then $\{J_m : m = 1, \ldots, M\}$ is a partition of Ω. Clearly $\hat{P}(J_m) = 1/M$ for all m. For this $(\hat{\mathcal{P}}, \hat{P})$, (6) simplifies to

$$U(c, \hat{\mathcal{P}}, \hat{P}) = \frac{1}{M} \min_{p \in \mathcal{P}} \left\{ \sum_{m=1}^{M} \phi_m \left(\sum_{i=1}^{N} p(y_m(s_i)) \ln \frac{p(y_m(s_i))}{P(y_m(s_i))} \right) + \sum_{i=1}^{N} p(s_i) u(c_i(s_i)) \right\}. \quad (7)$$

Note that

$$\sum_{i=1}^{N} p(y_m(s_i)) \ln \frac{p(y_m(s_i))}{P(y_m(s_i))}$$

is the relative entropy of Y_m. So the first term in (7) is the weighted sum of the relative entropies of Y_m. The weights are determined by ϕ_k which are preference parameters. Fox and Tversky (1995) find that people may have different reaction towards different sources of uncertainty depending on whether they believe they have better knowledge about a particular source of uncertainty. The utility function in (7) can accommodate that finding when ϕ_k are different. As another potential application, the utility function can be viewed as the static counterpart of the utility function used in Uppal and Wang (2002).

The final special case is where the decision maker has perfect knowledge about one source of uncertainty but not the other. Using the notation above, let $M = 2$ so that there are only two sources of uncertainty represented by two random variables Y_1 and Y_2. If the probability law of Y_1 is perfectly known, then equation (6) reduces to

$$U(c, \hat{\mathcal{P}}, \hat{P}) = 0.5 \min_{p \in \mathcal{P}} \left\{ \phi_2 \sum_{s \in S} p(y_2(s)) \ln \frac{p(y_2(s))}{P(y_2(s))} + \sum_{i=1}^{n} p(s) u(c(s)) \right\}.$$

The dynamic version of this utility function is used in Liu, Pan and Wang (2002) to study the effect of uncertainty about rare events on equilibrium equity premium.

A Appendix A

Proof of Theorem 3.1: Observe that, by Axiom 2, $U(c, \mathcal{P}, P) \leq U(c, p, P)$ for any c, $P \in \Delta_n$, $\mathcal{P} \in \mathbf{L}$ and $p \in \mathcal{P}$. The necessity claim of the theorem follows from this observation and Axioms 1 and 3. For sufficiency, the continuity of U and the compactness of $\mathcal{P}$ imply the existence of p and $q \in \mathcal{P}$ that solves the minimization problem in (3) and the symmetric maximization problem. For this pair, p and q, $U(c, q, P) \geq U(c, \mathcal{P}, P) \geq U(c, p, P)$. Thus Axiom 1 is satisfied. Continuity and Uncertainty Aversion are obvious. ∎

Proof of Theorem 4.1: The proof begins with two propositions which are useful in illustrating the role of the axioms.

Proposition A.1 *The preference $U(c, \mathcal{P}, P)$ satisfies Axioms 1-4 if and only if there exist continuous functions $f_i : X \times [0,1] \times \Delta_n \to R$ and $A_i : \Delta_n \to R$ such that*

$$U(c, \mathcal{P}, P) = \sum_i A_i(P) + \min_{p \in \mathcal{P}} \sum_{i=1}^{n} f_i(c_i, p_i, P), \tag{8}$$

and $f_i(c_i, 0, P) = 0$.

Proof: In light of Theorem 3.1, for necessity, it suffices to show that

$$U(c, p, P) = \sum_i A_i(P) + \sum_{i=1}^{n} f_i(c_i, p_i, P) \tag{9}$$

represents the preference over prospects of the form (c, p, P). But this follows from Axiom 4 and Theorem 3 of Debreu (1959), see also Koopmans (1972). Note the functions $A_i(P)$ and the restriction that $f_i(c_i, 0, P) = 0$ come from normalization. Again, for sufficiency, we only need to show that if $U(c, p, P)$ is given by (9), then Axiom 4 is satisfied. But this follows from

$$\sum_{i \in J} f_i(c_i, p_i, P) + \sum_{i \in J^c} f_i(c_i, p_i, P) \geq \sum_{i \in J} f_i(c'_i, p'_i, P) + \sum_{i \in J^c} f_i(c_i, p_i, P)$$

if and only if

$$\sum_{i \in J} f_i(c_i, p_i, P) + \sum_{i \in J^c} f_i(c'_i, p'_i, P) \geq \sum_{i \in J} f_i(c'_i, p'_i, P) + \sum_{i \in J^c} f_i(c'_i, p'_i, P). \quad \blacksquare$$

The following lemma is adapted from Lemma 1 of Csiszar (1991).

Lemma A.2 *Suppose that the preference $\succeq$ satisfies Axioms 1-3 and 5. Let $c \in \mathcal{C}$ and $P \in \Delta_n$. Then for every $\mathcal{P}' \in \mathbf{L}$ of dimension less than $n-2$, meaning that the linear span of $\mathcal{P}'$ in R^n is of dimension less than $n-2$, there exists a unique pair $(p, \mathcal{P})$ with $\mathcal{P} \in \mathcal{M}$ such that $p \in \mathcal{P}' \subset \mathcal{P}$ and $(c, \mathcal{P}' \cup \{P\}, P) \sim (c, p, P) \sim (c, \mathcal{P} \cup \{P\}, P)$, provided that $p \neq P$. As a result, if $p \in \mathcal{P}'$, $p \in \mathcal{P}$, $p \neq P$, $(c, \mathcal{P}' \cup \{P\}, P) \sim (c, p, P) \sim (c, \mathcal{P} \cup \{P\}, P)$ and $\mathcal{P} \in \mathcal{M}$, then $\mathcal{P}' \subset \mathcal{P}$ unless $p = p_0$ where p_0 is such that $(c, \Delta_n, P) \sim (c, p^0, P)$.*

Proof: Fix $c \in \mathcal{C}$. Let $\mathcal{P}' \in \mathbf{L}$. We prove that if $\dim(\mathcal{P}') = d$ with $0 \leq d < n-2$, then there exists a pair $(p, \mathcal{P})$ with $\mathcal{P} = \{p \in \Delta_n : Ap = b\} \in \mathbf{L}$, $p \in \mathcal{P}'$ and $p \in \mathcal{P}$ such that $\mathcal{P} \supset \mathcal{P}'$, $\dim(\mathcal{P}) = d+1$, and $U(c, \mathcal{P} \cup \{P\}, P) = U(c, p, P) = U(c, \mathcal{P}' \cup \{P\}, P)$, provided that $p \neq P$. Further, instead of the last equality, it suffices to show that $p \in \mathcal{P}'$ because of Axiom 2.

Now pick any $(d+1)$-dimensional $\mathcal{P}_1 = \{p \in \Delta_n : Ap = b\} \supset \mathcal{P}'$. Let $p_1 \in \mathcal{P}_1$ be such that $U(c, \mathcal{P}_1 \cup \{P\}, P) = U(c, p_1, P)$. Such p_1 always exists since $\mathcal{P}_1$ is closed and compact and U is continuous. If $p_1 \in \mathcal{P}'$, then there is nothing to prove. So suppose that $p_1 \notin \mathcal{P}'$. We will "rotate" $\mathcal{P}_1$ to obtain a family $\{\mathcal{P}_t : 0 \leq t \leq 2\}$ and show that for some t, there exists a $p_t \in \mathcal{P}_t$ such that $U(c, \mathcal{P}_t \cup \{P\}, P) = U(c, p_t, P)$ and $p_t \in \mathcal{P}'$. To this end, pick any p_0, $p_2 \notin \mathcal{P}_1$ such that some interior point of the segment $[p_0, p_2]$ is in $\mathcal{P}'$. Set $p_t = (1-t)p_0 + tp_1$ if $0 \leq t \leq 1$ and $p_t = (2-t)p_1 + (t-1)p_2$ if $1 \leq t \leq 2$. Finally, let $\mathcal{P}_t$ denote the subset of Δ_n given by $\{q \in \Delta_n : q = \alpha p + \beta p_t,\ p \in \mathcal{P}',\ \alpha, \beta \in R\}$. Then $\mathcal{P}' \subset \mathcal{P}_t$ for all t, $\mathcal{P}_0 = \mathcal{P}_2$ (because $[p_0, p_2]$ goes through $\mathcal{P}'$) and $\mathcal{P}_{t_1} \cap \mathcal{P}_{t_2} = \mathcal{P}'$ if $0 \leq t_1 < t_2 < 2$.

Now let q_t be such that $U(c, \mathcal{P}_t \cup \{P\}, P) = U(c, q_t, P)$. If $q_t = P$, then there is nothing to prove. So we assume that $q_t \neq P$. If $d = n-3$ so that $\mathcal{P}_t$ are $(n-2)$ dimensional and hence in $\mathcal{M}$, then by Axiom 5, this q_t is unique. By the continuity axiom, $\{q_t : 0 \leq t \leq 2\}$ is a continuous closed curve in the subspace $\tilde{\mathcal{P}}$ spanned by $\mathcal{P}_1$ and p_0 (because $\mathcal{P}_0 = \mathcal{P}_2$). For $\epsilon > 0$ sufficiently small, $q_{1-\epsilon}$ and $q_{1+\epsilon}$—which are arbitrarily close to p_1—are separated by $\mathcal{P}_1$ within $\tilde{\mathcal{P}}$. Since $\{q_t : 0 \leq t \leq 2\}$ is a closed curve, there exists some t with $|t-1| > \epsilon$ for which $q_t \in \mathcal{P}_1$. Then $q_t \in \mathcal{P}_t \cap \mathcal{P}_1 = \mathcal{P}'$. Hence $q_t \in \mathcal{P}' \subset \mathcal{P}_t$.

For uniqueness, suppose that $(p, \mathcal{P})$ and $(\hat{p}, \hat{\mathcal{P}})$ both satisfy the claim of the lemma. Then we must have

$$(c, p, P) \sim (c, \mathcal{P}, P) \sim (c, \hat{\mathcal{P}}, P) \sim (c, \hat{p}, P),$$

and $\hat{p} \in \mathcal{P}$ and $p \in \hat{\mathcal{P}}$. If $p \neq \hat{p}$, it is a contradiction to Axiom 5. If $p = \hat{p}$, then $\mathcal{P} \neq \hat{\mathcal{P}}$. It is also a contradiction to Axiom 5.

So the first part of the lemma is true for $\mathcal{P}'$ of dimension $n-3$. Now we induct on d. Suppose the claim is true for $\mathcal{P}'$ of dimension $d+1 < n-2$. Let $\mathcal{P}'$ be of dimension d. Repeat the argument above upto the existence of q_t such that $U(c, \mathcal{P}_t \cup \{P\}, P) = U(c, q_t, P)$. Now since $\mathcal{P}_t$ is $(d+1)$-dimensional, by induction assumption, there exists a unique pair $(\hat{q}_t, \hat{\mathcal{P}}_t)$ with $\hat{\mathcal{P}}_t \in \mathcal{M}$ such that $U(c, \mathcal{P}_t \cup \{P\}, P) = U(c, \hat{q}_t, P) = U(c, \hat{\mathcal{P}}_t \cup \{P\}, P)$, provided that $\hat{q}_t \neq P$. So we in fact can set $q_t = \hat{q}_t$. Since such $\hat{q}_t$ is unique by the induction assumption, the rest of the proof is the same as above. This completes the proof of the first part of the lemma.

For the second part of the lemma, let $(\hat{p}, \hat{\mathcal{P}})$ be the unique pair whose existence is guaranteed by the first part of the lemma. For this pair $\hat{p} \in \mathcal{P}' \subset \hat{\mathcal{P}}$ and $(c, \mathcal{P}' \cup \{P\}, P) \sim (c, \hat{p}, P) \sim (c, \hat{\mathcal{P}} \cup \{P\}, P)$. On the other hand, by the assumption of the lemma, $p \in \mathcal{P}'$, $p \in \mathcal{P}$, $(c, \mathcal{P}' \cup \{P\}, P) \sim (c, p, P) \sim (c, \mathcal{P} \cup \{P\}, P)$ and $\mathcal{P} \in \mathcal{M}$. By exactly the same argument as in the uniqueness proof above, this is impossible unless $\mathcal{P} = \hat{\mathcal{P}}$. The proof of the lemma is complete. ∎

Proposition A.3 *The preference $\succeq$ satisfies Axioms 1-6(a) if and only if there exist m functions $h_k(c, p)$, $k = 1, \ldots, m$ that are continuously differentiable and strictly quasiconvex in p and n continuous functions $A_i : \Delta_n \to R$ such that*

$$U(c, \mathcal{P}, P) = \sum_i A_i(P) + \min_{p \in \mathcal{P}} \left\{ \sum_{i \in J_1} P_i h_1(c_i, p_i/P_i) + \ldots + \sum_{i \in J_m} P_i h_m(c_i, p_i/P_i) \right\}. \qquad (10)$$

Proof: We first show that $U(c, p, P)$ is strictly quasiconvex in p. Let $p \in \Delta_n$ and $p \neq P$. By Lemma A.2 there is a unique hyperplane $\mathcal{P}(p)$ such that $U(c, \mathcal{P}(p), P) = U(c, p, P)$. Since $U(c, p, P)$ is continuously differentiable in p, this hyperplane is defined by the equation,

$$\nabla U(c, p, P)(q - p) = 0. \qquad (11)$$

That is,

$$\mathcal{P}(p) = \{q \in \Delta_n : \nabla U(c, p, P)(q - p) = 0\}$$

which is the tangent plane of $U(c, p, P)$. Now for any $q \neq p$ and $q \neq P$ that satisfies (11), we must have

$$\nabla U(c, q, P)(q - p) \neq 0.$$

(Otherwise, the set $\mathcal{P}'$ defined as the straight line going through p and q must be a subset of the tangent plane $\mathcal{P}(q)$. On the other hand, $\mathcal{P}'$ is also in $\mathcal{P}(p)$. This is a contradiction to the uniqueness claim of Lemma A.2.) By continuity, this nonzero inner product must be of constant sign for $q \neq p$ satisfying (11) when p is fixed. Further, again by continuity, the sign cannot actually depend on p. In fact,

$$\nabla U(c, q, P)(q - p) > 0 \tag{12}$$

must be true. This follows from the fact that for any q satisfying (11),

$$0 > U(c, p, P) - U(c, q, P) = \nabla U(c, q, P)(p - q) + o(\|p - q\|),$$

and hence for q close to p, $\nabla U(c, q, P)(p - q) < 0$. Thus we have obtained that (11) with $q \neq p$ always implies (12). This means that in $\mathcal{P}(p)$, $U(c, q, P)$ is strictly increasing as q moves away from p in any direction. This immediately implies that $U(c, p, P)$ is strictly quasiconvex.

Next we show that $U(c, p, P)$ has the structure claimed in the theorem. Fix a J_k. Let i and j be both in J_k and c be such that $c_i = c_j$. Let

$$\mathcal{P}_{ij} = \{p \in \Delta_n : p_i + p_j = \alpha\}.$$

Let $p^* \neq P$ be the solution of the minimization problem in (8) with $\mathcal{P} = \mathcal{P}_{ij} \cup \{P\}$. Since U is continuously differentiable in p, by Luenberger (1973,Theorem 10.3), there exist $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$ such that

$$\nabla U(c, p^*, P) = \lambda_1 1_{\{i,j\}} + \lambda_2 1_{\{1,\ldots,n\}},$$

where λ_2 is the Lagrangian multiplier for the constraint $\sum_i p_i = 1$. In light of Proposition A.1,

$$\partial f_i(c_i, p_i^*, P)/\partial p_i = \lambda_1 + \lambda_2 = \partial f_j(c_j, p_j^*, P)/\partial p_j.$$

Then Axiom 6(b) implies that, assuming that p^* is a solution to (8),

$$\partial f_i(c_i, p_i^*, P_i)/\partial p_i = \partial f_j(c_j, p_j^*, P_j)/\partial p_j \quad \text{if and only if} \quad p_i^*/P_i = p_j^*/P_j.$$

But this means there is a continuous function g_k such that

$$g_k\left(c, \frac{p^*}{P_i}\right) = \partial f_i(c_i, p_i^*, P_i)/\partial p_i,$$

provided that $P_i + P_j = \alpha$ and $p_i^* + p_j^* = t$. Since Axiom 6(a) applies to all $P_i + P_j = \alpha$, the possible range of p_i^* is $[0,1]$. Thus, as claimed,

$$f_i(c_i, p, P) - f_i(c_i, 0, P) = \int_0^p g_k(c, s/P_i)ds = P_i h_k\left(c, \frac{p}{P_i}\right), \quad h_k(c,t) = \int_0^t g_k(c,s)ds. \quad \blacksquare$$

Now we prove the first claim of the theorem. Observe that

$$\nabla U(c,q,p) = (((g_1(c_i, q_i/p_i), i \in J_1), \ldots, ((g_m(c_i, q_i/p_i), i \in J_m))$$

Observe next that $g_k(c_i, q_i/p_i)$ has a unique representation of the form

$$g_k(c_i, q_i/p_i) = \hat{g}_k(c_i, q_i/p_i) + a_k(c_i), \quad \hat{g}_k(c_i, 1) = 0.$$

Thus

$$g_k(c_i, q_i/p_i) = \hat{g}_k(c_i, q_i/p_i) + g_k(c_i, 1), \quad \hat{g}_k(c_i, 1) = 0. \tag{13}$$

Let $G(c) = ((g_1(c_1, 1), i \in J_1), \ldots, (g_m(c_m, 1), i \in J_m))$.

Suppose that c is of the form $((c_1, i \in J_1), \ldots, (c_m, i \in J_m))$. Suppose that $(c, \mathcal{P}, p) \sim (c, q, p)$ and $q \neq p$. By Lemma A.2, there is a unique $\mathcal{P}(q)$ such that $(c, \mathcal{P}(q), p) \sim (c, q, p)$. The hyperplane $\mathcal{P}(q)$ is fully characterized by the vector $\nabla U(c,q,p)$ through the equation

$$\nabla U(c,q,p)(x-q) = 0.$$

Moreover, the characterizing vector $\nabla U(c,q,p)$ is unique upto a linear transform of the form: $H_1(c)\nabla U(c,q,p) + H_2(c)$, where $H_1(c)$ and $H_2(c)$ are scalar functions. So

$$[\nabla U(c,q,p) - G(c)](x-q) = (((\hat{g}_1(c_1, q_i/p_i), i \in J_1), \ldots, ((\hat{g}_m(c_m, q_i/p_i), i \in J_m))(x-q) = 0.$$

By Axiom 6(b), this hyperplane is independent of c. This implies that the vector $(((\hat{g}_1(c_1, q_i/p_i), i \in J_1), \ldots, ((\hat{g}_m(c_m, q_i/p_i), i \in J_m))$ can be written as $F(c)((G_1(q_i/p_i), i \in J_1), \ldots, (G_m(q_i/p_i), i \in J_m))$ where $F(c)$ and $G_k(t)$, $k = 1, \ldots, m$, are scalar functions with $G_k(1) = 0$. Or $F(c)G_k(q_i/p_i) = \hat{g}_k(c_k, q_i/p_i) = g_k(c_k, q_i/p_i) - g_k(c_k, 1)$ for all k and $i \in J_k$.

Next let $\mathcal{P}$ be convex and have the property described in Axiom 6(b) and let $p \notin \mathcal{P}$. Suppose that $(c, \mathcal{P} \cup \{p\}, p) \sim (c, q, p)$ and $q \neq p$. Then

$$\nabla U(c,q,p)(x-q) = 0,$$

for all $x \in \mathcal{P}$. Since $\sum_{i\in J_k} x_i = t_k$, the above equation also holds for

$$[\nabla U(c,q,p) - G(c)](x-q) = 0.$$

Or

$$\sum_k F(c_k) \sum_{i\in J_k} G_k(q_i/p_i)(x_i - q_i) = 0. \tag{14}$$

Since $U(c,q,P)$ is strictly quasiconvex in q and $\mathcal{P}$ is convex, the above equation (first order condition) fully characterize the solution to the minimization problem

$$\min_{q\in\mathcal{P}} U(c,q,p).$$

Now we pick p and $\mathcal{P}$ such that $\sum_{i\in J_k} G_k(q_i/p_i)(x_i - q_i) \neq 0$ for at least two ks. By Axiom 6(b), q is independent of c. Then if $F(c)$ is not a constant, the equation (14) will be violated when we change $c = ((c_1, i \in J_1), \ldots, (c_m, i \in J_m))$, which is a contradiction to Axiom 6(b). Therefore we have shown

$$\nabla U(c,q,p) - G(c) = (((\hat{g}_1(q_i/p_i), i \in J_1), \ldots, ((\hat{g}_m(q_i/p_i), i \in J_m))$$

which is independent of c.

Now substitute (13) into the expression of $U(c,p,P)$ to get

$$\begin{aligned} U(c,p,P) &= \sum_i A_i(P) + \sum_k \sum_{i\in J_k} P_i h_k(c_i, p_i/P_i) = \sum_i A_i(P) + \sum_k \sum_{i\in J_k} P_i \int_0^{p_i/P_i} g_k(c_i,s)ds \\ &= \sum_i A_i(P) + \sum_k \sum_{i\in J_k} P_i \left[\int_0^{p_i/P_i} \hat{g}_k(s)ds + \frac{p_i}{P_i} g_k(c_i,1) \right]. \end{aligned}$$

Let $p = P$, we have

$$U(c,p,P) = \sum_i A_i(P) + \sum_k \sum_{i\in J_k} P_i \left[\int_0^1 \hat{g}_k(s)ds + g_k(c_i,1) \right]$$

By Axiom 7, $U(c,p,P)$ is expected utility when $p = P$. Thus $\sum_i A_i(P)$ is a constant and

$$\int_0^1 \hat{g}_k(s)ds + g_k(c,1) = u(c)$$

for some $u(c)$, independent of k. A substitution yields,

$$U(c,p,P) \;=\; \sum_k \sum_{i\in J_k} P_i \left[\int_0^{p_i/P_i} \hat{g}_k(s)ds - \frac{p_i}{P_i}\int_0^1 \hat{g}_k(s)ds\right] + \sum_i p_i u(c_i)$$

Thus,

$$U(c,p,P) \;=\; \sum_k \sum_{i\in J_k} P_i v_k(p_i/P_i) + \sum_i p_i u(c_i)$$

where v_k, $k = 1, \ldots, m$, are continuously differentiable, strictly convex, and satisfy $v_k(0) = v_k(1) = 0$. This completes the proof of the first part of Theorem 4.1.

The rest of the proof is a generalization of that of Theorem 3 of Csiszar (1991). Fix a $c = ((c_1, i \in J_1), \ldots, (c_m, i \in J_m))$. We show first that for any p, q and r in Δ_n with $\sum_{i\in J_k} r_i = \sum_{i\in J_k} q_i = t_k$, $k = 1,\ldots,m$, if

$$\nabla U(c,q,p)(r-q) = 0, \tag{15}$$

then for any $x \in \Delta_n$ with $\sum_{i\in J_k} x_i = t_k$, $k = 1,\ldots,m$,

$$\nabla U(c,q,p)(x-q) = 0, \quad \text{and} \quad \nabla U(c,r,q)(x-r) = 0, \quad \text{implies} \quad \nabla U(c,r,p)(x-r) = 0. \tag{16}$$

To see that the claim is true, let $\mathcal{P}_0$ denote the set of $x \in \Delta_n$ that satisfies the first two equations in (16). Let $\mathcal{P}_1$, $\mathcal{P}_2$ and $\mathcal{P}_3$ denote the tangent planes defined by the three equations in (16) respectively. By (15) and the first two equations of (16), $r \in \mathcal{P}_0$. Since $\mathcal{P}_0$ lies in $\mathcal{P}_2$, $(c, \mathcal{P}_0 \cup \{P\}, q) \sim (c, r, q)$ by Axiom 2. Next let $\mathcal{P}'_i$ be the subset of $\mathcal{P}_i$, $i = 0$, 1, and 2, such that all $p \in \mathcal{P}'_i$ satisfy $\sum_{i\in J_k} p_i = t_k$. By Axiom 2, $(c, \mathcal{P}'_0 \cup \{P\}, q) \sim (c, r, q)$ and $(c, \mathcal{P}'_1 \cup \{P\}, p) \sim (c, q, p)$. Apply Axiom 6 to $\mathcal{P}'_0 \subset \mathcal{P}'_1$ to get $(c, \mathcal{P}'_0 \cup \{P\}, p) \sim (c, r, p)$. But this implies, by the second part of Lemma A.2 and Axiom 2, that $\mathcal{P}'_0 \subset \mathcal{P}'_3 \subset \mathcal{P}_3$ as claimed.

Now assume that $p \gg 0$, $q \gg 0$ and $r \gg 0$ so that we do not have to worry about the nonnegativity constraint on components of p, q and r. By the claim surrounding (15) and (16), for any t_k, $k = 1, \ldots, m$, $\nabla U(c,r,p) - G(c)$ is spanned by $\nabla U(c,q,p) - G(c)$, $\nabla U(c,r,q) - G(c)$, $(1,\ldots,1)$, and $((0, i \in J_1), \ldots, (1, i \in J_k), \ldots (0, i \in J_m))$, $k = 1, \ldots, m$, where the last $m+1$ vectors come from the constraints: $\sum_i x_i = 1$, $\sum_{i\in J_k} x_i = t_k$, $k = 1, \ldots, m$. Thus, focusing on $i \in J_1$, for any p, q and r in Δ_n that satisfy (15), there exist $\alpha(c_1)$, $\beta(c_1)$ and $\gamma(c_1)$ such that

$$\alpha(c_1)[g_1(c_1, q_i/p_i) - g_1(c_1,1)] + \beta(c_1)[g_1(c_1, r_i/q_i) - g_1(c_1,1)] + \gamma(c_1) = g_1(c_1, r_i/p_i) - g_1(c_1,1) \tag{17}$$

for all $i \in J_1$, where $\gamma(c)$ is for the constraints $\sum_i x_i = 1$ and $\sum_{i \in J_1} x_i = t_1$. By assumption J_1 has more than three elements. Let $p_1 = q_1 = r_1$. Then we must have $\gamma(c) = 0$. Next pick $r_2 = q_2 \neq p_2$. Then we must have $\alpha(c_1) = 1$. Finally, pick $p_3 = q_3 \neq r_3$ and infer $\beta(c_1) = 1$. Thus we have

$$[g_1(c_1, q_i/p_i) - g_1(c_1, 1)] = g_1(c_1, r_i/p_i) - g_1(c_1, 1) - [g_1(c_1, r_i/q_i) - g_1(c_1, 1)]. \tag{18}$$

Since the left-hand side of this equation does not depend on r_i, this equation implies that

$$[g_1(c_1, q/p) - g_1(c_1, 1)] = \psi(c_1, q) - \psi(c_1, p). \tag{19}$$

But this implies

$$[g_1(c_1, ab) - g_1(c_1, 1)] = [g_1(c_1, a) - g_1(c_1, 1)] - [g_1(c_1, b) - g_1(c_1, 1)], \tag{20}$$

for a and b in R_+. By Aczel (1966, section 2.1.2),

$$g_1(c_1, a) - g_1(c_1, 1) = \phi_1(c_1) \ln a.$$

However, by the first part of the theorem, $g_1(c_1, a) - g_1(c_1, 1)$ does not depend on c, hence, $\phi_1(c_1) = \phi_1$ is a constant. Furthermore, because $U(c, p, P)$ is strictly quasiconvex in p, $\phi_1 > 0$. Integration yields

$$h_1(c, t) = \int_0^t g_1(c, a) da = \int_0^t \phi_1 \ln a da + g_1(c, 1)t = \phi_1 t \ln t - \phi_1 t + g_1(c, 1)t.$$

Letting $u_1(c) = g_1(c, 1) - \phi_1$, we have

$$\sum_k \sum_{i \in J_k} P_i h_k(c_i, p_i/P_i) = \sum_k \sum_{i \in J_k} \phi_k p_i \ln(p_i/P_i) + \sum_k \sum_{i \in J_k} p_i u_k(c_i).$$

Thus $U(c, \mathcal{P}, P) = \min_{p \in \mathcal{P}} \sum_k \sum_{i \in J_k} \phi_k p_i \ln(p_i/P_i) + \sum_k \sum_{i \in J_k} p_i u_k(c_i)$. As in the proof of the first claim of the theorem above, by Axiom 7, u_k does not depend on k. This completes the sufficiency part of the proof.

For necessity, suppose the U is represented by (6). We only need to show that it satisfies Axiom 6(c). Fix a $c = ((c_1, i \in J_1), \ldots, (c_m, i \in J_m))$ and a P. Since $\mathcal{P}'$ is convex, $(c, \mathcal{P}' \cup \{q\}, q) \sim (c, p, q)$ with $q \neq p$ if and only if for all $x \in \mathcal{P}'$,

$$0 = \nabla U(c, p, q)(x - p) = \sum_{k=1}^{m} \sum_{i \in J_k} \phi_k \ln \frac{p_i}{q_i}(x_i - p_i), \tag{21}$$

where in the second equality we have used the fact that $c = ((c_1, i \in J_1), \ldots, (c_m, i \in J_m))$ and $\sum_{i \in J_k} p_i = t_k$, $k = 1, \ldots, m$, for all $p \in \mathcal{P}'$. Since $\mathcal{P}$ is also convex, $(c, \mathcal{P} \cup \{P\}, P) \sim (c, q, P)$ with $q \neq P$ if and only if for all $x \in \mathcal{P}$,

$$\nabla U(c, q, P)(x - q) = \sum_{k=1}^{m} \sum_{i \in J_k} \phi_k \ln \frac{q_i}{P_i}(x_i - q_i) = 0. \tag{22}$$

In particular this is true for $x = p$ because $p \in \mathcal{P}' \subset \mathcal{P}$. Now consider, for $x \in \mathcal{P}'$,

$$\nabla U(c, p, P)(x - p) = \sum_{k=1}^{m} \sum_{i \in J_k} \phi_k \ln \frac{p_i}{P_i}(x_i - p_i)$$

$$= \sum_{k=1}^{m} \sum_{i \in J_k} \phi_k \ln \frac{p_i}{q_i}(x_i - p_i) + \sum_{k=1}^{m} \sum_{i \in J_k} \phi_k \ln \frac{q_i}{P_i}(x_i - p_i)$$

$$= \sum_{k=1}^{m} \sum_{i \in J_k} \phi_k \ln \frac{p_i}{q_i}(x_i - p_i) + \sum_{k=1}^{m} \sum_{i \in J_k} \phi_k \ln \frac{q_i}{P_i}(x_i - q_i) + \sum_{k=1}^{m} \sum_{i \in J_k} \phi_k \ln \frac{q_i}{P_i}(q_i - p_i)$$

This expression is equal to zero because of (21), (22) and that $q \in \mathcal{P}$. Therefore, $(c, \mathcal{P}' \cup \{P\}, P) \sim (c, p, P)$ as desired. ∎

References

[1] Aczel (1966): *Lectures on Functional Equations and Their Applications.* Academic, New York.

[2] Anderson, E., L. Hansen and T. Sargent (1999): "Robustness, Detection and the Price of Risk," working paper, University of Chicago.

[3] Camerer, C.F. and M. Weber (1992): "Recent Developments in Modeling Preferences: Uncertainty and Ambiguity," *Journal of Risk and Uncertainty*, 8, pp.167-196.

[4] Casadesus-Masanell, R., P. Klibanoff and E. Ozdenoren (2000): "Maxmin Expected Utility over Savage Acts with a Set of Priors," *Journal of Economic Theory*, vol.92, pp.53-65.

[5] Chen, Z. and L. Epstein (2002): "Ambiguity, risk and asset returns in continuous time," forthcoming *Econometrica*, vol.70, pp.1403-1443 .

[6] Csiszar, I. (1991): "Why Least Squares and Maximum Entropy? an Axiomatic Approach to Inference for Linear Inverse Problems," *Annals of Statistics*, Vol.19, pp.1032-2066.

[7] Debreu, G. (1954): "Representation of a Preference Ordering by a Numerical Function," in *Decision Processes,* R. Thrall, C. Coombs and R. David, eds., New York, John Wiley.

[8] Debreu, G. (1959): "Topological Methods in Cardinal Utility Theory," in *Mathematical Methods in Social Sciences*, ed. by Arrow, Karlin and Suppes.

[9] Dow, J. and S. Werlang (1992): "Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio", *Econometrica,* 60, pp.197-204.

[10] Ellsberg, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics,* 75, pp.643-669.

[11] Epstein, L. (2001): "Sharing Ambiguity," *American Economic Review,*, Vol. 91, pp.45-50.

[12] Esptein, L. and J. Miao (2001): "A two-person dynamic equilibrium under ambiguity," forthcoming *Journal of Economic Dynamics and Control.*

[13] Epstein, L. and T. Wang (1994): "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 62, pp.283-322.

[14] Epstein, L. and T. Wang (1995): "Uncertainty, Risk-Neutral Measures and Security Price Booms and Crashes," *Journal of Economic Theory*, 67, pp.40-80.

[15] Fox, C. and A. Tversky (1995): "Ambiguity Aversion and Comparative Ignorance," *Quarterly Journal of Economics*, 110, pp.585-603.

[16] Gilboa, I. (1987): "Expected Utility Theory with Purely Subjective Non-Additive Probabilities," *Journal of Mathematical Economics,* 16, pp.65-88.

[17] Gilboa, I. and D. Schmeidler (1989): "Maxmin Expected Utility Theory with Non-Unique Prior," *Journal of Mathematical Economics,* 18, pp.141-153.

[18] Hansen, L. and T. Sargent (2001): "Robust Control and Model Uncertainty," *American Economic Review,*, Vol. 91, pp.60-66.

[19] Klibanoff, P., M. Marinacci and S. Mukerji (2002): " A Smooth Model of Decision Making Under Ambiguity," Northwestern University working paper.

[20] Knight, F. (1921): *Risk, Uncertainty and Profit*, Houghton, Mifflin, Boston.

[21] Koopmans, T. (1972): "Representation of Preference Orderings with Independent Components of Consumption," in *Decision and Organization*, ed. by McGuire and Radner.

[22] Kreps, D. (1988): *Notes on the Theory of Choice*, Westview Press.

[23] Luenberger, D. (1973) *Introduction to Linear and Nonlinear Programming.* Addison-Wesley.

[24] Machina, M. and D. Schmeidler (1992): "A More Robust Definition of Subjective Probability," *Econometrica*, 60, pp.745-780.

[25] Maenhout, P. (1999): "Robust portfolio rules and asset pricing," Working paper, Harvard University.

[26] Nakamura, Y. (1990): "Subjective Expected Utility with Non-additive Probabilities on Finite State Spaces," *Journal of Economic Theory*, 51, pp.346-366.

[27] Sarin, R. and P. Wakker (1998): "A Simple Axiomatization of Nonadditive Expected Utility," *Econometrica,* 60, pp.1255-1272.

[28] Sarin, R. and P. Wakker (1998): "Revealed Likelihood and Knightian Uncertainty," *Journal of Risk and Uncertainty,* 16, pp.223-250.

[29] Savage, L. (1954): *Foundations of Statistics,* Wiley: New York.

[30] Schmeidler, D. (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica,* 57, pp.571-587.

[31] Uppal, R. and T. Wang (2002): "Model Misspecification and Under-Diversification," forthcoming in *Journal of Finance.*