Axiomatic Foundations of Multiplier Preferences*

Tomasz Strzalecki†
Northwestern University

Abstract

This paper axiomatizes the robust control criterion of multiplier preferences introduced by Hansen and Sargent (2001). The axiomatization relates multiplier preferences to other classes of preferences studied in decision theory. Some properties of multiplier preferences are generalized to the broader class of variational preferences, recently introduced by Maccheroni, Marinacci and Rustichini (2006). The paper also establishes a link between the parameters of the multiplier criterion and the observable behavior of the agent. This link enables measurement of the parameters on the basis of observable choice data and provides a useful tool for applications.

*I am indebted to my advisor Eddie Dekel for his continuous guidance, support, and encouragement. I am grateful to Peter Klibanoff and Marciano Siniscalchi for many discussions which resulted in significant improvements of the paper. I would also like to thank Jeff Ely and Todd Sarver for helpful comments and suggestions. This project started after a very stimulating conversation with Tom Sargent and was further shaped by conversations with Lars Hansen. All errors are my own.

†Department of Economics, Northwestern University. E-mail: `tomasz@northwestern.edu`.
This version: November 7, 2007.

1 Introduction

The concept of uncertainty has been studied by economists since the work of Keynes (1921) and Knight (1921). As opposed to risk, where probability is well specified, uncertainty, or ambiguity, is characterized by the decision maker's inability to formulate a single probability or by his lack of trust in any unique probability.

Indeed, as demonstrated by Ellsberg (1961), people often make choices that cannot be justified by a unique probability, thereby exhibiting a preference for risky choices over those involving ambiguity. Such *ambiguity aversion* has been one of the central issues in decision theory, motivating the development of axiomatic models of such behavior.[1]

The lack of trust in a single probability has also been a source of concern in macroeconomics. In order to capture concern about model misspecification, Hansen and Sargent (2001) formulated an important model of *multiplier preferences*. Thanks to their great tractability, multiplier preferences are now being adopted in applications.[2]

Despite their importance in macroeconomics, multiplier preferences have not been fully understood at the level of individual decision making. Although Maccheroni et al. (2006a) showed that they are a special case of the variational preferences that they axiomatized, an axiomatization of multiplier preferences has so far been elusive. Indeed, some authors even doubted the existence of behaviorally meaningful axioms that would pin down multiplier preferences within the broad class of variational preferences.

The main contribution of this paper is precisely a set of axioms satisfying this property. The proposed axiomatic characterization is important for three reasons. First, it provides a set of testable predictions of the model that allow for its empirical verification. This will help evaluate whether multiplier preferences, which are useful in modeling behavior at the macro level, are an accurate model of individual behavior. Second, the axiomatization establishes a link between the parameters of the multiplier criterion and the observable behavior of the agent. This link en-

[1]See, e.g., Gilboa and Schmeidler (1989); Schmeidler (1989); Ergin and Gul (2004); Klibanoff, Marinacci, and Mukerji (2005); Maccheroni, Marinacci, and Rustichini (2006a).

[2]See, e.g. Woodford (2006); Barillas, Hansen, and Sargent (2007); Karantounias, Hansen, and Sargent (2007); Kleshchelski and Vincent (2007).

ables measurement of the parameters on the basis of observable choice data alone, without relying on unverifiable assumptions. Finally, the axiomatization is helpful in understanding the relation between multiplier preferences and other axiomatic models of preferences and ways in which they can and cannot be used for modeling Ellsberg-type behavior.

1.1 Background and Overview of Results

The Expected Utility criterion ranks payoff profiles f according to

$$V(f) = \int u(f)\,\mathrm{d}q, \tag{1}$$

where u is a utility function and q is a subjective probability distribution on the states of the world. A decision maker with such preferences is considered ambiguity neutral, because he is able to formulate a single probability that governs his choices.

In order to capture lack of trust in a single probability, Hansen and Sargent (2001) formulated the following criterion

$$V(f) = \min_{p} \int u(f)\,\mathrm{d}p + \theta R(p \,\|\, q), \tag{2}$$

where $\theta \in (0, \infty]$ is a parameter and function $R(p \,\|\, q)$ is the *relative entropy* of p with respect to q. Relative entropy, otherwise known as Kullback-Leibler divergence, is a measure of "distance" between two probability distributions. An interpretation of (2) is that the decision maker has some best guess q of the true probability distribution, but does not fully trust it. Instead, he considers other probabilities p to be plausible, with plausibility diminishing proportionally to their "distance" from q. The role of the proportionality parameter θ is to measure the degree of trust of the decision maker in the reference probability q. Higher values of θ correspond to more trust; in the limit, when $\theta = \infty$, the decision maker fully trusts his reference probability and uses the expected utility criterion (1).

Multiplier preferences also belong to the more general class of variational preferences studied by Maccheroni et al. (2006a); those preferences have the following representation:

$$V(f) = \min_p \int u(f)\,\mathrm{d}p + c(p), \tag{3}$$

where $c(p)$ is a "cost function". The interpretation of (3) is like that of (2), and multiplier preferences are a special case of variational preferences with $c(p) = \theta R(p \| q)$. In general, the conditions that the function $c(p)$ in (3) has to satisfy are very weak, which makes variational preferences a very broad class. In addition to expected utility preferences and multiplier preferences, this class also nests the maxmin expected utility preferences of Gilboa and Schmeidler (1989), as well as the mean-variance preferences of Markowitz (1952) and Tobin (1958).

An important contribution of Maccheroni et al. (2006a) was to provide an axiomatic characterization of variational preferences. However, because variational preferences are a very broad class of preferences, it would be desirable to establish an observable distinction between multiplier preferences and other subclasses of variational preferences. Ideally, an axiom, or set of axioms, would exist that, when added to the list of axioms of Maccheroni et al. (2006a), would deliver multiplier preferences. This is, for example, the case with the maxmin expected utility preferences of Gilboa and Schmeidler (1989): a strengthening of one of the Maccheroni et al.'s (2006a) axioms restricts the general cost function $c(p)$ to be in the class used in Gilboa and Schmeidler's (1989) model. The reason for skepticism about the existence of an analogous strengthening in the case of multiplier preferences has been that the relative entropy $R(p \| q)$ is a very specific functional-form assumption, which does not seem to have any behaviorally significant consequences. The main finding of this paper is that these consequences *are* behaviorally significant. The main theorem shows that standard axioms characterize the class of multiplier preferences within the class of variational preferences.

1.2 Ellsberg's Paradox and Measurement of Parameters

Ellsberg's (1961) experiment demonstrates that most people prefer choices involving risk (i.e., situations in which the probability is well specified) to choices involving ambiguity (where the probability is not specified). Consider two urns containing colored balls. The decision maker can bet on the color of the ball drawn from each urn. Urn I contains 100 red and black balls in unknown proportion, while Urn II contains 50 red and 50 black balls.

In this situation, most people are indifferent between betting on red from Urn I and on black from Urn I. This reveals that, in the absence of evidence against symmetry, they view those two contingencies as interchangeable. Moreover, most people are indifferent between betting on red from Urn II and on black from Urn II. This preference is justified by their knowledge of the composition of Urn II. However, most people strictly prefer betting on red from Urn II to betting on red from Urn I, thereby displaying ambiguity aversion.

Ambiguity aversion cannot be reconciled with a single probability governing the distribution of draws from Urn I. For this reason, expected utility preferences are incapable of explaining the pattern of choices revealed by Ellsberg's experiment. Such pattern can, however, be explained by multiplier preferences. Recall that

$$V(f) = \min_{p} \int u(f)\,\mathrm{d}p + \theta R(p\| q). \tag{2}$$

The curvature of the utility function u measures the decision maker's risk aversion and governs his choices when probabilities are well specified—for example, choices between bets on red and black from Urn II. In contrast, the parameter θ measures the decision maker's attitude towards ambiguity, and influences his choices when probabilities are not well specified—for example, choices between bets on red and black from Urn I.

Formally, betting \$100 on red from Urn II corresponds to an objective lottery r_{II} paying \$100 with probability $\frac{1}{2}$ and \$0 with probability $\frac{1}{2}$. Betting \$100 on black from Urn II corresponds to lottery b_{II}, which is equivalent to r_{II}. The decision maker values r_{II} and b_{II} at

$$V(r_{II}) = V(b_{II}) = \frac{1}{2}u(100) + \frac{1}{2}u(0).$$

Moreover, let x denote the certainty equivalent of r_{II} and b_{II}, i.e., the amount of money that, when received for sure, would be indifferent to r_{II} and b_{II}. Formally

$$V(x) = u(x) = V(r_{II}) = V(b_{II}). \tag{4}$$

On the other hand, betting \$100 on red from Urn I corresponds to r_I, which pays \$100 when a red ball is drawn and \$0 otherwise. Similarly, betting \$100 on black from Urn I corresponds to b_I, which pays \$100 when a black ball is drawn and \$0 otherwise. The decision maker values r_I and b_I at

$$V(r_I) = V(b_I) = \min_{p \in [0,1]} pu(100) + (1-p)u(0) + \theta R(p \| q)$$

where q is the reference measure, assumed to put equal weights on red and black. Moreover, let y be the certainty equivalent of r_I and b_I, i.e., the amount of money that, when received for sure, would be indifferent to r_I and b_I. Formally

$$V(y) = u(y) = V(r_I) = V(b_I). \tag{5}$$

In Ellsberg's experiments most people prefer objective risk to subjective uncertainty, implying that $y < x$. This pattern of choices is implied by multiplier preferences with $\theta < \infty$. The equality $y = x$ holds only when $\theta = \infty$, i.e., when preferences are expected utility and there is no ambiguity aversion.

Ellsberg's paradox provides a natural setting for experimental measurement of parameters of the model. The observable choice data reveals the decision maker's preferences over objective lotteries, and hence his aversion toward pure risk embodied in the utility function u. The observed value of certainty equivalent x allows to infer the curvature of u.[3] Similarly, decision maker's choices between uncertain gambles reveal his attitude toward subjective uncertainty, represented by parameter θ. The observed "ambiguity premium" $x - y$ enables inferences about the value of θ: a big difference $x - y$ reveals that the decision maker has low trust in his reference probability, i.e., θ is low.[4]

The procedure described above suggests that simple choice experiments could be used for empirical measurement of both u and θ. Such revealed-preference measurement of parameters would be a useful tool in applied settings, where it is important to know the numerical values of parameters, and would be complementary to the heuristic method of detection error probabilities developed by Anderson, Hansen, and Sargent (2000) and Hansen and Sargent (2007).

[3] For example, let $u(z) = (w+z)^{1-\gamma}$, where w is the initial level of wealth. Then (4) establishes a 1-1 relationship between x and γ. The value of γ can be derived from observed values of x and w.

[4] Continuing the example from footnote 3, holding γ and w fixed, (5) establishes a 1-1 relationship between y and θ. Thus, the value of θ can be derived from observed values of y, x, and w.

1.3 Outline of the Paper

The paper is organized as follows. After introducing some notation and basic concepts in Section 2, Section 3 defines static multiplier preferences, discusses their properties in the classic setting of Savage, and indicates that richer choice domains are needed for axiomatization. Section 4 uses one of such richer domains, introduced by Anscombe-Aumann, and discusses the class of variational preferences, which nests multiplier preferences. Section 4 presents axioms that characterize the class of multiplier preferences within the class of variational preferences. Additionally, extending a result of Marinacci (2002), Section 4 discusses the extent to which variational preferences can be used for modelling the Allais paradox. Section 5 studies a different enrichment of choice domain and presents an axiomatization of multiplier preferences in a setting introduced by Ergin and Gul (2004), thereby obtaining a fully subjective axiomatization of multiplier preferences. Section 6 presents an axiomatization of dynamic multiplier preferences, which are central to applications in macroeconomics and finance. Because of its relation to the model of Kreps and Porteus (1978), this class exhibits a preference for earlier resolution of uncertainty. Section 6 establishes that such preference is exhibited by all stationary variational preferences, except for the subclass of maxmin expected utility preferences. Section 7 concludes.

2 Preliminaries

Decision problems considered in this paper involve a set S of *states of the world*, which represents all possible contingencies that may occur. One of the states, $s \in S$, will be realized, but the decision maker has to choose the course of action before learning s. His possible choices, called acts, are mappings from S to Z, the set of consequences. Each act is a complete description of consequences, contingent on states.

Formally, let Σ be a sigma-algebra of subsets of S. An act is a finite-valued, Σ-measurable function $f : S \to Z$; the set of all such acts is denoted $\mathcal{F}(Z)$. If $f, g \in \mathcal{F}(Z)$ and $E \in \Sigma$, then fEg denotes an act with $fEg(s) = f(s)$ if $s \in E$ and $fEg(s) = g(s)$ if $s \notin E$. The set of all finitely additive probability measures on (S, Σ) is denoted $\Delta(S)$; the set of all countably additive probability measures is denoted $\Delta^\sigma(S)$; its subset consisting of all measures absolutely continuous with respect to $q \in \Delta^\sigma(S)$ is denoted $\Delta^\sigma(q)$.

The choices of the decision maker are represented by a preference relation $\succsim$, where $f \succsim g$ means that the act f is weakly preferred to the act g. A functional $V : \mathcal{F}(Z) \to \mathbb{R}$ *represents* $\succsim$ if for all $f, g \in \mathcal{F}(Z)$ $f \succsim g$ if and only if $V(f) \geq V(g)$.

An important class of preferences are Expected Utility (EU) preferences, where the decision maker has a probability distribution $q \in \Delta(S)$ and a utility function which evaluates each consequence $u : Z \to \mathbb{R}$. A preference relation $\succsim$ has an *EU representation* (u, q) if there exists a functional $V : \mathcal{F}(Z) \to \mathbb{R}$ that represents $\succsim$ with $V(f) = \int_S (u \circ f) \, \mathrm{d}q$.

Let $Z = \mathbb{R}$, i.e., acts have monetary payoffs. Risk aversion is the phenomenon where sure payoffs are preferred to ones that are stochastic but have the same expected monetary value. Risk averse EU preferences have concave utility functions u. Likewise, one preference relation is more risk averse than another if it has a "more concave" utility function. More formally, a preference relation represented by (u_1, q_1) is *more risk averse* than one represented by (u_2, q_2) if and only if $q_1 = q_2$ and $u_1 = \phi \circ u_2$, where $\phi : \mathbb{R} \to \mathbb{R}$ is an increasing concave transformation.

A special role will be played by the class of transformations ϕ_θ, indexed by $\theta \in (0, \infty]$

$$\phi_\theta(u) = \begin{cases} -\exp\left(-\frac{u}{\theta}\right) & \text{for } \theta < \infty, \\ u & \text{for } \theta = \infty. \end{cases} \tag{6}$$

Lower values of θ correspond to "more concave" transformations, i.e., more risk aversion.

3 Concern about model misspecification

3.1 Model Uncertainty

A decision maker with expected utility preferences formulates a probabilistic model of the world, embodied by the subjective distribution $q \in \Delta(S)$. However, in many situations, a single probability cannot explain people's choices, as illustrated by the Ellsberg paradox.

Example 1 (Ellsberg Paradox)**.** Consider two urns containing colored balls; the decision maker can bet on the color of the ball drawn from each urn. Urn I contains 100 red and black balls in unknown proportion, while Urn II contains 50 red and 50 black balls.

In this situation, most people are indifferent between betting on red from Urn I and on black from Urn I. This reveals that they view those two contingencies as interchangeable. Moreover, most people are indifferent between betting on red from Urn II and on black from Urn II. This preference is justified by their knowledge of the composition of Urn II. However, most people strictly prefer betting on red from Urn II to betting on red from Urn I, thereby avoiding decisions based on imprecise information. Such a pattern of preferences cannot be reconciled with a single probability distribution, hence the paradox. ▲

In addition to this descriptive failure, a single probabilistic model of the world may also be too strong an assumption from a normative, or frequentist point of view. In many situations the decision maker may not have enough information to formulate a single probabilistic model. For example, it may be hard to statistically distinguish between similar probabilistic models, and thus hard to select one model and have full confidence in it. Hansen, Sargent, and coauthors (Hansen and Sargent, 2001; Hansen, Sargent, Turmuhambetova, and Williams, 2006) introduced a way of modelling such situations. In their model the decision maker does not know the true probabilistic model p, but has a "best guess", or *approximating model* q, also called a *reference probability*. The decision maker thinks that the true probability p is somewhere near to the approximating probability q. The notion of distance used by Hansen and Sargent is relative entropy.

Definition 1. Let a reference measure $q \in \Delta^\sigma(S)$ be fixed. The *relative entropy* $R(\cdot \| q)$ is a mapping from $\Delta(S)$ into $[0, \infty]$ defined by

$$R(p\| q) = \begin{cases} \int_S (\log \frac{\mathrm{d}p}{\mathrm{d}q})\,\mathrm{d}p & \text{if } p \in \Delta^\sigma(q), \\ \infty & \text{otherwise.} \end{cases}$$

A decision maker who is concerned with model misspecification computes his expected utility according to all probabilities p, but he does not treat them equally. Probabilities closer to his "best guess" have more weight in his decision.

Definition 2. A relation $\succsim$ has a *multiplier representation* if it is represented by

$$V(f) = \min_{p \in \Delta(S)} \int_S (u \circ f)\, \mathrm{d}p + \theta R(p \| q),$$

where $u : Z \to \mathbb{R}$, $q \in \Delta^\sigma(S)$, and $\theta \in (0, \infty]$. In this case, $\succsim$ is called a *multiplier preference.*

The multiplier representation of $\succsim$ may suggest the following interpretation. First, the decision maker chooses an act without knowing the true distribution p. Second, "Nature" chooses the probability p in order to minimize the decision maker's expected utility. Nature is not free to choose, but rather it incurs a "cost" for using each p. Probabilities p that are farther from the reference measure q have a larger potential for lowering the decision maker's expected utility, but Nature has to incur a larger cost in order to select them.

This interpretation suggests that a decision maker with such preferences is concerned with model misspecification and makes decisions that are robust to such misspecification. He is pessimistic about the outcome of his decision which leads him to exercise caution in choosing the course of action.[5] Such cautious behavior is reminiscent of Ellsberg's paradox from Example 1. However, the following theorem shows that such caution is equivalent to increased risk aversion.

3.2 Link to Increased Risk Aversion

The following variational formula (see, e.g., Proposition 1.4.2 of Dupuis and Ellis, 1997) plays a critical role in the analysis and applications of multiplier preferences.

$$\min_{p \in \Delta S} \int_S (u \circ f)\, \mathrm{d}p + \theta R(p \| q) = -\theta \log \left(\int_S \exp\left(-\frac{u \circ f}{\theta} \right) \mathrm{d}q \right). \tag{7}$$

This formula links model uncertainty, as represented by the left hand side of formula (7), to increased risk aversion, as represented by the right hand side of for-

[5]Hansen and Sargent also study a closely related class of *constraint preferences*, represented by $V(f) = \min_{\{p | R(p\|q) \leq \eta\}} \int_S (u \circ f)\, \mathrm{d}p$, which are a special case of Gilboa and Schmeidler's (1989) maxmin expected utility preferences. Due to their greater analytical tractability, multiplier, rather than constraint, preferences are used in the analysis of economic models (see, e.g., Woodford, 2006; Barillas et al., 2007; Karantounias et al., 2007; Kleshchelski and Vincent, 2007).

mula (7). Jacobson (1973), Whittle (1981), Skiadas (2003), and Maccheroni, Marinacci, and Rustichini (2006b) showed that in dynamic settings this link manifests itself as an observational equivalence between dynamic multiplier preferences and a (subjective analogue of) Kreps and Porteus (1978) preferences. As a consequence, in a static Savage setting multiplier preferences become expected utility preferences.

Observation 1. *The relation $\succsim$ has a multiplier representation (θ, u, q) if and only if $\succsim$ has an EU representation V with*

$$V(f) = \int_S (\phi_\theta \circ u \circ f) \, \mathrm{d}q, \tag{8}$$

where the transformation ϕ_θ is defined by (6).

Corollary 1. *If $\succsim$ has a multiplier representation, then it has an EU representation with utility bounded from above. Conversely, if $\succsim$ has an EU representation with utility bounded from above, then for any $\theta \in (0, \infty]$ preference $\succsim$ has a multiplier representation with that θ.*[6]

This observation suggests that multiplier preferences do not reflect model uncertainty, because the decision maker bases his decisions on a well specified probability distribution. For the same reason such preferences cannot be used for modeling Ellsberg's paradox in the Savage setting.

More importantly, given a multiplier preference $\succsim$, only the function $\phi_\theta \circ u$ is identified in absence of additional assumptions. Because of this lack of identification, there is no way of disentangling risk aversion (curvature of u) from concern about model misspecification (value of θ).

Example 2. Consider a multiplier preference $\succsim_1$ with $u_1(x) = -\exp(-x)$ and $\theta_1 = \infty$. This representation suggests that the decision maker $\succsim_1$ is risk averse, while not being concerned about model misspecification or ambiguity. In contrast, consider a multiplier preference $\succsim_2$ with $u_2(x) = x$ and $\theta_2 = 1$. This representation suggests that the decision maker with $\succsim_2$ is risk neutral, while being concerned about model misspecification or ambiguity.

[6] It can be verified that $\succsim$ has an EU representation with utility bounded from above if and only if $\succsim$ has an EU representation and the following axiom is satisfied: There exist $z \prec z'$ in Z and a non-null event E, such that $wEz \prec z'$ for all $w \in Z$. According to Corollary 1, in the Savage setting this axiom is the only behavioral consequence of multiplier preferences beyond expected utility.

Despite the apparent differences between $\succsim_1$ and $\succsim_2$, it is true that $\phi_{\theta_1} \circ u_1 = \phi_{\theta_2} \circ u_2$, so, by 1, the two preference relations are identical. Hence, the two decision makers behave in exactly the same way and there are no observable differences between them. ▲

This lack of identification means that, within this class of models, choice data alone is not sufficient to distinguish between risk aversion and ambiguity. As a consequence, any econometric estimation of a model involving such decision makers would not be possible without additional ad-hoc assumptions about parameters. Likewise, policy recommendations based on such a model would depend on a somewhat arbitrary choice of the representation. Different representations of the same preferences could lead to different welfare assessments and policy choices, but such choices would not be based on observable data.[7]

Sections 4 and 5 present two ways of enriching the domain of choice and thereby making the distinction between model uncertainty and risk aversion based on observable choice data. In both axiomatizations the main idea is to introduce a subdomain of choices where, either by construction or by revealed preference, the decision maker is not concerned about model misspecification. This subdomain serves as a point of reference and makes it possible to distinguish between concern for model misspecification (and related to it Ellsberg-type behavior) and Expected Utility maximization.

4 Axiomatization with Objective Risk

This section discusses an extension of the domain of choice to the Anscombe-Aumann setting, where objective risk coexists with subjective uncertainty. In this setting a recent model of *variational preferences* (introduced and axiomatized by Maccheroni et al., 2006a) nests multiplier preferences as a special case. Despite this classification, additional axioms that, together with the axioms of Maccheroni et al. (2006a), would deliver multiplier preferences have so far been elusive. This section presents such axioms. It is also shown that in the Anscombe-Aumann setting multiplier preferences can be distinguished from expected utility on the basis of Ellsberg-type experiments.

[7] See, e.g., Barillas et al. (2007), who study welfare consequences of eliminating model uncertainty. The evaluation of such consequences depends on the value of parameter θ.

4.1 Introducing Objective Risk

One way of introducing objective risk into the present model is to replace the set Z of consequences with (simple) probability distributions on Z, denoted $\Delta(Z)$.[8] An element of $\Delta(Z)$ is called a lottery. A lottery paying off $z \in Z$ for sure is denoted δ_z. For any two lotteries $\pi, \pi' \in \Delta(Z)$ and a number $\alpha \in (0,1)$ the lottery $\alpha\pi + (1-\alpha)\pi'$ assigns probability $\alpha\pi(z) + (1-\alpha)\pi'(z)$ to each prize $z \in Z$.

Given this specification, preferences are defined on acts in $\mathcal{F}(\Delta(Z))$. Every such act $f : S \to \Delta(Z)$ involves two sources of uncertainty: first, the payoff of f is contingent on the state of the world, for which there is no objective probability given; second, given the state, f_s is an objective lottery.

The original axioms of Anscombe and Aumann (1963) and Fishburn (1970) impose the same attitude towards those two sources. They imply the existence of a utility function $u : Z \to \mathbb{R}$ and a subjective probability distribution $q \in \Delta(S)$ such that each act is evaluated by

$$V(f) = \int_S \left(\sum_{z \in Z} u(z) f_s(z) \right) \mathrm{d}q(s). \tag{9}$$

Thus, in each state of the world s the decision maker computes the expected utility of the lottery f_s and then averages those values across states. By slightly abusing notation, define $u : \Delta(Z) \to \mathbb{R}$ by $u(\pi) = \sum_{z \in Z} u(z)\pi(z)$. Using this definition, the Anscombe-Aumann Expected Utility criterion can be written as

$$V(f) = \int_S u(f_s)\, \mathrm{d}q(s).$$

4.2 Multiplier Preferences

In this environment, the multiplier preferences take the following form

$$V(f) = \min_{p \in \Delta S} \int_S u(f_s)\, \mathrm{d}p + \theta R(p \| q), \tag{10}$$

[8] This particular setting was introduced by Fishburn (1970); settings of this type are usually named after Anscombe and Aumann (1963), who were the first to work with them.

The decision maker with such preferences makes a distinction between objective risk and subjective uncertainty: he uses the expected utility criterion to evaluate lotteries, while using the multiplier criterion to evaluate acts.

4.3 Variational Preferences

To capture ambiguity aversion, Maccheroni et al. (2006a) introduce a class of variational preferences, with representation

$$V(f) = \min_{p \in \Delta S} \int_S u(f_s)\, \mathrm{d}p + c(p), \tag{11}$$

where $c : \Delta S \to [0, \infty]$ is a *cost function.*

Multiplier preferences are a special case of variational preferences where $c(p) = \theta R(p \,\|\, q)$. The variational criterion (11) can be given the same interpretation as the multiplier criterion (10): Nature wants to reduce the decision maker's expected utility by choosing a probability distribution p, but she is not entirely free to choose. Using different p's leads to different values of the decision maker's expected utility $\int_S u(f_s)\, \mathrm{d}p$, but comes at a cost $c(p)$.

In order to characterize variational preferences behaviorally, Maccheroni et al. (2006a) use the following axioms.

Axiom A1 *(Weak Order).* The relation $\succsim$ is transitive and complete.

Axiom A2 *(Weak Certainty Independence).* For all $f, g \in \mathcal{F}(\Delta(Z))$, $\pi, \pi' \in \Delta(Z)$, and $\alpha \in (0, 1)$,

$$\alpha f + (1-\alpha)\pi \succsim \alpha g + (1-\alpha)\pi \Rightarrow \alpha f + (1-\alpha)\pi' \succsim \alpha g + (1-\alpha)\pi'.$$

Axiom A3 *(Continuity).* For any $f, g, h \in \mathcal{F}(\Delta(Z))$ the sets $\{\alpha \in [0,1] \mid \alpha f + (1-\alpha)g \succsim h\}$ and $\{\alpha \in [0,1] \mid h \succsim \alpha f + (1-\alpha)g\}$ are closed.

Axiom A4 *(Monotonicity).* If $f, g \in \mathcal{F}(\Delta(Z))$ and $f(s) \succsim g(s)$ for all $s \in S$, then $f \succsim g$.

Axiom A5 *(Uncertainty Aversion).* If $f, g \in \mathcal{F}(\Delta(Z))$ and $\alpha \in (0, 1)$, then

$$f \sim g \Rightarrow \alpha f + (1-\alpha)g \succsim f.$$

Axiom A6 *(Nondegeneracy).* $f \succ g$ for some $f, g \in \mathcal{F}(\Delta(Z))$.

Axiom A7 *(Unboundedness).* There exist $\pi' \succ \pi$ in $\Delta(Z)$ such that, for all $\alpha \in (0,1)$, there exists $\rho \in \Delta(Z)$ that satisfies either $\pi \succ \alpha\rho + (1-\alpha)\pi'$ or $\alpha\rho + (1-\alpha)\pi \succ \pi'$.

Axiom A8 *(Weak Monotone Continuity).* If $f, g \in \mathcal{F}(\Delta(Z))$, $\pi \in \Delta(Z)$, $\{E_n\}_{n\geq 1} \in \Sigma$ with $E_1 \supseteq E_2 \supseteq \cdots$ and $\bigcap_{n\geq 1} E_n = \emptyset$, then $f \succ g$ implies that there exists $n_0 \geq 1$ such that $\pi E_{n_0} f \succ g$.

Maccheroni et al. (2006a) show that preference $\succsim$ satisfies Axioms A1-A6 if and only if $\succsim$ is represented by (11) with a non-constant $u : \Delta(Z) \to \mathbb{R}$ and $c : \Delta S \to [0, \infty]$ that is convex, lower semicontinuous, and grounded (achieves value zero). Moreover, Axiom A7 implies unboundedness of the utility function u, which guarantees uniqueness of the cost function c, while Axiom A8 guarantees that function c is concentrated only on countably additive measures.

The conditions that the cost function c satisfies are very general. For example, if $c(p) = \infty$ for all measures $p \neq q$, then (11) reduces to (9), i.e., preferences are expected utility. Axiomatically, this can be obtained by strengthening Axiom A2 to

Axiom A2' *(Independence).* For all $f, g, h \in \mathcal{F}(\Delta(Z))$ and $\alpha \in (0,1)$,

$$f \succsim g \Leftrightarrow \alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h.$$

Similarly, setting $c(p) = 0$ for all measures p in a closed and convex set C and $c(p) = \infty$ otherwise, denoted $c = \delta_C$, reduces (11) to

$$V(f) = \min_{p\in C} \int_S \left(\sum_{z\in Z} u(z) f_s(z) \right) \mathrm{d}p,$$

which is a representation of the Maxmin Expected Utility preferences introduced by Gilboa and Schmeidler (1989). Axiomatically, this can be obtained by strengthening Axiom A2 to

Axiom A2" *(Certainty Independence).* For all $f, g \in \mathcal{F}(\Delta(Z))$, $\pi \in \Delta(Z)$ and $\alpha \in (0,1)$,

$$f \succsim g \Leftrightarrow \alpha f + (1-\alpha)\pi \succsim \alpha g + (1-\alpha)\pi.$$

As mentioned before, multiplier preferences also are a special case of variational preferences. They can be obtained by setting $c(p) = \theta R(p \,\|\, q)$. However, because relative entropy is a specific functional form assumption, Maccheroni et al. (2006a) were skeptical that a counterpart of Axiom A2' or Axiom A2" exists that would deliver multiplier preferences:

> [...] we view entropic preferences as essentially an analytically convenient specification of variational preferences, much in the same way as, for example, Cobb-Douglas preferences are an analytically convenient specification of homothetic preferences. As a result, in our setting there might not exist behaviorally significant axioms that would characterize entropic preferences (as we are not aware of any behaviorally significant axiom that characterizes Cobb-Douglas preferences).

Despite this seeming impasse, the next section shows that pinning down the functional form is possible with behaviorally significant axioms. In fact, somewhat unexpectedly, they are the well known Savage's P2 and P4 axioms (together with his technical axiom of continuity—P6).[9]

4.4 Axiomatization of Multiplier Preferences

Axiom P2 *(Savage's Sure-Thing Principle).* For all $E \in \Sigma$ and $f, g, h, h' \in \mathcal{F}(\Delta(Z))$

$$fEh \succsim gEh \Rightarrow fEh' \succsim gEh'.$$

Axiom P4 *(Savage's Weak Comparative Probability).* For all $E, F \in \Sigma$ and $\pi, \pi', \rho, \rho' \in \Delta(Z)$ such that $\pi \succ \rho$ and $\pi' \succ \rho'$

$$\pi E\rho \succsim \pi F\rho \Rightarrow \pi' E\rho' \succsim \pi' F\rho'.$$

Axiom P6 *(Savage's Small Event Continuity).* For all acts $f \succ g$ and $\pi \in \Delta(Z)$, there exists a finite partition $\{E_1, \ldots, E_n\}$ of S such that for all $i \in \{1, \ldots, n\}$

$$f \succ \pi E_i g \quad \text{and} \quad \pi E_i f \succ g.$$

[9] Those axioms, together with axioms A1-A8, imply other Savage axioms.

Theorem 1. *Suppose $\succsim$ is a variational preference. Then Axioms P2, P4, and P6, are necessary and sufficient for $\succsim$ to have a multiplier representation (10). Moreover, two triples (θ', u', q') and (θ'', u'', q'') represent the same multiplier preference $\succsim$ if and only if q' and q'' are identical and there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $u' = \alpha u'' + \beta$ and $\theta' = \alpha\theta''$.*

The two cases: $\theta = \infty$ (lack of concern for model misspecification) and $\theta < \infty$ (concern for model misspecification) can be distinguished on the basis of the Independence Axiom (Axiom A2').[10] In the case when θ is finite, its numerical value is uniquely determined, given u. A positive affine transformation of u changes the scale on which θ operates, so θ has to change accordingly. This is reminiscent of the necessary adjustments of the CARA coefficient when units of account are changed.

Alternative axiomatizations are presented in Appendix A.2.9. It is shown there that Axiom A7 can be dispensed with in the presence of another of Savage's axioms—P3. Also, Savage's axiom P6 can be be dispensed with if Axiom A8 is strengthened to Arrow's (1970) Monotone-Continuity axiom and an additional axiom of nonatomicity is assumed.

4.5 Discussion

Any Anscombe-Aumann act can be viewed as a Savage act where prizes have an internal structure: they are lotteries. Because of this, an Anscombe-Aumann setting with the set of prizes Z can be viewed as a Savage setting with the set of prizes $\Delta(Z)$. Compared to a Savage setting with the set of prizes Z, more choice-observations are available in the Anscombe-Aumann setting. This additional information makes it possible to distinguish EU preferences from multiplier preferences.

To understand this distinction, observe that by 1, multiplier preferences have the following representation.

$$V(f) = \int_S \phi_\theta\left(\sum_{z\in Z} u(z) f_s(z)\right) \mathrm{d}q(s), \tag{12}$$

[10]The weaker Certainty Independence Axiom (Axiom A2") is also sufficient for making such a distinction. Alternatively, Machina and Schmeidler's (1995) axiom of Horse/Roulette Replacement could be used.

Because of the introduction of objective lotteries, this equation does not reduce to (8). The existence of two sources of uncertainty enables a distinction between *purely objective lotteries*, i.e., acts which pay the same lottery $\pi \in \Delta(Z)$ irrespectively of the state of the world and *purely subjective acts*, i.e., acts that in each state of the world pay off δ_z for some $z \in Z$.

From representation (12) it follows that for any two purely objective lotteries $\pi' \succsim \pi$ if and only if

$$\sum_{z\in Z} u(z)\pi'(z) \succsim \sum_{z\in Z} u(z)\pi(z).$$

On the other hand, each purely subjective act f induces a lottery $\pi_f(z) = q(f^{-1}(z))$. However, for any two such acts $f' \succsim f$ if and only if

$$\sum_{z\in Z} \phi_\theta(u(z))\pi_{f'}(z) \succsim \sum_{z\in Z} \phi_\theta(u(z))\pi_f(z).$$

What is crucial here is that the decision maker has a different attitude towards objective lotteries and subjective acts. In particular, if $\theta < \infty$ he is more averse towards subjective uncertainty than objective risk. The coexistence of those two sources in one model permits a joint measurement of those two attitudes.

It has been observed in the past that differences in attitudes towards risk and uncertainty lead to Ellsberg-type behavior. Neilson (1993) showed that the following *Second-Order Expected Utility* representation

$$V(f) = \int_S \phi\left(\sum_{z\in Z} u(z)f_s(z)\right) \mathrm{d}q(s), \tag{13}$$

can be obtained by a combination of von Neumann-Morgenstern axioms on lotteries and Savage axioms on acts.[11] A similar model was studied by Ergin and Gul (2004), see Section 5. From this perspective, multiplier preferences are a special case of (13) where $\phi = \phi_\theta$. Theorem 1 shows that this specific functional form of the function ϕ is implied by Weak Certainty Independence (Axiom A2) and by Uncertainty Aversion (Axiom A5).[12] Thus, the class of multiplier preferences is the intersection of the class

[11] I am grateful to Peter Klibanoff for this reference.

[12] This stems from the fact that, as elucidated by Grant and Polak (2007), variational preferences display constant absolute ambiguity aversion,

of variational preferences and the class of second-order expected utility preferences. The following example shows that, because of this property, multiplier preferences can be used for modelling Ellsberg-type behavior.

Example 3 (Ellsberg's Paradox revisited)**.** Suppose Urn I contains 100 red and black balls in unknown proportion, while Urn II contains 50 red and 50 black balls. Let the state space $S = \{R, B\}$ represent the possible draws from Urn I. Betting \$100 on red from Urn I corresponds to an act $f_R = (\delta_{100}, \delta_0)$ while betting \$100 on black from Urn I corresponds to an act $f_B = (\delta_0, \delta_{100})$. On the other hand, betting \$100 on red from Urn II corresponds to a lottery $\pi_R = \frac{1}{2}\delta_{100} + \frac{1}{2}\delta_0$, while betting \$100 on black from Urn II corresponds to a lottery $\pi_B = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_{100}$. These correspondences reflect the fact that betting on Urn I involves subjective uncertainty, while betting on Urn II involves objective risks. Note in particular, that $\pi_R = \pi_B$.

Consider the two multiplier preferences from Example 2: $\succsim_1$ with $u_1(x) = -\exp(-x)$ and $\theta_1 = \infty$, and $\succsim_2$ with $u_2(x) = x$ and $\theta_2 = 1$. Suppose also, that they both share the probability assessment $q(B) = q(R) = \frac{1}{2}$.

As explained in Example 2, the representation of $\succsim_1$ suggests that the decision maker is not concerned about model misspecification or ambiguity. Indeed, his choices reveal that $\pi_B \sim \pi_R \sim f_R \sim f_B$. This decision maker is indifferent between objective risk and subjective uncertainty, avoiding the Ellsberg paradox.

In contrast, the representation of $\succsim_2$ suggests that the decision maker is concerned about model misspecification or ambiguity. And indeed, his choices reveal that $\pi_B \sim \pi_R \succ f_R \sim f_B$. This decision maker prefers objective risk to probabilistically equivalent subjective uncertainty, displaying behavior typical in Ellsberg's experiments.

This means that introducing objective uncertainty makes it possible to disentangle risk aversion from concern about model misspecification and thus escape the consequences of 1. As a consequence, the interpretations of representations of $\succsim_1$ and $\succsim_2$ become behaviorally meaningful. ▲

It is worthwile to notice that for $\theta < \infty$ the decision maker behaves according to EU on the subdomain of objective lotteries and also on the subdomain of purely subjective acts. What leads to Ellsberg-type behavior are violations of EU across those domains: the decision maker's aversion towards objective risk (captured by u) is lower than his aversion towards objective risk (captured by $\phi_\theta \circ u$). This phenomenon is called *Second Order Risk Aversion*.[13]

4.6 Probabilistically Sophisticated Variational Preferences

Probabilistic sophistication, introduced by Machina and Schmeidler (1992), means that the decision maker treats subjective uncertainty in the same way as objective risk. In order to do so, the decision maker formulates a subjective measure on the state space. To evaluate an act, he first computes the distribution that the act induces on prizes. Second, he uses some criterion evaluating objective lotteries over prizes. This criterion may be expected utility, but it can also be one of many nonexpected utility criteria, which allow for modeling choices consistent with the Allais (1953) paradox.

In a setting where objective risk is explicitly present, such as in the Anscombe-Aumann setup, the requirement that the decision maker treats subjective uncertainty in the same way as objective risk imposes a uniform risk attitude towards both sources. Such uniformity is a critical requirement of the definition of probabilistic sophistication in the Anscombe-Aumann setting formulated by Machina and Schmeidler (1995).[14] In the class of variational preferences, this uniformity requirement implies that preferences are Anscombe-Aumann expected utility, because variational preferences use expected utility to evaluate objective lotteries.

A less restrictive notion of *second-order probabilistic sophistication* requires that the decision maker's preferences satisfy probabilistic sophistication on the subdomain of purely subjective acts (in accordance with Machina and Schmeidler's (1992) definition) and also on the subdomain of lotteries (by construction), but allow those two criteria to differ.[15] Preferences do not have to be probabilistically sophisticated overall because the decision maker's attitude towards those two sources of uncertainty is not required to be uniform. In particular, the decision maker could use a nonexpected utility criterion for evaluating subjective acts, while using an expected utility criterion for evaluating objective lotteries. Another possibility is when expected util-

[13] This notion was introduced by Ergin and Gul (2004) in a setting with two subjective sources of uncertainty (see Section 5).

[14] See also a recent analysis of Grant and Polak (2006).

[15] This notion was introduced by Ergin and Gul (2004), see Section 5. A related, but different, notion was discussed by Halevy and Ozdenoren (2007).

ity is used for both, but risk aversion depends on the source of uncertainty. The latter case is exemplified by multiplier preferences (12) and, more generally, second order expected utility (13).

For the class of variational preferences, Maccheroni et al. (2006a) obtain a general characterization of second-order probabilistic sophistication. A corollary of their result is that multiplier preferences are second-order probabilistically sophisticated. As representation (12) reveals, they are in fact second-order expected utility, which means that the criterion used by the decision maker to evaluate subjective acts is expected utility. This raises the question of whether there exist variational preferences that are second-order probabilistically sophisticated, but are not second-order EU. In other words: Can the Allais (1953) paradox be modeled using variational preferences? Marinacci (2002) showed that for the subclass of Maxmin Expected Utility preferences the answer is negative under a weak assumption of agreement of probabilities. Theorem 2 below extends Marinacci's (2002) result to the whole class of variational preferences under an appropriately extended notion of agreement of probabilities.

Assumption 1. For any $r \in [0, \infty)$ there exists an event $A_r \in \Sigma$ such that if $c(p) = c(p') = r$, then $0 < p(A_r) = p'(A_r) < 1$.

Assumption 1 requires that all measures with the same cost agree on some event. This assumption is equivalent to Marinacci's (2002) assumption for the subclass of Maxmin Expected Utility preferences.[16]

Theorem 2. *Suppose that $\succsim$ satisfies Axioms A1-A8. If Assumption 1 holds, then the following two statements are equivalent*

(i) $\succsim$ *is Second-Order Probabilistically Sophisticated*

(ii) $\succsim$ *is an Anscombe-Aumann Expected Utility preference.*

[16]In principle, another (stronger) generalization of Assumption 3 could be considered:
Assumption 2. There exists an event $A \in \Sigma$ such that if $c(p), c(p') < \infty$, then $0 < p(A) = p'(A) < 1$.
This assumption means that the state space contains outcomes of an objective randomizing device, for example of a coin. All measures that the decision maker considers plausible attach the same probability to events generated by such a device. This is a stronger requirement than Assumption 1 and it may be harder to verify for a given cost function. However, as Theorem 2 shows, the weaker Assumption 1 is sufficient.

Theorem 2 extends the result of Marinacci (2002) to the class of variational preferences.[17] The proof of Theorem 2 relies on different techniques than Marinacci's (2002) proof; it builds on the characterization of second-order probabilistically sophisticated variational preferences obtained by Maccheroni et al. (2006a).

4.7 Second-Order Variational Preferences

Multiplier preferences are an example of variational preferences having two representations:

$$V_1(f) = \min_{p \in \Delta(S)} \int_S u(f)\,\mathrm{d}p + \theta R(p\|q) \tag{10}$$

and

$$V_2(f) = \int_S \phi_\theta\big(u(f)\big)\,\mathrm{d}q. \tag{12}$$

One interpretation of this dichotomy is that model uncertainty in (10) manifests itself as second order risk aversion in (12). This motivates the following definition.

Definition 3. Preference relation $\succsim$ is a *Second-Order Variational Preference* if $\succsim$ is a variational preference with representation

$$V_1(f) = \min_{p \in \Delta S} \int_S u(f)\,\mathrm{d}p + c_1(p)$$

[17]Strictly speaking, Theorem 2 is not a generalization, because Marinacci's (2002) result holds also for α-MEU preferences which are not ambiguity averse (they violate Axiom A5). Moreover, he uses a weaker notion of probabilistic sophistication, that of probabilistic beliefs, and his results for maxmin expected utility preferences do not rely on countable additivity. Additionally, his theorem is presented in a Savage setting and makes an assumption about range-convexity of the utility function. One way to escape the consequences of Marinacci's (2002) theorem is to relax that assumption. However, Theorem 2 can be formulated only in the Anscombe-Aumann setting, because in the Savage setting the variational representation of preferences is not unique (as exemplified by 1 and Theorem 3 in the next section). Thus, because of the lottery structure, range convexity is implicit in the Anscombe-Aumann setting and the consequences of the Theorem cannot be avoided in the aforementioned way.

and it also has representation

$$V_2(f) = \min_{p \in \Delta S} \int_S \phi_\theta\big(u(f)\big)\,\mathrm{d}p + c_2(p)$$

for $\theta \in (0, \infty)$ and some grounded, convex, and lower semicontinous cost function c_2.

The following theorem characterizes this class of variational preferences.

Theorem 3. *Suppose that S is a Polish space and that $\succsim$ satisfies A1-A8. Preference $\succsim$ is a second-order variational preference if and only if $c_1(p) = \min_{q \in Q} \theta R(p\| q)$ for some closed and convex set of measures $Q \subseteq \Delta^\sigma(S)$. In this case c_2 can be chosen to satisfy $c_2 = \delta_Q$, i.e., $V_2(f) = \min_{p \in Q} \int_S \phi_\theta\big(u(f_s)\big)\,\mathrm{d}p$.*[18]

The analysis of probabilistic sophistication of Section 4.6 can be extended to second-order variational preferences. In order to do so, the following weak agreement assumption will be used.

Assumption 3. There exists an event $A_0 \in \Sigma$ such that if $c(p) = c(p') = 0$, then $0 < p(A_0) = p'(A_0) < 1$.

This is the agreement assumption used by Marinacci (2002). It means that there exists an event A_0, such that any two measures with zero cost agree on A_0.

Theorem 4. *Suppose that $\succsim$ is a Second-Order Variational Preference. If Assumption 3 holds, then the following two statements are equivalent*

(i) $\succsim$ is Second-Order Probabilistically Sophisticated

(ii) $\succsim$ is a Second-Order Expected Utility preference.

As a corollary of Theorem 2 another characterization of multiplier preferences is obtained.

Corollary 2. *Suppose that $\succsim$ satisfies Axioms A1-A8 and Assumption 3 holds. Then $\succsim$ is a multiplier preference if and only if $\succsim$ is a Second-Order Variational Preference and it is Second-Order Probabilistically Sophisticated.*

[18]The function c_2 in representation V_2 may not be unique. Uniqueness is guaranteed if the function u is unbounded from below.

5 Axiomatization within Ergin-Gul's model

This section discusses another enrichment of the domain of choice, which does not rely on the assumption of objective risk. Instead, it is assumed that there are two sources of subjective uncertainty, towards which the decision maker may have different attitudes. This type of environment was discussed by Chew and Sagi (2007), Ergin and Gul (2004), and Nau (2001, 2006); for an empirical application see Abdellaoui, Baillon, and Wakker (2007).

5.1 Subjective Sources of Uncertainty

Assume that the state space has a product structure $S = S_a \times S_b$, where a and b are two separate *issues*, or sources of uncertainty, towards which the decision maker may have different attitudes. In comparison with the Anscombe-Aumann framework, where objective risk is one of the sources, here both sources are subjective. Let $\mathcal{A}_a$ be a sigma algebra of subsets of S_a and $\mathcal{A}_b$ be a sigma algebra of subsets of S_b. Let Σ_a be the sigma algebra of sets of the form $A \times S_b$ for all $A \in \mathcal{A}_a$, Σ_b be the sigma algebra of sets of the form $S_a \times B$ for all $B \in \mathcal{A}_b$, and Σ be the sigma algebra generated by $\Sigma_a \cup \Sigma_b$. As before, $\mathcal{F}(Z)$ is the set of all simple acts $f : S \to Z$. In order to facilitate the presentation, it will be assumed that certainty equivalents exist, i.e., for any $f \in \mathcal{F}(Z)$ there exists $z \in Z$ with $z \sim f$. The full analysis without this assumption is contained in Appendices A.6 and A.7.

Ergin and Gul (2004) axiomatized preferences which are general enough to accommodate probabilistic sophistication and even second-order probabilistic sophistication. An important subclass of those preferences are second-order expected utility preferences.

$$V(f) = \int_{S_b} \phi\left(\int_{S_a} u(f(s_a, s_b))\, \mathrm{d}q_a(s_a) \right) \mathrm{d}q_b(s_b) \tag{14}$$

where $u : Z \to \mathbb{R}$, $\phi : \mathbb{R} \to \mathbb{R}$ is a strictly increasing and continuous function, and the measures $q_a \in \Delta(S_a)$ and $q_b \in \Delta(S_b)$ are nonatomic.

To characterize preferences represented by (14), Ergin and Gul (2004) assume Axioms A1, A6, and P3, together with weakenings of P2 and P4 and a strengthening of P6. There is a close relationship between (14) and Neilson's (1993) representation (13). The role of objective risk is now taken by a subjective source: issue a. For each

s_b, the decision maker computes the expected utility of $f(\cdot, s_b)$ and then averages those values using function ϕ.

5.2 Second-Order Risk Aversion

In the Anscombe-Aumann framework, concavity of the function ϕ is responsible for second-order risk aversion, i.e., higher aversion towards subjective uncertainty than towards objective risk. This property is a consequence of the axiom of Uncertainty Aversion (Axiom A5).[19] Similarly, in the present setup, concavity of function ϕ is responsible for higher aversion towards issue b than towards issue a. This property was introduced by Ergin and Gul (2004) who formally defined it in terms of mean-preserving spreads. However, this definition refers to the probability measures obtained from the representation and hence is not directly based on preferences. Theorems 2 and 5 of Ergin and Gul (2004) characterize second-order risk aversion in terms of induced preferences over induced Anscombe-Aumann acts and an analogue of Axiom A5 in that induced setting. However, just as with mean-preserving spreads, those induced Anscombe-Aumann acts are constructed using the subjective probability measure derived from the representation. As a consequence, the definition is not expressed directly in terms of observables.

In the presence of other axioms, the following purely behavioral axiom is equivalent to Ergin and Gul's (2004) definition.

Axiom A5' *(Second Order Risk Aversion).* For any $f, g \in \mathcal{F}_b$ and any $E \in \Sigma_a$ if $f \sim g$, then $fEg \succsim f$.

This axiom is a direct subjective analogue of Schmeidler's (1989) axiom of Uncertainty Aversion (Axiom A5).

Theorem 5. *Suppose $\succsim$ has representation (14). Then Axiom A5' is satisfied if and only if the function ϕ in (14) is concave.*

5.3 Axiomatization of Multiplier Preferences

The additional axiom that delivers multiplier preferences in this framework is Constant Absolute Second Order Risk Aversion.

[19] This follows from the proof of Theorem 1, see section A.2.6 in Appendix A.2.

Axiom A2''' *(Constant Absolute Second Order Risk Aversion).* There exists a non-null event $E \in \Sigma_a$ such that for all $f, g \in \mathcal{F}_b(Z)$, $x, y \in Z$

$$fEx \succsim gEx \Rightarrow fEy \succsim gEy.$$

In addition, two technical axioms, similar to Axioms 7 and 8, are needed.

Axiom A7' *($\mathcal{F}_a$-Unboundedness).* There exist $x \succ y$ in Z such that, for all non-null $E_a \in \Sigma_a$ there exist $z \in Z$ that satisfies either $y \succ zE_a x$ or $zE_a y \succ x$.

Axiom A8' *($\mathcal{F}_b$-Monotone Continuity).* If $f, g \in \mathcal{F}(Z)$, $x \in Z$, $\{E_n\}_{n \geq 1} \in \Sigma_b$ with $E_1 \supseteq E_2 \supseteq \cdots$ and $\bigcap_{n \geq 1} E_n = \emptyset$, then $f \succ g$ implies that there exists $n_0 \geq 1$ such that $xE_{n_0} f \succ g$.

Theorem 6. *Suppose $\succsim$ has representation (14). Then Axioms A2''', A5', A7, and A8 are necessary and sufficient for $\succsim$ to be represented by V, where*

$$V(f) = \min_{p_b \in \Delta S_b} \int_{S_b} \left(\int_{S_a} u(f(s_a, s_b))\, \mathrm{d}q_a(s_a) \right) \mathrm{d}p_b(s_b) + \theta R(p_b \| q_b)$$

and $u : Z \to \mathbb{R}$, $\theta \in (0, \infty]$, and q_a, q_b are nonatomic measures.

6 Dynamic Models

The main contributions of this section are an axiomatization of dynamic multiplier preferences and a characterization of preference for earlier resolution of uncertainty in the class of variational preferences. After the domain of choice is introduced in Section 6.1, dynamic variational preferences are defined in Section 6.2 and dynamic multiplier preferences are defined in Section 6.3. Section 6.4 presents an axiomatization of the latter class. Section 6.5 extends the notion of IID ambiguity to the class of variational preferences; this prepares the ground for studying the timing of uncertainty. Finally, Section 6.6 presents a characterization of preference for earlier resolution of uncertainty in the class of variational preferences.

6.1 Domain of Choice

Hayashi (2005), who studied a dynamic model of stationary maxmin expected utility preferences, used a domain of choice $\mathcal{H}$, which proves useful also for studying variational preferences. In each period the state space S is finite and the set of outcomes is a compact set Z.[20] The domain of *temporal Anscombe-Aumann acts*, $\mathcal{H}$, is constructed inductively

$$\mathcal{H}_0 = \mathcal{F}(\Delta(Z))$$

and

$$\mathcal{H}_t = \mathcal{F}(\Delta(Z \times \mathcal{H}_{t-1}))$$

for each $t \geq 1$.[21] Define $f \in \prod_{t=0}^{\infty} \mathcal{H}_t$ to be coherent if for any t the act f_{t+1} induces the same consumption process as f_t. As asserted by Theorem 1 of Hayashi (2005), the set $\mathcal{H}$ of such coherent acts satisfies the following homeomorphism

$$\mathcal{H} \simeq \mathcal{F}(\Delta(Z \times \mathcal{H})).$$

This recursive property facilitates axiomatizations of stationary preferences, because $\mathcal{H}$ is a mixture space under the usual state-by-state mixing of Anscombe-Aumann acts. An important subdomain of $\mathcal{H}$ is the space $\mathcal{D}$ of *temporal lotteries* of Kreps and Porteus (1978) and Epstein and Zin (1989)

$$\mathcal{D} \simeq \mathcal{F}(\Delta(Z \times \mathcal{D})).$$

Another important subdomain consists of *one-step-ahead acts* $\mathcal{H}_{+1}$ where all subjective uncertainty resolves in the first period.

$$\mathcal{H}_{+1} = \{h_{+1} \in \mathcal{F}(\Delta(Z \times \mathcal{H})) \mid h_{+1}(s) \in \mathcal{D} \text{ for all } s \in S\}.$$

[20] Finiteness of S and compactness of Z can be relaxed. This requires a different construction of space $\mathcal{H}$, which will be included in future versions of this paper.

[21] For any compact metric space X, the set of Borel probability measures $\Delta(X)$ is a compact metric space with the Prohorov metric and the set $\mathcal{F}(X) = X^S$ is a compact metric space under the product metric.

6.2 Dynamic Variational Preferences

Following Hayashi (2005), for each $t \geq 0$ and history $s^t = (s_1, \ldots, s_t) \in S^t$ the decision maker's preference $\succsim_{s^t}$ over $\Delta(Z \times \mathcal{H})$ is observed. For any $z \in Z$ and $h \in \mathcal{H}$ the degenerate lottery $\delta_{(z,h)}$ will, with a slight abuse of notation, be denoted (z, h).

Definition 4. Family $\{\succsim_{s^t}\}$ is a Dynamic Variational Preference if it is represented by a family of continuous, nonconstant functions $U_{s^t} : \Delta(Z \times \mathcal{H}) \to \mathbb{R}$ such that

$$U_{s^t}(\mu) = \int_{Z\times\mathcal{H}} \left\{ u(z) + \beta \left[\min_{p\in\Delta(S)} \int_S U_{(s^t,s)}\big(h(s)\big)\,\mathrm{d}p(s) + c_{s^t}(p) \right] \right\} \mathrm{d}\mu(z,h)$$

for any $p \in \Delta(Z\times\mathcal{H})$, where $u : Z \to \mathbb{R}$ is continuous and nonconstant and $\beta \in (0,1)$ and cost functions c_{s^t} are grounded, convex, and lower semicontinous. Moreover, β is unique and the function u is unique up to positive affine transformations.

An axiomatization of dynamic variational preferences can be obtained by modifying Hayashi's (2005) axiomatization of dynamic maxmin expected utility preferences, in particular by relaxing certainty independence to weak certainty independence. Appendices A.8 and A.9 present this axiomatization.

6.3 Dynamic Multiplier Preferences

Definition 5 (Dynamic Multiplier Preference)**.** Family $\{\succsim_{s^t}\}$ is a dynamic multiplier preference if it is represented by

$$U_{s^t}(\mu) = \int_{Z\times\mathcal{H}} \left\{ u(z) + \beta \left[\min_{p\in\Delta(S)} \int_S U_{(s^t,s)}\big(h(s)\big)\,\mathrm{d}p(s) + \theta R(p\|q_{s^t}) \right] \right\} \mathrm{d}\mu(z,h) \quad (15)$$

where $u : Z \to \mathbb{R}$ is continuous and nonconstant, $\beta \in (0,1)$, $\theta \in (0,\infty]$ and $q \in \Delta(S)$.

The reference probability q_{s^t} in representation (15) can be history dependent, which is natural in non-stationary environments or when learning takes place. However, the parameter θ is history-independent. Thus, a separation is achieved between the attitude towards model uncertainty, which is constant, and the uncertainty itself can depend on the history of shocks, reflecting possible persistence of shocks or learning about the environment.

6.4 Axiomatization of Dynamic Multiplier Preferences

In the static version of the model, Savage's axioms were used to characterize multiplier preferences. Because those axioms rely on infiniteness of the state space and in the present setting S is finite, a different approach will be used, that of Wakker's *tradeoff consistency* (see, e.g., Köbberling and Wakker, 2003).[22]

Relation $\sim^*_{s^t}$ introduced below compares tradeoffs between pairs of temporal lotteries. Pair $[d_1, d_2]$ is in relation with pair $[d_3, d_4]$ if the utility difference between d_1 and d_2 is the same as the utility difference between d_3 and d_4.

Definition 6. For any $d_1, d_2, d_3, d_4 \in \mathcal{D}$ define $[d_1, d_2] \sim^*_{s^t} [d_3, d_4]$ if there exist acts $h'_{+1}, h''_{+1} \in \mathcal{H}_{+1}$, and a s^t-nonnull state[23] $s \in S$ such that

$$(d_1)s(h'_{+1}) \sim_{s^t} (d_2)s(h''_{+1}) \text{ and } (d_3)s(h'_{+1}) \sim_{s^t} (d_4)s(h''_{+1}).$$

Axiom B1 *(Tradeoff Consistency).* For any $s^t \in S^t$ and $d_1, d_2, d_3, d_4 \in \mathcal{D}$ if $[d_1, d_2] \sim^*_{s^t} [d_3, d_4]$, then improving any of the outcomes breaks the relation.

Axiom B1 implies multiplier representation of preferences in each period

$$U_{s^t}(\mu) = \int_{Z\times\mathcal{H}} \left\{ u(z) + \beta \left[\min_{p\in\Delta S} \int_S U_{(s^t,s)}\big(h(s)\big)\,\mathrm{d}p(s) + \theta_{s^t} R(p\|\, q_{s^t}) \right] \right\} \mathrm{d}\mu(z,h),$$

but allows the concern for model misspecification to be time- and state- dependent. The following axiom guarantees constant θ.

Axiom B2 *(Stationary Tradeoff Consistency).* Relation $\sim^*_{s^t}$ is independent of s^t.

Theorem 7. *Suppose that $\{\succsim_{s^t}\}$ is a dynamic variational preference. Then Axioms B1 and B2 are necessary and sufficient for $\{\succsim_{s^t}\}$ to be a dynamic multiplier preference. Moreover, $\big(\theta, u, \{q_{s^t}\}\big)$ and $\big(\theta', u', \{q'_{s^t}\}\big)$ represent the same dynamic multiplier preference if and only if $q'_{s^t} = q_{s^t}$ for all s^t and there exists $a > 0$ and $b \in \mathbb{R}$ such that $u' = au + b$ and $\theta' = a\theta$.*

[22] Using a construction of the space $\mathcal{H}$ that accommodates infinite S (as described footnote 20) will allow to replace tradeoff consistency with Savage axioms in the future versions of this paper.

[23] A state s is s^t-non-null if there exist $h'_{+1}, h''_{+1}, g_{+1} \in \mathcal{H}_{+1}$ such that $(h'_{+1})s(g_{+1}) \succ_{s^t} (h''_{+1})s(g_{+1})$.

6.5 Stationary Variational Preferences and IID Ambiguity

The discussion so far has concentrated on variational preferences where the utility function u and the discount factor β are constant, but the cost function c_{s^t} is allowed to depend on the history s^t. For example, in the case of multiplier preferences, the reference measure q_{s^t} can be history-dependent. This section introduces a class of *stationary* preferences, where the preference on one-step-ahead acts is the same in every time period. This permits writing $\succsim$ instead of $\succsim_{s^t}$.

Definition 7. Relation $\succsim$ is a Stationary Variational Preference if it is represented by function $U : \mathcal{H} \to \mathbb{R}$

$$U(\mu) = \int_{Z\times\mathcal{H}} \left\{ u(z) + \beta \left[\min_{p\in\Delta S} \int_S U\big(h(s)\big)\, \mathrm{d}p(s) + c(p) \right] \right\} \mathrm{d}\mu(z,h) \tag{16}$$

for $\beta \in (0,1)$, $u : Z \to \mathbb{R}$, and some grounded, convex, and lower semicontinous cost function c.

This definition extends the notion of *IID Ambiguity* studied by Chen and Epstein (2002) and Epstein and Schneider (2003) in the context of maxmin expected utility to the class of variational preferences. Intuitively, IID ambiguity means that every period the decision maker faces a new Ellsberg urn. His ex-ante beliefs about each urn are identical, but because he observes only one draw from each urn, he cannot make inferences across urns.[24]

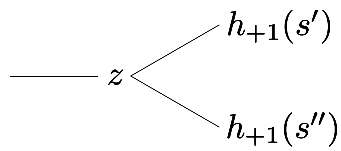
Because the uncertainty that the decision maker faces in period t is identical to the uncertainty in period $t+1$, and the only property that distinguishes them is the timing of their resolution, attitudes towards such timing of resolution can be studied.

6.6 Attitudes Towards the Timing of Subjective Uncertainty

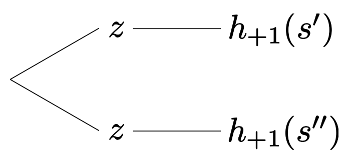
The main objective of this section is to determine which of the stationary variational preferences exhibit preference for earlier resolution of uncertainty. In order to do so, some notation will be introduced. Let $h_{+1} \in \mathcal{H}_{+1}$ be a one-step-ahead act and $z \in Z$ be a deterministic payoff. Define $(1; z, h_{+1})$ to be a temporal act where the

[24] This failure of inference is known in econometrics as the problem of incidental parameters (see, e.g., Neyman and Scott, 1948).

subjective uncertainty about h_{+1} is resolved in period 1, i.e., whose chance node for period 0 is degenerate. Formally, define $(1; z, h_{+1})(s) = (z, h_{+1})$ for all $s \in S$.



On the other hand, define $(0; z, h_{+1})$ to be a one-step-ahead act where the subjective uncertainty about h_{+1} is resolved already in period 0, i.e., whose chance node for period 0 is not degenerate. Formally, define $(0; z, h_{+1})(s) = (z, h_{+1}(s))$ for all $s \in S$.



Note that both in $(0; z, h_{+1})$ and in $(1; z, h_{+1})$ the payoffs of h_{+1} are delivered in period 1. The difference is *when* the decision maker learns about them. Some decision makers may prefer one to the other.

Definition 8. Relation $\succsim$ exhibits *preference for [resp., indifference to, preference against] earlier resolution of uncertainty* if

$$(0; z, h_{+1}) \succsim \text{ [resp., } \sim, \precsim] \, (1; z, h_{+1})$$

for all $z \in Z$, and $h_{+1} \in \mathcal{H}_{+1}$.

Given this definition, the preference for earlier resolution of uncertainty can be studied in the class of stationary variational preferences. One initial observation is that stationary multiplier preferences, which are represented by

$$U(\mu) = \int_{Z \times \mathcal{H}} \left[u(z) + \beta \phi_\theta^{-1} \left(\int_S \phi_\theta \big(U \big(h(s) \big) \big) \, \mathrm{d}q(s) \right) \right] \mathrm{d}\mu(z, h) \tag{17}$$

exhibit strict preference for earlier resolution of uncertainty (unless $\theta = \infty$).

Theorem 8. *Suppose $\succsim$ is a stationary multiplier preferencewith $\theta < \infty$. Then for all $z \in Z$, and $h_{+1} \in \mathcal{H}_{+1}$ such that $h_{+1}(s) \nsim h_{+1}(s')$ for some $s \neq s'$*

$$(0; z, h_{+1}) \succ (1; z, h_{+1}).$$

Similarly to stationary multiplier preferences, stationary second-order variational preferences, which are represented by

$$U(\mu) = \int_{Z \times \mathcal{H}} \left[u(z) + \beta \phi_\theta^{-1} \left(\min_{q \in Q} \int_S \phi_\theta \big(U\big(h(s)\big)\big) \, \mathrm{d}q(s) \right) \right] \mathrm{d}\mu(z, h) \qquad (18)$$

exhibit strict preference for earlier resolution of uncertainty (unless $\theta = \infty$).

Theorem 9. *Suppose $\succsim$ is a stationary second-order risk-averse variational preference with $\theta < \infty$. Then for all $z \in Z$, and $h_{+1} \in \mathcal{H}_{+1}$ such that $h_{+1}(s) \nsim h_{+1}(s')$ for some $s \neq s'$*

$$(0; z, h_{+1}) \succ (1; z, h_{+1}).$$

Both in Theorem 8 and in Theorem 9 the preference for earlier resolution of uncertainty appears to be connected to the function ϕ_θ. Indeed, the strength of the preference depends on the parameter θ; in the extreme case of $\theta = \infty$ the indifference obtains. By Theorem 3, the second-order risk-averse variational preferences are the largest subclass of variational preferences with representation

$$U(\mu) = \int_{Z \times \mathcal{H}} \left[u(z) + \beta \phi_\theta^{-1} \left(\min_{p \in \Delta(S)} \int_S \phi_\theta \big(U\big(h(s)\big)\big) \, \mathrm{d}p(s) + c(p) \right) \right] \mathrm{d}\mu(z, h)$$

For this reason, it may be tempting to conclude that all other variational preferences satisfy indifference to the timing of resolution of uncertainty. However, as the next theorem shows, quite the opposite is true.

Theorem 10. *Suppose that $\succsim$ is a stationary variational preference. Relation $\succsim$ satisfies indifference to the timing of resolution of uncertainty if and only if it is a stationary maxmin expected utility preference, i.e., it is represented by*

$$U(\mu) = \int_{Z \times \mathcal{H}} \left[u(z) + \beta \min_{q \in Q} \int_S U\big(h(s)\big) \, \mathrm{d}q(s) \right] \mathrm{d}\mu(z, h). \qquad (19)$$

Theorem 10 asserts that stationary variational preferences typically exhibit preference for earlier resolution of uncertainty. The only class that satisfies indifference is precisely the class of stationary maxmin expected utility preferences studied by Chen and Epstein (2002) and Epstein and Schneider (2003).

7 Conclusion

One of the challenges in decision theory lies in finding decision models that would do better than Expected Utility in describing individual choices, but would at the same time be easy to incorporate into economic models of aggregate behavior.

This paper studies the model of multiplier preferences which is known to satisfy the latter requirement. By obtaining an axiomatic characterization of this model, the paper studies its individual choice properties, which helps to determine whether it also satisfies the first requirement mentioned above. The axiomatization provides a set of testable implications of the model, which will be helpful in its empirical verification. The axiomatization also enables measurement of the parameters of the model on the basis of observable choice data alone, thereby providing a useful tool for applications of the model.

In addition, the paper generalizes some of the properties of multiplier preferences to the broader class of variational preferences and studies the extent to which this more general class of preferences can be used for modelling the Allais paradox.

In a dynamic setting, the paper obtains an axiomatization of recursive multiplier preferences. This class displays an interesting property of preference for earlier resolution of uncertainty. The paper shows that this property is generally shared by dynamic variational preferences; the only subclass that displays indifference to timing is that of recursive maxmin expected utility preferences.

A Appendix: Proofs

Let $B_0(\Sigma)$ denote the set of all real-valued Σ-measurable simple functions and let $B_0(\Sigma, K)$ be the set of all functions in $B_0(\Sigma)$ that take values in a convex set $K \subseteq \mathbb{R}$.

A.1 Proof of Observation 1

Because $\theta^{-1} \cdot (u \circ f)$ is a bounded measurable function on (S, Σ), from Proposition 1.4.2 of Dupuis and Ellis (1997) it follows that

$$\min_{p \in \Delta S} \int_S (u \circ f) \,\mathrm{d}p + \theta R(p \| q) = -\theta \log \left(\int_S \exp \left(-\frac{u \circ f}{\theta} \right) \mathrm{d}q \right).$$

Thus, $\succsim$ is a multiplier preference with θ, u, and q iff it is represented by U with

$$U(f) = -\theta \log \left(\int_S \exp \left(-\frac{u \circ f}{\theta} \right) \mathrm{d}q \right).$$

Rewrite using the definition of ϕ_θ:

$$U(f) = \phi_\theta^{-1} \left(\int_S (\phi_\theta \circ u \circ f) \,\mathrm{d}q \right).$$

Since ϕ_θ is a monotone transformation, $\succsim$ is also represented by $V := \phi_\theta \circ U$, i.e.,

$$V(f) = \int_S (\phi_\theta \circ u \circ f) \,\mathrm{d}q.$$

A.2 Proof of Theorem 1

A.2.1 Niveloidal Representation

By Lemmas 25 and 28 of Maccheroni et al. (2006a), Axioms A1-A7 imply that there exists an unbounded affine function $u : \Delta(Z) \to \mathbb{R}$ and a normalized concave niveloid $I : B_0(\Sigma, u(\Delta(Z))) \to \mathbb{R}$ such that for all $f \succsim g$ iff $I(u \circ f) \geq I(u \circ g)$. Moreover, within this class, u is unique up to positive affine transformations. Define $\mathcal{U} := u(\Delta(Z))$. After normalization, there are three possible cases: $\mathcal{U} \in \{\mathbb{R}_+, \mathbb{R}_-, \mathbb{R}\}$.

A.2.2 Utility Acts

For each act f, define the *utility act* associated with f as $u \circ f \in B_0(\Sigma, \mathcal{U})$. The preference on acts induces a preference on utility acts: for any $\xi', \xi'' \in B_0(\Sigma, \mathcal{U})$ define $\xi' \succsim_u \xi''$ iff $f' \succsim f''$, for some $\xi' = u \circ f'$ and $\xi'' = u \circ f''$. The choice of particular versions of f' and f'' is irrelevant, because $\xi' \succsim_u \xi''$ iff $I(\xi') \geq I(\xi'')$.

By Lemma 22 in Maccheroni, Marinacci, and Rustichini (2004), for all $k \in \mathcal{U}$ and $\xi \in B_0(\Sigma,\mathcal{U})$ we have $I(\xi + k) = I(\xi) + k$. Thus, $\xi' \succsim_u \xi''$ iff $I(\xi') \geq I(\xi'')$ iff $I(\xi' + k) \geq I(\xi'' + k)$ iff $\xi' + k \succsim_u \xi'' + k$ for all $k \in \mathcal{U}$ and $\xi', \xi'' \in B_0(\Sigma,\mathcal{U})$.

A.2.3 Savage's P3

In order to show that $\succsim$ have an additive representation (12), Savage's theorem will be used in A.2.4. To do this, it is necessary to show that his P3 axiom holds.

Definition 9. An event $E \in \Sigma$ is *non-null* if there exist $f, g, h \in \mathcal{F}$ such that $fEh \succ gEh$.

Axiom P3 *(Savage's Eventwise Monotonicity).* For all $x, y \in Z$, $h \in \mathcal{F}$, and non-null $E \in \Sigma$

$$x \succsim y \Leftrightarrow xEh \succsim yEh.$$

Lemma 1. *Axioms A1-A7, together with Axiom P2 imply axiom P3.*

Proof. First, suppose that $x \succsim y$. It follows from Axiom A4 (Monotonicity) that $xEh \succsim yEh$ for any $h \in \mathcal{F}$ and any E. Second, suppose that $y \succ x$. It follows from Monotonicity that $yEh \succsim xEh$ for any $h \in \mathcal{F}$ and any E. Towards contradiction, suppose that $yEh \sim xEh$ for a non-null $E \in \Sigma$ and some $h \in \mathcal{F}$.

Because E is non-null, there exist $f, g \in \mathcal{F}$ such that $fEh \succ gEh$. Let $\{E_1, \ldots, E_n, E\}$ be a partition of S with respect to which both fEh and gEh are measurable. Let y' be the most preferred element among $\{f(E_i) \mid i = 1, \ldots, n\}$ and let x' be the least preferred element among $\{g(E_i) \mid i = 1, \ldots, n\}$. By Monotonicity, $y'Eh \succsim fEh$ and $gEh \succsim x'Eh$. Thus $y'Eh \succ x'Eh$.

Observe that there exist $a, a' \in \mathcal{U}$ and $k, k' > 0$, such that $a = u(x), a + k = u(y), a' = u(x')$ and $a' + k' = u(y')$. Thus there exists $\xi \in B_0(\Sigma,\mathcal{U})$, such that $aE\xi = u \circ (xEh)$, $(a + k)E\xi = u \circ (yEh)$, $a'E\xi = u \circ (x'Eh)$, and $(a' + k')E\xi = u \circ (y'Eh)$. It follows that

$$I((a + k)E\xi) = I(aE\xi) \tag{20}$$

$$I((a' + k')E\xi) > I(a'E\xi). \tag{21}$$

Suppose that $\mathcal{U} = \mathbb{R}_+$. By translation invariance, it follows from (20) that $I((a+2k)E(\xi+k)) = I((a+k)E(\xi+k))$ and by P2, that $I((a+2k)E\xi) = I((a+k)E\xi)$. Hence, $I((a+2k)E\xi) = I(aE\xi)$. By induction $I((a+nk)E\xi) = I(aE\xi)$ for all $n \in \mathbb{N}$, and by Monotonicity $I((a+r)E\xi) = I(aE\xi)$ for all $r \in \mathbb{R}_+$. In particular, letting $r = k'$, we have

$$I((a+k')E\xi) = I(aE\xi). \tag{22}$$

Suppose that $a' \geq a$. By translation invariance, $I((a'+k')E(\xi+a'-a) = I(a'E(\xi+a'-a))$ and by P2, $I((a'+k')E\xi) = I(a'E\xi)$. Contradiction with (22). Thus, it must be that $a > a'$. By translation invariance, it follows from (21), that $I((a+k')E(\xi+a-a')) > I(aE(\xi+a-a'))$ and by P2, $I((a+k')E\xi) > I(aE\xi)$. Contradiction with (22). The proof is analogous in case when $\mathcal{U} = \mathbb{R}_-$ or $\mathcal{U} = \mathbb{R}$. □

A.2.4 Application of Savage's Theorem

It follows from Chapters 1-5 of Savage (1972) that there exists a (not necessarily affine) function $\psi : \Delta(Z) \to \mathbb{R}$ and a measure $q \in \Delta S$, such that for any $f, g \in \mathcal{F}$, $f \succsim g$ iff $\int_S (\psi \circ f)\, \mathrm{d}q \geq \int_S (\psi \circ g)\, \mathrm{d}q$. Moreover, ψ is unique up to positive affine transformations. From Theorem 1 in Section 1 of Villegas (1964) it follows that Axiom A8 implies that $q \in \Delta^\sigma(S)$.

A.2.5 Proof of representation (13)

By A.2.2, $f \succsim g$ iff $\int_S (\psi \circ f)\, \mathrm{d}q \geq \int_S (\psi \circ g)\, \mathrm{d}q$. In particular, $x \succsim y$ iff $\psi(x) \geq \psi(y)$. From axioms A1-A6 it follows that $x \succsim y$ iff $u(x) \geq u(y)$. Thus, there exists a unique strictly increasing function $\phi : \mathbb{R} \to \mathbb{R}$ such that $\psi = \phi \circ u$. Thus, $f \succsim g$ iff $\int_S (\phi \circ u \circ f)\, \mathrm{d}q \geq \int_S (\phi \circ u \circ g)\, \mathrm{d}q$. This leads to the following representation of $\succsim_u$: $\xi' \succsim_u \xi''$ iff $\int_S (\phi \circ \xi')\, \mathrm{d}q \geq \int_S (\phi \circ \xi'')\, \mathrm{d}q$.

A.2.6 Concavity of ϕ

Let $a, b \in \mathcal{U}$. Let $\pi, \rho \in \Delta(Z)$ be such that $a = u(\pi)$ and $b = u(\rho)$. Because q is range convex, there exists a set E with $q(E) = \frac{1}{2}$. Let $f = \pi E \rho$ and $g = \rho E \pi$ and observe that $V(f) = \frac{1}{2}\phi(a) + \frac{1}{2}\phi(b) = V(g)$; thus, $f \sim g$. By Axiom A5, $\frac{1}{2}f + \frac{1}{2}g \succsim f$, i.e., $\phi\big(\frac{1}{2}a + \frac{1}{2}b\big) = V\big(\frac{1}{2}f + \frac{1}{2}g\big) \geq V(f) = \frac{1}{2}\phi(a) + \frac{1}{2}\phi(b)$. Thus,

$$\phi\left(\frac{1}{2}a + \frac{1}{2}b\right) \geq \frac{1}{2}\phi(a) + \frac{1}{2}\phi(b). \tag{23}$$

Let $\alpha \in (0,1)$. Let the sequence $\{\alpha_n\}$ be a dyadic approximation of α. By induction, inequality (23) implies that $\phi\big(\alpha_n a + (1-\alpha_n)b\big) \geq \alpha_n\phi(a) + (1-\alpha_n)\phi(b)$ for all n. By continuity of ϕ, $\lim_{n\to\infty} \phi\big(\alpha_n a + (1-\alpha_n)b\big) = \phi\big(\alpha a + (1-\alpha)b\big)$. Thus, $\phi\big(\alpha a + (1-\alpha)b\big) \geq \alpha\phi(a) + (1-\alpha)\phi(b)$.

A.2.7 Proof that $\phi = \phi_\theta$

By defining $\phi^k(x) := \phi(x+k)$ for all $k, x \in \mathcal{U}$, it follows from A.2.2 and A.2.5 that $\int_S \phi^k \circ \xi' \,\mathrm{d}q \geq \int_S \phi^k \circ \xi'' \,\mathrm{d}q$ iff $\int_S \phi \circ \xi' \,\mathrm{d}q \geq \int_S \phi \circ \xi'' \,\mathrm{d}q$. Thus, (ϕ, q) and (ϕ^k, q) are EU representations of the same preference on $B_0(\Sigma, \mathcal{U})$. By uniqueness, $\phi(x+k) = \alpha(k)\phi(x) + \beta(k)$ for all $k, x \in \mathcal{U}$. This is a generalization of Pexider's equation (see equation (3) of Section 3.1.3, p. 148 of Aczél, 1966). If $\mathcal{U} \in \{\mathbb{R}, \mathbb{R}_+\}$, then by Corollary 1 in Section 3.1.3 of Aczél (1966), up to positive affine transformations, the only strictly increasing concave solutions are of the form ϕ_θ, for $\theta \in (0, \infty]$. It is easy to prove that the same is true for $\mathcal{U} = \mathbb{R}_-$.

A.2.8 Conclusion of the Proof

Combining Steps 4 and 5, $f \succsim g$ iff $\int_S (\phi_\theta \circ u \circ f)\,\mathrm{d}q \geq \int_S (\phi_\theta \circ u \circ g)\,\mathrm{d}q$. Because $q \in \Delta^\sigma$, by 1, it follows that $f \succsim g$ iff $\min_{p \in \Delta S} \int_S (u \circ f)\,\mathrm{d}p + \theta R(p \,\|\, q) \geq \min_{p \in \Delta S} \int_S (u \circ g)\,\mathrm{d}p + \theta R(p \,\|\, q)$. □

A.2.9 Alternative Axiomatizations

Removing P6

Instead of Axiom P6, the following two axioms could be assumed:

Axiom A8" *(Arrow's Monotone Continuity).* If $f, g \in \mathcal{F}$, $x \in Z$, $\{E_n\}_{n\geq 1} \in \Sigma$ with $E_1 \supseteq E_2 \supseteq \cdots$ and $\bigcap_{n\geq 1} E_n = \emptyset$, then $f \succ g$ implies that there exists $n_0 \geq 1$ such that $xE_{n_0}f \succ g$ and $f \succ xE_{n_0}g$.

Axiom A9 *(Nonatomicity).* Every nonnull event can be partitioned into two nonnull events.

Axiom A8" is stronger than Axiom A8 and is necessary to obtain a countably additive probability. Axiom A9 (see Villegas, 1964) is needed to obtain fineness and tightness of the qualitative probability.

This leads to the following theorem: Axioms A1-A7, A8", together with P2, P4, and A9 are necessary and sufficient for $\succsim$ to have a multiplier representation. The proof is analogous, but instead of Savage's Theorem, as in A.2.4, Arrow's (1970) theorem is used (cf. Chapter 2 of his book).

Removing Unboundedness

Instead of Axiom A7, Savage's axiom P3 could be assumed. as verified by Klibanoff et al. (2005) in the proof of their Proposition 2, the family of functions ϕ_θ remains to be the only solution of Pexider's functional equation when domain is restricted to an interval.

Savage Axioms Only on Purely Objective Acts

If the existence of certainty equivalents for lotteries is assumed, i.e., for any $\pi \in \Delta(Z)$ there exists $z \in Z$ with $z \sim \pi$, then the Savage axioms can be weakened in the following sense. In Theorem 1 Axioms P2, P4, and P6 were assumed to hold on all (Anscombe-Aumann) acts. Assuming the existence of certainty equivalents makes it possible to impose Savage axioms only on Savage acts, i.e., acts paying out a degenerate lottery in each state.

A.3 Proof of Theorem 2

Let $q \in \Delta^\sigma(S, \Sigma)$ and let $L^1(S, \Sigma, q)$ denote the set of all nonnegative measurable functions on (S, Σ) with $\int_S f \, \mathrm{d}q = 1$. For $f, g \in L^1(S, \Sigma, q)$ define $f \sim_{cx} g$ iff

$$q(s \in S \mid f(s) \leq t) = q(s \in S \mid g(s) \leq t)$$

for any $t \geq 0$. Similarly, for any measures $p, p' \in \Delta^\sigma(S, \Sigma)$ define $p \sim_{cx} p'$ iff $\frac{\mathrm{d}p}{\mathrm{d}q} \sim_{cx} \frac{\mathrm{d}p'}{\mathrm{d}q}$. For $p \in \Delta^\sigma(S, \Sigma)$, the set $O(p) = \{p' \in \Delta^\sigma(S, \Sigma) \mid p' \sim_{cx} p\}$ is called the *orbit* of p. A set of measures $\Gamma \subseteq \Delta^\sigma(q)$ is called *orbit-closed* iff $p \in \Gamma \Rightarrow O(p) \subseteq \Gamma$.

Lemma 2. *Let $f \in L^1(S, \Sigma, q)$ and let $F, G \in \Sigma$ be disjoint events, with $q(F) = q(G)$. Then, there exists $g \in L^1(S, \Sigma, q)$ such that $f = g$ on $(F \cup G)^c$, $\int_F f \, \mathrm{d}q = \int_G g \, \mathrm{d}q$, and $f \sim_{cx} g$.*

Proof. For each $n \in \mathbb{N}$ and for $1 \leq k \leq n2^n$ define sets

$$
{}_nF_0 = \{s \in F \mid f(s) \geq n\}, \; {}_nF_k = \left\{ s \in F \,\middle|\, \frac{k-1}{2^n} \leq f(s) \leq \frac{k}{2^n} \right\},
$$
$$
{}_nG_0 = \{s \in G \mid f(s) \geq n\}, \; {}_nG_k = \left\{ s \in G \,\middle|\, \frac{k-1}{2^n} \leq f(s) \leq \frac{k}{2^n} \right\}.
$$

Because q is nonatomic, it is also convex-ranged (see, e.g., Villegas, 1964). Thus, for each n, partitions $\{{}_nF'_k\}_{k=0}^{n2^n}$ of F and $\{{}_nG'_k\}_{k=0}^{n2^n}$ of G can be constructed such that

$$
q(F'_{n,k}) = q(G_{n,k}) \text{ and } q(G'_{n,k}) = q(F_{n,k})
$$

for all $0 \leq k \leq n2^n$ and

$$
{}_{(n+1)}G'_{(2k)} \subseteq {}_{(n+1)}G'_{(k)} \text{ and } {}_{(n+1)}G'_{(2k+1)} \subseteq {}_{(n+1)}G'_{(k)}
$$

for all $0 \leq k \leq n2^n$ and $n \in \mathbb{N}$.

Define functions

$$
f_n = \sum_{k=1}^{n2^n} \left(\frac{k-1}{2^n} \mathbf{1}_{{}_nF_k} \right) + n\mathbf{1}_{{}_nF_k} + f_{|(E \cup G)^c} + \sum_{k=1}^{n2^n} \left(\frac{k-1}{2^n} \mathbf{1}_{{}_nG_k} \right) + n\mathbf{1}_{{}_nG_k},
$$
$$
g_n = \sum_{k=1}^{n2^n} \left(\frac{k-1}{2^n} \mathbf{1}_{{}_nF'_k} \right) + n\mathbf{1}_{{}_nF'_k} + f_{|(E \cup G)^c} + \sum_{k=1}^{n2^n} \left(\frac{k-1}{2^n} \mathbf{1}_{{}_nG'_k} \right) + n\mathbf{1}_{{}_nG'_k}.
$$

Observe, that functions f_n satisfy $0 \leq f_n \leq f_{n+1}$, and converge pointwise to f. Similarly, functions g_n satisfy $0 \leq g_n \leq g_{n+1}$. Define $g = \lim_{n \to \infty} g_n$. Observe that $f = g$ on $(E \cup G)^c$. Moreover, $\int_S f_n \, \mathrm{d}q = \int_S g_n \, \mathrm{d}q$, so by the Monotone Convergence Theorem $\int_S f \, \mathrm{d}q = \int_S g \, \mathrm{d}q$.

To see that $f \sim_{cx} g$, let $t \geq 0$ and define sets

$$A_n = \{s \in S \mid f_n(s) \leq t\},\ A = \{s \in S \mid f(s) \leq t\},$$
$$B_n = \{s \in S \mid g_n(s) \leq t\},\ B = \{s \in S \mid g(s) \leq t\},$$

Verify, that by construction of f_n and g_n $A_n \downarrow A$, $B_n \downarrow B$, and $q(A_n) = q(B_n)$ for all n. By countable additivity of q, $\lim_{n\to\infty} q(A_n) = q(A)$ and $\lim_{n\to\infty} q(B_n) = q(B)$. □

Lemma 3. *Suppose that $\Gamma \subseteq \Delta^\sigma(q)$ is an orbit-closed set of measures. Suppose also that there exists $A \in \Sigma$ such that $0 < p(A) = p'(A) < 1$ for all $p, p' \in \Gamma$. Then $\Gamma = \{q\}$.*

Proof. Let $\alpha = q(A)$. Observe, that wlog $\alpha \leq \frac{1}{2}$, because if all measures in Γ agree on A, then they also agree on A^c. Also, if $\alpha = 0$, then for any $p \in \Gamma$ $q(A) = 0 \Rightarrow p(A) = 0$, contradicting the assumption. Thus, $\alpha \in (0, \frac{1}{2}]$.
Step 1: *$p(E) = p(A)$ for all $p \in \Gamma$ and for all events $E \in \Sigma$ with $q(E) = \alpha$.*

Let $E \in \Sigma$ be such that $q(E) = \alpha$ and observe that $q(A - E) = q(E - A)$. Let $p \in \Gamma$ and define $f = \frac{\mathrm{d}p}{\mathrm{d}q}$. By Lemma 2 applied to $(E - A)$ and $(A - E)$, there exists $g \in L^1(S, \Sigma, q)$ such that $f = g$ on $(A \cup E)^c \cup (A \cap E)$, $\int_{(E-A)} f \,\mathrm{d}q = \int_{(A-E)} g \,\mathrm{d}q$, and $f \sim_{cx} g$. Define measure $p' \in \Delta^\sigma(S, \Sigma)$ by $p'(F) = \int_F g \,\mathrm{d}q$ and observe that $p' \sim_{cx} p$. Moreover, $p(E - A) = p'(A - E)$ and $p(A \cap E) = p'(A \cap E)$. Thus, $p(E) = p(E - A) + p(A \cap E) = p'(A - E) + p'(A \cap E) = p'(A) = p(A)$, where the last equality holds by orbit-closedness of Γ.
Step 2: *$p(F) = p(F')$ for all $p \in \Gamma$ and for all disjoint events $F, F' \in \Sigma$ with $q(F) = q(F') = \beta < \alpha$.*

Observe that $\beta < \frac{1}{2}$, so $\alpha - \beta < 1 - 2\beta$. Thus, by range-convexity of q, there exists $H \subseteq (F \cup F')^c$ with $q(H) = \alpha - \beta$. By Step 1 applied to sets $F \cup H$ and $F' \cup H$, it follows that $p(F) + p(H) = p(F \cup H) = p(A) = p(F' \cup H) = p(F') + p(H)$; hence, $p(F) = p(F')$.
Step 3: *$p(G) = q(G)$ for all $p \in \Gamma$ and for $G \in \Sigma$..*

Let $\gamma = q(G)$ and for each $n \in \mathbb{N}$ define $k_n = \sup\{k \mid \frac{k}{n} \leq \gamma\}$. Observe, that $\lim_{n\to\infty} \frac{k_n}{n} = \gamma$. For each $n \in \mathbb{N}$, by range-convexity of q, there exists a partition $\{F_1, \ldots, F_n\}$ of F such that $q(F_k) = \frac{1}{n}$ for $k = 1, \ldots, n$, sets $F_1, \ldots, F_{k_n} \subseteq G$, and sets $F_{k_n+2}, \ldots, F_n \subseteq G^c$. By Step 2, $p(F_k) = \frac{1}{n}$ for $k = 1, \ldots n$, so $\frac{k_n}{n} \leq p(G) \leq \frac{k_n+1}{n}$. By letting n to infinity, $p(G) = \gamma$. □

Proof of Theorem 2. The direction $(ii) \Rightarrow (i)$ is trivial. For $(i) \Rightarrow (ii)$, observe that for any $r \in \mathbb{R}_+$ let $C_r = \{p \in \Delta(S, \Sigma) \mid c(p) = r\}$ denote the level set of the cost function c. Observe that

$$V(f) = \min_{p \in \Delta(S,\Sigma)} \int_S (u \circ f)\, \mathrm{d}p + c(p) = \min_{r \in \mathbb{R}_+} \min_{p \in C_r} \int_S (u \circ f)\, \mathrm{d}p + r$$

From the proof of Corollary 4 in Sarin and Wakker (2000) it follows that Axiom A8 implies that $\succsim$ is probabilistically sophisticated with respect to some $q \in \Delta^\sigma(S)$. By Theorem 14 of Maccheroni et al. (2006a), if $\succsim$ is probabilistically sophisticated with respect to $q \in \Delta^\sigma(S)$, then c is rearrangement invariant, i.e., $p \sim_{cx} p' \Rightarrow c(p) = c(p')$ for all $p, p' \in \Delta(S, \Sigma)$. Thus, each C_r is orbit-closed. Therefore, by Assumption 1 and Lemma 3, $C_r = \{q\}$ for all $r \in \mathbb{R}_+$. Thus,

$$V(f) = \min_{r \in \mathbb{R}_+} \int_S (u \circ f)\, \mathrm{d}q + r = \int_S (u \circ f)\, \mathrm{d}q. \qquad \square$$

A.4 Proof of Theorem 3

Lemma 4 establishes that $c_1(p) = \min_{q \in Q} \theta R(p \,\|\, q)$ is a legitimate cost function. Lemma 5 is the main step in proving necessity. The rest of the proof deals with sufficiency.

Lemma 4. *Suppose S is a Polish space. For any convex closed set $Q \subseteq \Delta^\sigma(S)$ the function $c_1(p) = \min_{q \in Q} \theta R(p \| q)$ is nonnegative, convex, lower semicontinuous, and $\{p \in \Delta(S) \mid c_1(p) \leq r\} \subseteq \Delta^\sigma(S)$ for each $r \geq 0$. Moreover, the function c_1 is grounded and $\{p \in \Delta(S) \mid c_1(p) = 0\} = Q$.*

Proof. Nonnegativity follows from $R(p\|q)$ being nonnegative for any $p, q \in \Delta(S)$.

By Lemma 1.4.3 (b) in Dupuis and Ellis (1997), $R(\cdot \| \cdot)$ is a convex, lower semicontinuous function on $\Delta^\sigma(S) \times \Delta^\sigma(S)$. Thus, $\arg\min_{q \in Q} \theta R(p \,\|\, q)$ is a nonempty compact and convex set for any $p \in \Delta^\sigma(S)$. Let $\lambda \in (0, 1)$ and $p', p'' \in \Delta^\sigma(S)$. Let $q' \in \arg\min_{q \in Q} \theta R(p' \,\|\, q)$ and $q'' \in \arg\min_{q \in Q} \theta R(p'' \,\|\, q)$. Convexity follows from:

$$\begin{aligned} c_1(\lambda p' + (1-\lambda)p'') &= \min_{q\in Q} \theta R(\lambda p' + (1-\lambda)p'' \,\|\, q) \\ &\le \theta R\big(\lambda p' + (1-\lambda)p'' \,\|\, \lambda q' + (1-\lambda)q''\big) \\ &\le \lambda\theta R(p' \,\|\, q') + (1-\lambda)\theta R(p'' \,\|\, q'') \\ &= \lambda c_1(p') + (1-\lambda)c_1(p''). \end{aligned}$$

For lower semicontiuniuty define $\mathrm{Proj} : \Delta^\sigma(S)\times Q\times\mathbb{R} \to \Delta^\sigma(S)\times\mathbb{R}$ to be a projection $\mathrm{Proj}(p,q,r) = (p,r)$. Let $\mathrm{Epi}(R) = \{(p,q,r) \in \Delta^\sigma(S)\times Q\times\mathbb{R} \mid R(p\|\,q) \le r\}$ be the epigraph of R and $\mathrm{Epi}(c_1) = \{(p,r) \in \Delta^\sigma(S)\times\mathbb{R} \mid c_1(p) \le r\}$ be the epigraph of c_1. Observe that, by lower semicontinuity of R, the set $\mathrm{Epi}(R)$ is closed. Next, observe that $\mathrm{Epi}(c_1) = \mathrm{Proj}\big(\mathrm{Epi}(R)\big)$.

To verify that, let $(p,r) \in \mathrm{Epi}(c_1)$. Then $c_1(p) \le r$; thus $\min_{q\in Q} R(p \,\|\, q) \le r$. Let $q' \in \arg\min_{q\in Q} R(p \,\|\, q)$. It follows, that $R(p \,\|\, q') \le r$; thus, $(p,q,r) \in \mathrm{Epi}(R)$. Conclude that $(p,r) \in \mathrm{Proj}\big(\mathrm{Epi}(R)\big)$. Conversely, let $(p,r) \in \mathrm{Proj}\big(\mathrm{Epi}(R)\big)$. Then there exists q' such that $(p,q',r) \in \mathrm{Epi}(R)$, so that $R(p \,\|\, q') \le r$. Thus, $c_1(p) = \min_{q\in Q} R(p\|\,q) \le R(p\|\,q') \le r$. Conclude that $(p,r) \in \mathrm{Epi}(c_1)$.

Finally, observe that $\mathrm{Proj}(C)$ is closed for any closed set $C \in \Delta^\sigma(S)\times Q\times\mathbb{R}$. Let (p_n,r_n) be a sequence in $\mathrm{Proj}(C)$ with limit (p,r). Because $(p_n,r_n) \in \mathrm{Proj}(C)$, there exists a sequence q_n in Q such that $(p_n,q_n,r_n) \in C$. Because Q is a compact set subset of a metric space, $\lim_{n\to\infty} q_n = q \in Q$ by passing to a subsequence. By closedness of C, it follows that $\lim_{n\to\infty}(p_n,q_n,r_n) = (p,q,r) \in C$. Thus, $(p,r) \in C$.

To see that $\{p \in \Delta(S) \mid c_1(p) \le r\} \subseteq \Delta^\sigma(S)$ for each $r \ge 0$, observe that $\{p \in \Delta(S) \mid R(p \,\|\, q) \le r\} \subseteq \Delta^\sigma(S)$ and that by compactness of Q and lower-semicontinuity of $R(p\|\,\cdot)$

$$\{p \in \Delta(S) \mid c_1(p) \le r\} = \bigcup_{q\in Q} \{p \in \Delta(S) \mid R(p\|\,q) \le r\}.$$

For groundedness, recall that by Lemma 1.4.1 in Dupuis and Ellis (1997) $R(p\| q) = 0$ iff $p = q$. Thus, $c_1(q) \le R(q\,\|\,q) = 0$ for any $q \in Q$. Conversely, if $c_1(p) = 0$, then $\min_{q\in Q} R(p \,\|\, q) = 0$. By lower semincontinuity of R, there exists $q \in Q$ such that $0 = c_1(p) = R(p\|\,q)$. Thus, by Lemma 1.4.1 in Dupuis and Ellis (1997), $p = q$; hence, $p \in Q$. □

Lemma 5. *Suppose $\succsim$ is a variational preference and $Q \subseteq \Delta^\sigma(S)$ is a closed and convex set. Then V_1 with $c_1(p) = \min_{q\in Q} \theta R(p\| q)$ represents $\succsim$ if and only if V_2 with $c_2 = \delta_Q$ represents $\succsim$.*

Proof. Observe that

$$\begin{aligned} V_1(f) &= \min_{p\in\Delta S} \int_S u(f_s)\,\mathrm{d}p + \min_{q\in Q} \theta R(p\| q) \\ &= \min_{p\in\Delta S} \min_{q\in Q} \int_S u(f_s)\,\mathrm{d}p + \theta R(p\| q) \\ &= \min_{q\in Q} \min_{p\in\Delta S} \int_S u(f_s)\,\mathrm{d}p + \theta R(p\| q) \\ &= \min_{q\in Q} \phi_\theta^{-1}\left(\int_S \phi_\theta\big(u(f_s)\big)\,\mathrm{d}q \right) \\ &= \phi_\theta^{-1}\left(\min_{q\in Q} \int_S \phi_\theta\big(u(f_s)\big)\,\mathrm{d}q \right), \end{aligned}$$

where the fourth inequality follows from Proposition 1.4.2 in Dupuis and Ellis (1997) and the fifth from strict monotonicity of ϕ_θ^{-1}. Thus, V_1 is ordinally equivalent to $V_2(f) = \min_{q\in Q} \int_S \phi_\theta\big(u(f_s)\big)\,\mathrm{d}q = V_2(f) = \min_{p\in\Delta S} \int_S \phi_\theta\big(u(f_s)\big)\,\mathrm{d}p + c_2(p)$. □

Proof of Theorem 3. Suppose that V_1 with $c_1(p) = \min_{q\in Q} \theta R(p \,\|\, q)$ represents $\succsim$. By Lemma 4 an by Theorems 3 and 13 of Maccheroni et al. (2006a), $V_1(f) = \min_{p\in\Delta S} \int_S u(f_s)\,\mathrm{d}p + c_1(p)$ is a representation of a preference $\succsim$ that satisfies axioms A1-A8. By Lemma 5, V_2 with $c_2 = \delta_Q$ represents $\succsim$.

Conversely, suppose that $\succsim$ is a variational preference represented by $V_2(f) = \min_{p\in\Delta S} \int_S \phi_\theta\big(u(f_s)\big)\,\mathrm{d}p + c_2(p)$. Define niveloid $I : B_0(\Sigma, \phi_\theta(\mathcal{U})) \to \mathbb{R}$ by $I(\xi) = \min_{p\in\Delta S} \int_S \xi\,\mathrm{d}p + c_2(p)$ and observe that $V_2(f) = I\big(\phi_\theta(u(f))\big)$. Therefore,

$$\begin{aligned} V_2(\alpha f + (1-\alpha)\pi) &= I\bigg(\phi_\theta\big(\alpha u(f) + (1-\alpha)u(\pi)\big) \bigg) \\ &= I\bigg(-\phi_\theta\big((1-\alpha)u(\pi)\big) \cdot \phi_\theta\big(\alpha u(f_s)\big) \bigg) \end{aligned} \tag{24}$$

for any $f \in \mathcal{F}(\Delta(Z))$, $\pi \in \Delta(Z)$, and $\alpha \in (0, 1)$.

Niveloid I is homogeneous of degree one. To verify, suppose that $\mathcal{U} = u(\Delta(Z)) = \mathbb{R}_+$. (The case of $\mathcal{U} \in \{\mathbb{R}_-, \mathbb{R}\}$ is analogous.) Let $\xi \in \mathcal{B}_0(\Sigma, \phi_\theta(\mathbb{R}_+))$ and $b \in (0, 1]$ (the case $b \geq 1$ follows from this). Let scalar $r = b^{-1}I(b\xi)$; observe that $I(br) = I(I(b\xi)) = I(b\xi)$. Let $f \in \mathcal{F}(\Delta(Z))$ be such that $\phi_\theta(\frac{1}{2}u(f)) = \xi$ and $\pi \in \Delta(Z)$

be such that $\phi_\theta(\frac{1}{2}u(\pi)) = r$. Their existence is guaranteed by unboundedness of $\mathcal{U}$. Furthermore, let $\rho, \rho' \in \Delta(Z)$ be such that $b = -\phi_\theta(\frac{1}{2}u(\rho))$ and $u(\rho') = 0$. (In the case of $\mathcal{U} = \mathbb{R}_-$, prove homogeneity for $b \geq 1$ and deduce for $b \in (0, 1]$.) By (24), $I(b\xi) = I(br)$ this implies $V_2\big(\phi_\theta(\frac{1}{2}u(f) + \frac{1}{2}u(\rho))\big) = V_2\big(\phi_\theta(\frac{1}{2}u(\pi) + \frac{1}{2}u(\rho))\big)$. Because $\succsim$ satisfies Axiom A2, this implies $V_2\big(\phi_\theta(\frac{1}{2}u(f) + \frac{1}{2}u(\rho'))\big) = V_2\big(\phi_\theta(\frac{1}{2}u(\pi) + \frac{1}{2}u(\rho'))\big)$, which, by (24), implies $I(\xi) = I(r)$. Thus, $I(b\xi) = I(br) = bI(r) = bI(\xi)$.

If $\mathcal{U} = \mathbb{R}_+$ or $\mathcal{U} = \mathbb{R}_-$, then I is defined on $B_0(\Sigma, [-1, 0))$ or $B_0(\Sigma, (-\infty, -1])$, respectively. Extend I to $B_0(\Sigma, \mathbb{R}_-)$ by homogeneity. Note that I is monotone, homogeneous of degree one, and vertically invariant on $B_0(\Sigma, \mathbb{R}_-)$. If $\mathcal{U} = \mathbb{R}$, then I is already defined on $B_0(\Sigma, \mathbb{R}_-)$ and enjoys those properties.

By Lemma 23 of Maccheroni et al. (2004), I is niveloid on $B_0(\Sigma, \mathbb{R}_-)$. By Lemmas 21 and 22 of Maccheroni et al. (2004), the unique vertically invariant extension of I to $\mathcal{B}_0(\Sigma)$, defined by $\tilde{I}(\xi + k) = I(\xi) + k$ for any $\xi + k \in B_0(\Sigma, \mathbb{R})$ such that $\xi \in B_0(\Sigma, \mathbb{R}_-)$ is monotonic. Note that $\tilde{I}$ is monotone homogeneous of degree one on $B_0(\Sigma, \mathbb{R})$.

Therefore, $\tilde{I}$ satisfies the assumptions of Lemma 3.5 of Gilboa and Schmeidler (1989). Thus, there exists a closed, convex set $Q \subseteq \Delta(S)$ such that $\tilde{I}(\xi) = \min_{p \in Q} \int \xi \, \mathrm{d}p$. Hence, $I(\xi) = \min_{p \in Q} \int \xi \, \mathrm{d}p$ for all $\xi \in B_0(\Sigma, \phi_\theta(\mathcal{U}))$.

Let E_n be a vanishing sequence of events and let $x < y$ be elements of $\phi_\theta(\mathcal{U})$. Observe that by Axiom A8, for any k there exists a N such that $I(xE_n y) > I(y - \frac{1}{k})$ for all $n \geq N$. Thus, $\min_{p \in Q} \int xE_n y \, \mathrm{d}p > y - \frac{1}{k}$. Therefore, $(x - y) \max_{p \in Q} p(E_n) > \frac{1}{k}$. Hence, $p(E_n) < (k(y - x))^{-1}$ for any $p \in Q$. Therefore $\lim_{n \to \infty} p(E_n) = 0$ for any $p \in Q$. Thus, $Q \subseteq \Delta^\sigma(S)$.

Finally, by Lemma 5, $c_1(p) = \min_{q \in Q} \theta R(p \| q)$. □

A.5 Proof of Theorem 4

The direction $(ii) \Rightarrow (i)$ is trivial. For $(i) \Rightarrow (ii)$, observe that by Theorem 3, $\succsim$ can be represented by

$$V_1(f) = \min_{p \in \Delta(S, \Sigma)} \int_S (u \circ f) \, \mathrm{d}p + c_1(p)$$

with $c_1(p) = \min_{q \in Q} R(p \| q)$ for some closed and convex set $Q \subseteq \Delta^\sigma(S)$. From the proof of Corollary 4 in Sarin and Wakker (2000) it follows that Axiom A8 implies that $\succsim$ is probabilistically sophisticated with respect to some $q \in \Delta^\sigma(S)$. By Theorem 14 of Maccheroni et al. (2006a), if $\succsim$ is probabilistically sophisticated with respect to $q \in \Delta^{\sigma}(S)$, then c_1 is rearrangement invariant, i.e., $p \sim_{cx} p' \Rightarrow c_1(p) = c_1(p')$ for all $p, p' \in \Delta(S, \Sigma)$. Thus, in particular, the set $\{p \in \Delta(S) \mid c_1(p) = 0\}$ is orbit-closed. Therefore, by Assumption 1 and Lemma 3, $\{p \in \Delta(S) \mid c_1(p) = 0\} = \{q\}$. But, by Theorem 3, $\succsim$ can be represented by

$$V_2(f) = \min_{p \in Q} \int_S \phi_\theta(u \circ f) \, \mathrm{d}p.$$

Moreover, by Lemma 4, $Q = \{p \in \Delta(S) \mid c_1(p) = 0\}$. Conclude that $\succsim$ can be represented by

$$V_2(f) = \int_S \phi_\theta(u \circ f) \, \mathrm{d}q.$$

A.6 Proof of Theorem 5

In order to relax the assumption of existence of certainty equivalents, the following definition will be used.

Definition 10. Act $f \in \mathcal{F}_a(Z)$ is *symmetric with respect to* $E \in \Sigma_a$ if for all $z \in Z$

$$fEz \sim zEf.$$

Symmetric acts have the same expected utility on each "half" of the state space.[25]

Axiom A5" *(Second Order Risk Aversion).* If acts $f, g \in \mathcal{F}_a$ are symmetric with respect to $E \in \Sigma_a$, then for all $F \in \Sigma_b$

$$fFg \sim gFf \Rightarrow (fFg)E(gFf) \succsim fFg.$$

The proof of Theorem 5 follows from the proof of the following stronger theorem

Theorem 11. *Suppose $\succsim$ has representation (14). Then Axiom A5" is satisfied if and only if the function ϕ in (14) is concave.*

[25] Symmetric acts are acts that can be "subjectively mixed". Such subjective mixtures are different from subjective mixtures studied by Ghirardato, Maccheroni, Marinacci, and Siniscalchi (2003), whose construction relies on range-convexity of u. In the present setting, subjective mixtures are not needed under range-convexity of u.

Proof.

A.6.1 Necessity

Suppose $f \in \mathcal{F}_a(Z)$ is symmetric with respect to $E \in \Sigma_a$. Let $\alpha = q_a(E)$. Axiom A6 and representation (14) imply that there exist $z', z'' \in Z$ with $z' \succ z''$. Thus, $fEz' \sim z'Ef$ and $fEz'' \sim z''Ef$ imply that

$$\int_E (u \circ f)\, \mathrm{d}q_a + (1-\alpha)u(z') = \alpha u(z') + \int_{E^c} (u \circ f)\, \mathrm{d}q_a, \tag{25}$$

$$\int_E (u \circ f)\, \mathrm{d}q_a + (1-\alpha)u(z'') = \alpha u(z'') + \int_{E^c} (u \circ f)\, \mathrm{d}q_a. \tag{26}$$

By subtracting (26) from (25)

$$(1-\alpha)[u(z') - u(z'')] = \alpha[u(z') - u(z'')];$$

thus, $\alpha = \frac{1}{2}$ and therefore

$$\int_E (u \circ f)\, \mathrm{d}q_a = \int_{E^c} (u \circ f)\, \mathrm{d}q_a.$$

Let $f, g \in \mathcal{F}_a(Z)$. Denote $U(f) = \int_{S_a} (u \circ f)\, \mathrm{d}q_a$ and $U(g) = \int_{S_a} (u \circ g)\, \mathrm{d}q_a$. Because f and g are symmetric with respect to $E \in \Sigma_a$,

$$\int_E (u \circ f)\, \mathrm{d}q_a = \int_{E^c} (u \circ f)\, \mathrm{d}q_a = \frac{1}{2}U(f)$$

$$\int_E (u \circ g)\, \mathrm{d}q_a = \int_{E^c} (u \circ g)\, \mathrm{d}q_a = \frac{1}{2}U(g).$$

Let $F \in \Sigma_b$ and $\beta = q_b(F)$. If $fFg \sim gFf$, then

$$\beta\phi\big(U(f)\big) + (1-\beta)\phi\big(U(g)\big) = \beta\phi\big(U(g)\big) + (1-\beta)\phi\big(U(f)\big).$$

Thus,

$$(2\beta - 1)\phi\big(U(f)\big) = (2\beta - 1)\phi\big(U(g)\big).$$

If $\beta \neq \frac{1}{2}$, then $U(f) = U(g)$ and trivially

$$V\big((fFg)E(gFf)\big) = \beta\phi\left(\frac{1}{2}U(f) + \frac{1}{2}U(g)\right) + (1-\beta)\phi\left(\frac{1}{2}U(g) + \frac{1}{2}U(f)\right)$$
$$= \beta\phi\big(U(f)\big) + (1-\beta)\phi\big(U(g)\big) = V(fFg).$$

If $\beta = \frac{1}{2}$, then

$$V\big((fFg)E(gFf)\big) = \frac{1}{2}\phi\left(\frac{1}{2}U(f) + \frac{1}{2}U(g)\right) + \frac{1}{2}\phi\left(\frac{1}{2}U(g) + \frac{1}{2}U(f)\right)$$
$$= \phi\left(\frac{1}{2}U(f) + \frac{1}{2}U(g)\right) \geq \frac{1}{2}\phi\big(U(f)\big) + \frac{1}{2}\phi\big(U(g)\big) = V(fFg),$$

where the inequality follows from concavity of ϕ.

A.6.2 Sufficiency

Convexity of Domain of ϕ

Let D_ϕ be the domain of function ϕ, i.e., $D_\phi = \{U(f) \mid f \in \mathcal{F}_a\}$. Suppose $k, l \in D_\phi$ and $\alpha \in (0,1)$. Wlog $k < l$. Let $f, g \in \mathcal{F}_a$ be such that $k = U(f)$ and $l = U(g)$. Define $A = \min_{s\in S} f(s)$ and $B = \max_{s\in S} g(s)$ and let $x, y \in Z$ be such that $u(x) = A$ and $u(y) = B$. By nonatomicity of q_a, there exists $E \in \Sigma_a$ with $q_a(E) = \big(B - [\alpha k + (1-\alpha)l]\big)(B-A)^{-1}$. Verify, that $U(xEy) = \alpha k + (1-\alpha)l$. Hence, D_ϕ is a convex set.

Dyadic Convexity of ϕ

Suppose $k, l \in D_\phi$ and let $f, g \in \mathcal{F}_a$ be such that $k = U(f)$ and $l = U(g)$. Define $\underline{k} = \min_{s\in S} f(s)$, $\bar{k} = \max_{s\in S} f(s)$, $\underline{l} = \min_{s\in S} g(s)$, and $\bar{l} = \max_{s\in S} g(s)$. Let $\underline{x}, \bar{x}, \underline{y}, \bar{y}$ be such that $u(\underline{x}) = \underline{k}$, $u(\bar{x}) = \bar{k}$, $u(\underline{y}) = \underline{l}$, $u(\bar{y}) = \bar{l}$. Also, define $\kappa = \frac{\bar{k}-k}{\bar{k}-\underline{k}}$ and $\lambda = \frac{\bar{l}-l}{\bar{l}-\underline{l}}$. By nonatomicity of q_a there exist partitions $\{E_1^\kappa, E_2^\kappa, E_3^\kappa, E_4^\kappa\}$ and $\{E_1^\lambda, E_2^\lambda, E_3^\lambda, E_4^\lambda\}$ of S_a such that $E_1^\kappa \cup E_2^\kappa = E_1^\lambda \cup E_2^\lambda$, $q_a(E_1^\kappa \cup E_2^\kappa) = q_a(E_1^\lambda \cup E_2^\lambda) = \frac{1}{2}$, $q_a(E_1^\kappa \cup E_3^\kappa) = \frac{\kappa}{2}$, and $q_a(E_1^\lambda \cup E_3^\lambda) = \frac{\lambda}{2}$.

Define acts $f = \underline{x}E_1^\kappa\,\bar{x}E_2^\kappa\,\underline{x}E_3^\kappa\,\bar{x}E_4^\kappa$ and $g = \underline{y}E_1^\lambda\,\bar{y}E_2^\lambda\,\underline{y}E_3^\lambda\,\bar{y}E_4^\lambda$. Verify that f and g are symmetric with respect to $E = E_1^\kappa \cup E_2^\kappa = E_1^\lambda \cup E_2^\lambda$ and satisfy $U(f) = k$ and $U(g) = l$. By nonatomicity of q_b, there exists $F \in \Sigma_b$ with $q_b(F) = \frac{1}{2}$. Verify that

$V(fFg) = \frac{1}{2}\phi(k) + \frac{1}{2}\phi(l) = V(gFf)$. Hence, by Axiom A5',

$$\phi\left(\frac{1}{2}k + \frac{1}{2}l\right) = \frac{1}{2}\phi\left(\frac{1}{2}k + \frac{1}{2}l\right) + \frac{1}{2}\phi\left(\frac{1}{2}l + \frac{1}{2}k\right) = V\big((fFg)E(gFf)\big)$$
$$\geq V(fFg) = \frac{1}{2}\phi(k) + \frac{1}{2}\phi(l).$$

As a consequence,

$$\phi\left(\frac{1}{2}k + \frac{1}{2}l\right) \geq \frac{1}{2}\phi(k) + \frac{1}{2}\phi(l) \tag{27}$$

for all $k, l \in D_\phi$.

Limiting argument

Let $\alpha \in [0,1]$. From A.6.2 it follows that $\alpha k + (1-\alpha)l \in D_\phi$. Let the sequence $\{\alpha_n\}$ be a dyadic approximation of α. By induction, inequality (27) implies that $\phi(\alpha_n k + (1-\alpha_n)l) \geq \alpha_n\phi(k) + (1-\alpha_n)\phi(l)$ for all n. By continuity of ϕ, $\lim_{n\to\infty}\phi(\alpha_n k + (1-\alpha_n)l) = \phi(\alpha k + (1-\alpha)l)$. Thus, $\phi(\alpha k + (1-\alpha)l) \geq \alpha\phi(k) + (1-\alpha)\phi(l)$. □

A.7 Proof of Theorem 6

By Theorem 3 of Ergin and Gul (2004), Axioms A1, A6, P2', P3, P4', and P6' guarantee the existence of nonatomic measures $q_a \in \Delta S_a$ and $q_b \in \Delta S_b$, function $u : Z \to \mathbb{R}$, and a continuous and strictly increasing $\phi : \mathbb{R} \to \mathbb{R}$ such that $\succsim$ is represented by V with

$$V(f) = \int_{S_b} \phi\left(\int_{S_a} u(f(s_a, s_b))\, \mathrm{d}q_a(s_a)\right) \mathrm{d}q_b(s_b). \tag{28}$$

Let x, y be as in Axiom A7'. Wlog $u(y) = 0$, thus $u(x) > 0$. Nonatomicity of q_a guarantees that there exists a sequence of events $\{E_n\}_{n\geq 1}$ in Σ_a with $q_a(E_n) = \frac{1}{n}$. Axiom A7' guarantees that there exist a sequence $\{z'_n\}_{n\geq 1}$ with $\phi\big(0\big) > \phi\big(\frac{1}{n}u(z'_n) + \frac{n-1}{n}u(x)\big)$ or a sequence $\{z''_n\}_{n\geq 1}$ with $\phi\big(\frac{1}{n}u(z''_n)\big) > \phi\big(u(x)\big)$ (or both such sequences exist). By strict monotonicity of ϕ if follows that, in the first case, $-(n-1)u(x) > u(z'_n)$; thus $u(z'_n) \to -\infty$; hence, u is unbounded from below. In the second case, $u(z''_n) > nu(x)$; thus, $u(z''_n) \to +\infty$; hence, in this case u

is unbounded from above. Define $\mathcal{U} := u(Z)$. After normalization, there are three possible cases: $\mathcal{U} \in \{\mathbb{R}+, \mathbb{R}_-, \mathbb{R}\}$.

Let $E \in \Sigma_a$ be as in Axiom A2''' and let $p := q_a(E)$. For any $k \in \mathcal{U}$ define a preference $\succsim^k$ on $\mathcal{F}_b$ as follows. Let $z \in Z$ be such that $u(z) = k$ and for any $f, g \in \mathcal{F}_b(Z)$ define $f \succsim^k g$ iff $fEz \succsim gEz$. (Because of Axiom A2''', the choice of particular z does not matter.) Define $\phi^k(u) := \phi(u + (1-p)k)$. From representation (28), it follows that $\succsim^k$ is represented by V^k with

$$V^k(f) = \int_{S_b} \phi^k\left(\int_E u(f(s_a, s_b))\, \mathrm{d}q_a(s_a)\right) \mathrm{d}q_b(s_b).$$

By Axiom A2''', $\succsim^k = \succsim^0$ for all $k \in \mathcal{U}$. Hence, ϕ^k and ϕ^0 are equal up to positive affine transformations, i.e., $\phi(u + (1-p)k) = \alpha(k)\phi(u) + \beta(k)$ for all $u, k \in \mathcal{U}$. By changing variables: $k' := (1-p)k$, $\alpha'(k') = \alpha(\frac{k'}{p})$, and $\beta'(k') = \beta(\frac{k'}{p})$, it follows that $\phi(k'+u) = \alpha'(k')\phi(u) + \beta'(k')$ for all $u, k' \in \mathcal{U}$, which is is a generalization of Pexider's equation (see equation (3) of Section 3.1.3, p. 148 of Aczél, 1966). By Theorem 5, ϕ is concave. By Corollary 1 in Section 3.1.3 of Aczél (1966), up to positive affine transformations, the only strictly increasing quasiconcave solutions are of the form ϕ_θ, for $\theta \in (0, \infty]$.

It follows from Theorem 1 in Section 1 of Villegas (1964) that Axiom A8' delivers countable additivity of q_b. A reasoning similar to 1 of this paper concludes the proof. □

A.8 Axiomatization of Recursive Variational Preferences

Definition 11. Family $\{\succsim_{s^t}\}$ is a Recursive Variational Preference if it is represented by a family of continuous, nonconstant functions $U_{s^t} : \Delta(Z \times \mathcal{H}) \to \mathbb{R}$ such that

$$U_{s^t}(\mu) = \int_{Z \times \mathcal{H}} W\left(z,\ \min_{p \in \Delta S} \int_S U_{(s^t,s)}\big(h(s)\big)\, \mathrm{d}p(s) + c_{s^t}(p)\right) \mathrm{d}\mu(z, h) \qquad (29)$$

for any $p \in \Delta(Z \times \mathcal{H})$, where the aggregator $W : Z \times R_U \to R_U$ is continuous and strictly increasing in the second argument and cost functions c_{s^t} are grounded, convex, and lower semicontinous. Here, $R_U = \bigcup_{t \geq 1} \bigcup_{s^t \in S^t} \bigcup_{p \in \Delta(Z \times \mathcal{H})} U_{s^t}(p)$.

The axiomatization of recursive variational preferences combines Hayashi's (2005) axiomatization of recursive maxmin expected utility preferences with the axiomatization of Maccheroni et al. (2006a). All of Hayashi's (2005) axioms are retained, except that certainty independence is relaxed to weak certainty independence.

Axiom D1 *(Order).* For any $s^t \in S^t$ relation $\succsim_{s^t}$ is a continuous, complete, transitive, and there exist $y, y' \in Z^\infty$ such that $y \succ_{s^t} y'$.

Axiom D2 *(Consumption Separability).* For any $s^t \in S^t$, $z, z' \in Z$, and $h, h' \in \mathcal{H}$

$$(z, h) \succsim_{s^t} (z, h') \text{ if and only if } (z', h) \succsim_{s^t} (z', h').$$

Axiom D3 *(Risk Preference).* For any $s^t, \hat{s}^t \in S^t$, $z \in Z$, and $d, d' \in \mathcal{D}$

(i) (History-Independence)

$$d \succsim_{s^t} d' \text{ if and only if } d \succsim_{\hat{s}^t} d',$$

(ii) (Stationarity)

$$(z, d) \succsim_{s^t} (z, d') \text{ if and only if } d \succsim_{s^t} d'.$$

Axiom D4 *(Risk Equivalence Preservation).* For any $s^t \in S^t$, $p, p' \in \Delta(Z \times \mathcal{H})$, $d, d' \in \mathcal{D}$, and $\alpha \in (0, 1)$

$$[p \sim_{s^t} d \text{ and } p' \sim_{s^t} d'] \Longrightarrow [\alpha p + (1-\alpha)p' \sim_{s^t} \alpha d + (1-\alpha)d'].$$

By Axiom D2, for each $s^t \in S^t$ the preference $\succsim_{s^t}$ over degenerate lotteries of the form (z, h_{+1}) induces a preference over one-step-ahead acts. By a slight abuse of notation this induced preference will also be denoted $\succsim_{s^t}$.

Axiom D5 *(One-Step-Ahead Variational Preference).* For any $s^t \in S^t$, $h, h' \in \mathcal{H}_{+1}$, $d, d' \in \mathcal{D}$, and $\alpha \in (0, 1)$

(i) (Weak Certainty Independence)

$$\alpha h_{+1} + (1-\alpha)d \succsim_{s^t} \alpha h'_{+1} + (1-\alpha)d \Longrightarrow \alpha h_{+1} + (1-\alpha)d' \succsim_{s^t} \alpha h'_{+1} + (1-\alpha)d',$$

(ii) (Uncertainty Aversion)

$$h_{+1} \sim_{s^t} h'_{+1}$$
$$\Longrightarrow \alpha h_{+1} + (1-\alpha)h'_{+1} \succsim_{s^t} h_{+1}.$$

Axiom D6 *(Dynamic Consistency).* For any $s^t \in S^t$ and $h, h' \in \mathcal{H}_{+1}$

$$\left[h(s) \succsim_{s^t,s} h'(s) \text{ for all } s \in S\right] \Longrightarrow h \succsim_{s^t} h'.$$

Theorem 12. *Family* $\{\succsim_{s^t}\}$ *satisfies Axioms D1-D6 if and only if it has a variational representation* (29).

Proof. This proof adapts the proof of Hayashi's (2005) Theorem 1.

A.8.1 Lemmas

The following Lemmas of Hayashi (2005) hold for $\{\succsim_{s^t}\}$

Lemma H8. For any $s^t \in S^t$, $z \in Z$, and $h_{+1}, h'_{+1} \in \mathcal{H}_{+1}$ if $(c, h_{+1}(s)) \succsim_{s^t} (c, h'_{+1}(s))$ for every $s \in S$, then $(c, h_{+1}) \succsim_{s^t} (c, h'_{+1})$.

Lemma H9. For any $s \in S$, $h \in \mathcal{H}$, and $\mu \in \Delta(Z \times \mathcal{H})$ there exist risk equivalents $d, d' \in D$ such that $(z, h) \sim_{s^t} (z, d)$ and $\mu \sim_{s^t} d'$.

Lemma H10. For any $h \in \mathcal{H}$ there exists $h_{+1} \in \mathcal{H}_{+1}$ such that (i) $h(s) \sim_{s^t,s} h_{+1}(s)$ for all $s \in S$, (ii) $(z, h) \sim_{s^t} (z, h_{+1})$.

Hayashi (2005) Lemma 11 relies on C-independence and has to be weakended.

Lemma H11'. For any $s^t \in S^t$, $d, d', d'' \in \mathcal{D}$, and $\alpha \in (0,1)$ if $d \sim_{s^t} d'$ then $\alpha d + (1-\alpha)d'' \sim_{s^t} \alpha d' + (1-\alpha)d''$.

Proof. First show that $\frac{1}{2}d + \frac{1}{2}d'' \sim_{s^t} \frac{1}{2}d' + \frac{1}{2}d''$. This modifies part of the proof of Lemma 28 of Maccheroni et al. (2006a). Towards contradiction, suppose wlog $\frac{1}{2}d + \frac{1}{2}d'' \succ_{s^t} \frac{1}{2}d' + \frac{1}{2}d''$. By Axiom B5$(i)$, $\frac{1}{2}d + \frac{1}{2}d \succ_{s^t} \frac{1}{2}d' + \frac{1}{2}d$ and, by Axiom B5(i) again, $\frac{1}{2}d + \frac{1}{2}d' \succ_{s^t} \frac{1}{2}d' + \frac{1}{2}d'$; thus $d \succ_{s^t} d'$; contradiction. Second, because continuity implies mixture continuity, the conclusion follows from Theorem 2 of Herstein and Milnor (1953). □

From Axiom B4 and Lemma H11' follows

Lemma H12'. For any $s^t \in S^t$, $\mu, \mu', \mu'' \in \Delta(Z \times \mathcal{H})$, and $\alpha \in (0,1)$ if $\mu \sim_{s^t} \mu'$ then $\alpha\mu + (1-\alpha)\mu'' \sim_{s^t} \alpha\mu' + (1-\alpha)\mu''$.

Following Hayashi (2005), risk preference is uniquely determined by a history-independent preference $\succsim$ over $\mathcal{D}$. By Theorem 2 of Grandmont (1972), $\succsim$ is represented by $U : \mathcal{D} \to \mathbb{R}$ where $U(d) = \int u(z, d') \, \mathrm{d}d(z, d')$. By continuity and compactness, U can be chosen so that $U(\mathcal{D}) = [-M, M]$.

By continuity and Lemma H12', Theorem 2 of Grandmont (1972) implies that $\{\succsim_{s^t}\}$ is represented by a family $\{U_{s^t}\}$ where $U_{s^t}(\mu) : \Delta(Z \times \mathcal{H}) \to \mathbb{R}$ has $U_{s^t} = \int u_{s^t}(z, h) \, \mathrm{d}\mu(z, h)$ with $u_{s^t} : Z \times \mathcal{H} \to \mathbb{R}$ continuous.

By Axiom B2, $u_{s^t} = W_{s^t}(z, u_{s^t}(\hat{z}, h))$ for some fixed $\hat{z} \in Z$. Moreover, as argued by Hayashi (2005), W_{s^t} can be chosen to be independent of history and time. It will be denoted W.

A.8.2 Representation over one-step-ahead acts

As before, with a slight abuse of notation let $h_{+1} \succsim_{s^t} h'_{+1}$ iff $(z, h_{+1}) \succsim_{s^t} (z, h'_{+1})$ for some $z \in Z$ (which doesn't matter). By Axiom B1, $\succsim_{s^t}$ is a continuous, non-degenerate preference relation.

Thus, by Axiom B5 and Lemma H8 the assumptions of Maccheroni et al.'s (2006a) Theorem 3 are satisfied. Therefore, there exists a nonconstant affine function $v_{s^t} : \mathcal{D} \to \mathbb{R}$ and a grounded, convex and lower semicontinuous function $c_{s^t} : \Delta S \to [0, \infty]$ such that on H_{+1} preference $\succsim_{s^t}$ is represented by $V_{s^t}(h_{+1}) = \min_{p \in \Delta S} \int v_{s^t} \circ h_{+1} \, \mathrm{d}p + c_{s^t}(p)$ for all $h_{+1} \in H_{+1}$. By Axiom B3(i), preference $\succsim_{s^t}$ on $\mathcal{D}$ is history independent, so wlog $v_{s^t} = U$. Thus, on H_{+1} preference $\succsim_{s^t}$ is represented by $V_{s^t}(h_{+1}) = \min_{p \in \Delta S} \int U \circ h_{+1} \, \mathrm{d}p + c_{s^t}(p)$ for all $h_{+1} \in H_{+1}$.

Define function $U_{(s^t,h)}$ by $U_{(s^t,h)}(s) = U_{(s^t,s)}\big(h(s)\big)$. The following lemma is proved by Hayashi (2005).

Lemma H13. For any $h \in \mathcal{H}$ there exists $h_{+1} \in \mathcal{H}_{+1}$ such that $U_{(s^t,h)} = U_{(s^t,h_{+1})}$.

Thus, V_{s^t} represents $\succsim_{s^t}$ on the whole of $\mathcal{H}$. The aggregator W and full support of measures obtained as in Hayashi (2005) □

A.9 Attitudes Towards the Timing of Objective Risk

As in the model of Kreps and Porteus (1978), the aggregator W in representation (29) is responsible for preference for earlier resolution of objective risk. For any $s^t \in S$, $z \in Z$, $d_1, d_2 \in \mathcal{D}$, and $\alpha \in [0, 1]$ define $(1, \alpha; z, d_1, d_2)$ to be a temporal lottery where risk is resolved in period 1, i.e., whose chance node for period 0 is degenerate. Formally, define $(1,\alpha;z,d_1,d_2)=(z,\alpha d_1+(1-\alpha)d_2)$. In contrast, define $(0,\alpha;z,d_1,d_2)$ to be a temporal lottery where risk is resolved already in period 0, i.e., whose chance node for period 0 is not degenerate. Formally , define $(0,\alpha;z,d_1,d_2)=\alpha(z,d_1)+(1-\alpha)(z,d_2)$.

Definition 12. Relation $\succsim$ exhibits *preference for [resp., indifference to, preference against] earlier resolution of risk* if

$$(0,\alpha;z,d_1,d_2)\ \succsim_{s^t}\ \text{[resp., }\ \sim_{s^t},\precsim_{s^t}]\,(1,\alpha;z,d_1,d_2)$$

for all $t\geq 0$, $s^t\in S^t$, $z\in Z$, $d_1,d_2\in\mathcal{D}$, and $\alpha\in(0,1)$.

Preference for earlier resolution of purely objective risk is an important feature of preferences studied by Kreps and Porteus (1978), but is conceptually unrelated to uncertainty about subjective states.

Axiom D7 *(Risk Timing Indifference).* Preference $\succsim$ exhibits indifference to earlier resolution of risk.

Another important property of preferences is that tradeoffs between consumption at period t and $t+1$ are independent from consumption at later periods.[26]

Axiom D8 *(Future Separability).* For any $s^t\in S^t$, $d_{0,1},d'_{0,1}\in\Delta(Z\times\Delta(Z))$, and $y,y'\in Z^\infty$

$$(d_{0,1},y)\succsim_{s^t}(d'_{0,1},y)\quad\Longleftrightarrow\quad(d_{0,1},y')\succsim_{s^t}(d'_{0,1},y').$$

The following theorem extends Hayashi's (2005).

Theorem 13. *The family $\{\succsim_{s^t}\}$ satisfies Axioms D1-D8 if and only if the aggregator $W:Z\times R_U\to\mathbb{R}_U$ in (29) has the form $W(z,r)=u(z)+\beta r$; thus, the family is represented by*

$$U_{s^t}(\mu)=\int_{Z\times\mathcal{H}}\left\{u(z)+\beta\left[\min_{p\in\Delta S}\int_S U_{(s^t,s)}\big(h(s)\big)\,\mathrm{d}p(s)+c_{s^t}(p)\right]\right\}\mathrm{d}\mu(z,h)$$

where $u:Z\to\mathbb{R}$ is continuous and nonconstant and $\beta\in(0,1)$. Moreover, β is unique and the function u is unique up to positive affine transformations.

[26]This is Assumption 5 in Epstein (1983) and Axiom 8 in Hayashi (2005).

Proof. Follows from the proof of Corollary 1 in Hayashi (2005), which does not rely on certainty independence. $\square$

A.10 Proof of Theorem 7

By Theorem 13, $\succsim_{s^t}$ is represented by

$$U_{s^t}(\mu) = \int_{Z\times\mathcal{H}} \left\{ u(z) + \beta \left[\min_{p\in\Delta S} \int_S U_{(s^t,s)}\big(h(s)\big)\, \mathrm{d}p(s) + c_{s^t}(p) \right] \right\} \mathrm{d}\mu(z,h).$$

Thus, $\succsim_{s^t}$ on $\mathcal{H}_{+1}$ is represented by V_{s^t} where $V_{s^t}(h_{+1}) = I_{s^t}(U_{(s^t,h_{+1})})$, where, as before $U_{(s^t,h_{+1})}$ is defined as $U_{(s^t,h_{+1})}(s) = U_{(s^t,s)}\big(h_{+1}(s)\big)$.

Observe that $\succsim_{s^t}$ on $\mathcal{H}_{+1}$ is continuous, monotone (by Lemma H8), and satisfies tradeoff consistency (by Axiom B1). Moreover, $\mathcal{D}$ is a connected topological space. Thus, by Corollary 10 of Köbberling and Wakker (2003) there exists a unique probability $q_{s^t} \in \Delta(S)$ and a continuous function $\phi : \mathcal{D} \to \mathbb{R}$ that represents $\succsim_{s^t}$. Moreover, function ϕ is unique up to positive affine transformations.

As in other proofs, translation invariance of I_{s^t} leads to the Pexider equation for ϕ. As verified by Klibanoff et al. (2005) in the proof of their Proposition 2, even when the domain of ϕ is a bounded interval, as is the case here because of the compactness of $\Delta(Z)$ and continuity of u, the only solutions of the Pexider equation are ϕ_θ, where $\theta \in (0,\infty]$ is uniquely pinned down.

Because relation $\sim^*_{s^t}$ is constant across s^t, the scalar θ_{s^t} is constant across s^t. To see that, for each θ define $x(\theta) < 0$ which satisfies $\phi_\theta(1) - \phi_\theta(0) = \phi_\theta(0) - \phi_\theta(x)$. Thus, $x(\theta)$ is implicitly defined by $\Phi(\theta,x) = \phi_\theta(1) + \phi_\theta(x) - 2\phi_\theta(0)$. By the implicit function theorem, $\frac{\mathrm{d}x}{\mathrm{d}\theta} = -\frac{\mathrm{d}\Phi}{\mathrm{d}\theta}/\frac{\mathrm{d}\Phi}{\mathrm{d}x} = \frac{\exp(-\theta^{-1}) - x\cdot\exp(-x\theta^{-1})}{\theta\cdot\exp(-x\theta^{-1})} > 0$. Thus, for any two different values of θ, the corresponding values of $x(\theta)$ are different.

Let $s^t, \hat{s}^{\hat{t}}$ be distinct histories of possibly different length and recall that $\succsim$ over $\mathcal{D}$ is history independent. Let $U(d) = 1, U(d^*) = 0$, and assume that $U(d_{s^t}) = x(\theta_{s^t})$ and $U(d_{\hat{s}^{\hat{t}}}) = x(\theta_{\hat{s}^{\hat{t}}})$. Observe that $[d,d^*] \sim^*_{s^t} [d^*, d_{s^t}]$ and $[d,d^*] \sim^*_{\hat{s}^{\hat{t}}} [d^*, d_{\hat{s}^{\hat{t}}}]$. If $\theta_{\hat{s}^{\hat{t}}} \neq \theta_{s^t}$, then, wlog, $x(\theta_{\hat{s}^{\hat{t}}}) > x(\theta_{s^t})$, so $d_{\hat{s}^{\hat{t}}}$ is an improvement over d_{s^t}. This contradicts the equality $\sim^*_{\hat{s}^{\hat{t}}} = \sim^*_{s^t}$ and tradeoff consistency of both $\sim^*_{\hat{s}^{\hat{t}}}$ and $\sim^*_{s^t}$. $\square$

A.11 Proof of Theorem 8

Let $u_s = U(h_{+1}(s))$ and let $q_s = q(\{s\})$. Observe that

$$\begin{aligned} U(0; z, h_{+1}) &= \phi_\theta^{-1}\Bigg(\sum_{s\in S} \phi_\theta(u(z) + \beta u_s) q_s\Bigg) \\ &= \phi_\theta^{-1}\Bigg(\sum_{s\in S} \big[-\phi_\theta(u(z)) \cdot \phi_\theta(\beta u_s)\big] q_s\Bigg) \\ &= u(z) + \phi_\theta^{-1}\Bigg(\sum_{s\in S} \phi_\theta(\beta u_s) q_s\Bigg) \end{aligned}$$

and

$$U(1; z, h_{+1}) = u(z) + \beta\phi_\theta^{-1}\Bigg(\sum_{s\in S} \phi_\theta(u_s) q_s\Bigg).$$

Thus, $U(0; z, h_{+1}) > U(1; z, h_{+1})$ if and only if

$$\frac{1}{\beta}\phi_\theta^{-1}\Bigg(\sum_{s\in S} \phi_\theta(\beta u_s) q_s\Bigg) > \phi_\theta^{-1}\Bigg(\sum_{s\in S} \phi_\theta(u_s) q_s\Bigg)$$

if and only if

$$\phi_{\frac{\theta}{\beta}}^{-1}\Bigg(\sum_{s\in S} \phi_{\frac{\theta}{\beta}}(u_s) q_s\Bigg) > \phi_\theta^{-1}\Bigg(\sum_{s\in S} \phi_\theta(u_s) q_s\Bigg). \tag{30}$$

Because $\beta < 1$, the function ϕ_θ is a strictly concave transformation of $\phi_{\frac{\theta}{\beta}}$. Moreover, $q_s > 0$ for all $s \in S$ and by assumption there exist $s', s'' \in S$ such that $u_{s'} \neq u_{s''}$. Thus, inequality (30) follows from Jensen's inequality. □

A.12 Proof of Theorem 9

Follows from the reasoning in the proof of Theorem 8. □

A.13 Proof of Theorem 10

Let $\succsim$ be a stationary variational preference represented by

$$U(\mu) = \int_{Z\times\mathcal{H}} \Bigg[u(z) + \beta \min_{p\in\Delta S} \int_S U\big(h(s)\big)\, \mathrm{d}p(s) + c(p)\Bigg] \mathrm{d}\mu(z, h)$$

As before, $U(\mathcal{D}) = [-M, M] =: \mathcal{V}$. Define niveloid $I : B_0(\Sigma, \mathcal{V}) \to \mathbb{R}$ as $I(\xi) = \min_{p\in\Delta(S)} \int \xi \, \mathrm{d}p + c(p)$.

Suppose that $\xi \in B_0(\Sigma, \mathcal{V})$. For each $s \in S$ the value $\xi(s) \in \mathcal{V}$, so there exists $d_s \in \mathcal{D}$ such that $U(d_s) = \xi(s)$. Define $h \in \mathcal{H}_{+1}$ by $h(s) = d_s$ for all $s \in S$. Let $z_0, z_1 \in Z$. Because $\succsim$ satisfies indifference to the timing of resolution of uncertainty, $(z_0, (0; z_1, h)) \sim (z_0, (1; z_1, h))$. Thus, $u(z_0) + \beta I\big(u(z_1) + \beta\xi\big) = u(z_0) + \beta\big(u(z_1) + \beta I(\xi)\big)$. Hence, by translation invariance, $I(\beta\xi) = \beta I(\xi)$ for any $\xi \in B_0(\Sigma, \mathcal{V})$.

Let $0 < b < \beta$ and suppose that there exists $\xi \in B_0(\Sigma, \mathcal{V})$ such that $I(b\xi) \neq bI(\xi)$. Observe that, $I(b\xi) = I(b\xi + (1-b)0) \geq bI(\xi)$, by concavity and because $I(0) = 0$. Thus, $I(b\xi) > bI(\xi)$. Moreover, $I(\beta^n\xi) = I(\beta\beta^{n-1}\xi) = \beta I(\beta^{n-1}\xi) = \cdots = \beta^n I(\xi)$ for any $n \in \mathbb{N}$. Choose n such that $\beta^n < b$. For this n it follows that $\beta^n I(\xi) = I(\beta^n\xi) = I\big(\frac{\beta^n}{b} b\xi + \frac{b-\beta^n}{b} 0\big) \geq \frac{\beta^n}{b} I(b\xi) > \beta^n I(\xi)$. Contradiction.

Let $\beta < b < 1$ and suppose that there exists $\xi \in B_0(\Sigma, \mathcal{V})$ such that $I(b\xi) \neq bI(\xi)$. As above $I(b\xi) > bI(\xi)$ follows. Moreover, $I(b^n\xi) = I(b^{n-1}b\xi) \geq b^{n-1}I(b\xi) > b^n I(\xi)$ for any $n \in \mathbb{N}$. Choose n such that $b^n < \beta$. Contradiction with the case $0 < b < \beta$.

As a consequence, I is a niveloid on $B_0(\Sigma, \mathcal{V})$ that is homogenous of degree one. Extend I to $B_0(\Sigma)$ by homogeneity. Observe that the extension is a normalized niveloid, thus it satisfies the assumptions of Lemma 3.5 of Gilboa and Schmeidler (1989); therefore, there exists a closed and convex set $C \subseteq \Delta(S)$ such that $I(\xi) = \min_{p\in C} \int \xi \, \mathrm{d}p$ for all $\xi \in B_0(\Sigma, \mathcal{V})$. □

References

Abdellaoui, M., A. Baillon, and P. P. Wakker (2007): “Combining Bayesian Beliefs and Willingness to Bet to Analyze Attitudes towards Uncertainty,” *mimeo.*

Aczél, J. (1966): *Lectures on Functional Equations and Their Applications.*, NewYork: Academic Press.

Allais, M. (1953): “Le Comportement de l’Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l’Ecole Americaine,” *Econometrica*, 21, 503–546.

Anderson, E. W., L. P. Hansen, and T. J. Sargent (2000): “Robustness, Detection, and the Price of Risk,” *mimeo.*

Anscombe, F. and R. Aumann (1963): "A Definition of Subjective Probability," *The Annals of Mathematical Statistics*, 34, 199–205.

Arrow, K. (1970): *Essays in the Theory of Risk-Bearing*, Amsterdam: North-Holland.

Barillas, F., L. P. Hansen, and T. J. Sargent (2007): "Doubts or Variability?" *mimeo.*

Chen, Z. and L. Epstein (2002): "Ambiguity, Risk, and Asset Returns in Continuous Time," *Econometrica*, 70, 1403–1443.

Chew, S. H. and J. S. Sagi (2007): "Small Worlds: Modeling Attitudes towards Sources of Uncertainty," *Journal of Economic Theory*, forthcoming.

Dupuis, P. and R. S. Ellis (1997): *A Weak Convergence Approach to the Theory of Large Deviations*, Wiley, New York.

Ellsberg, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *The Quarterly Journal of Economics*, 75, 643–669.

Epstein, L. (1983): "Stationary cardinal utility and optimal growth under uncertainty," *Journal of Economic Theory*, 31, 133–152.

Epstein, L. and S. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Epstein, L. G. and M. Schneider (2003): "IID: independently and indistinguishably distributed," *Journal of Economic Theory*, 113, 32–50.

Ergin, H. and F. Gul (2004): "A Subjective Theory of Compound Lotteries," *mimeo.*

Fishburn, P. C. (1970): *Utility Theory for Decision Making*, Wiley, New York.

Ghirardato, P., F. Maccheroni, M. Marinacci, and M. Siniscalchi (2003): "A Subjective Spin on Roulette Wheels," *Econometrica*, 71, 1897–1908.

Gilboa, I. and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18, 141–153.

Grandmont, J.-M. (1972): "Continuity Properties of a von Neumann-Morgenstern Utility," *Journal of Economic Theory*, 4, 45–57.

GRANT, S. AND B. POLAK (2006): "Bayesian beliefs with stochastic monotonicity: An extension of Machina and Schmeidler," *Journal of Economic Theory*, 130, 264–282.

——— (2007): "Absolute Ambiguity Aversion and Mean-Dispersion Preferences," *mimeo*.

HALEVY, Y. AND E. OZDENOREN (2007): "Uncertainty and Compound Lotteries: Calibration," *mimeo*.

HANSEN, L. P. AND T. J. SARGENT (2001): "Robust Control and Model Uncertainty," *The American Economic Review*, 91, 60–66.

——— (2007): *Robustness*, Princeton University Press.

HANSEN, L. P., T. J. SARGENT, G. TURMUHAMBETOVA, AND N. WILLIAMS (2006): "Robust control and model misspecification," *Journal of Economic Theory*, 128, 45–90.

HAYASHI, T. (2005): "Intertemporal substitution, risk aversion and ambiguity aversion," *Economic Theory*, 25, 933–956.

HERSTEIN, I. N. AND J. MILNOR (1953): "An Axiomatic Approach to Measurable Utility," *Econometrica*, 21, 291–297.

JACOBSON, D. J. (1973): "Optimal Linear Systems with Exponential Performance Criteria and their Relation to Differential Games," *IEEE Transactions on Automatic Control*, 18, 124–131.

KARANTOUNIAS, A. G., L. P. HANSEN, AND T. J. SARGENT (2007): "Ramsey taxation and fear of misspecication," *mimeo*.

KEYNES, J. (1921): *A Treatise on Probability*, MacMillan.

KLESHCHELSKI, I. AND N. VINCENT (2007): "Robust Equilibrium Yield Curves," *mimeo*.

KLIBANOFF, P., M. MARINACCI, AND S. MUKERJI (2005): "A smooth model of decision making under ambiguity," *Econometrica*, 73, 1849–1892.

KNIGHT, F. (1921): *Risk, Uncertainty and Profit*, Houghton Mifflin Company.

KÖBBERLING, V. AND P. P. WAKKER (2003): "Preference Foundations for Nonexpected Utility: A Generalized and Simplified Technique," *Mathematics of Operations Research*, 28, 395–423.

Kreps, D. and E. Porteus (1978): “Temporal Resolution of Uncertainty and Dynamic Choice Theory,” *Econometrica*, 46, 185–200.

Maccheroni, F., M. Marinacci, and A. Rustichini (2004): “Ambiguity Aversion, Malevolent Nature, and the Variational Representation of Preferences,” *mimeo*.

——— (2006a): “Ambiguity Aversion, Robustness, and the Variational Representation of Preferences,” *Econometrica*, 74, 1447 – 1498.

——— (2006b): “Dynamic variational preferences,” *Journal of Economic Theory*, 128, 4–44.

Machina, M. and D. Schmeidler (1995): “Bayes Without Bernoulli: Simple Conditions for Probabilistically Sophisticated Choice,” *Journal of Economic Theory*, 108–128.

Machina, M. J. and D. Schmeidler (1992): “A More Robust Definition of Subjective Probability,” *Econometrica*, 60, 745–780.

Marinacci, M. (2002): “Probabilistic Sophistication and Multiple Priors,” *Econometrica*, 70, 755–764.

Markowitz, H. M. (1952): “Portfolio Selection,” *Journal of Finance*, 7, 77–91.

Nau, R. (2001): “Uncertainty Aversion with SecondOrder Utilities and Probabilities,” in *Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications*, http://www.sipta.org/isipta01/proceedings/063.html.

——— (2006): “Uncertainty Aversion with SecondOrder Utilities and Probabilities,” *Management Science*, 52, 136–145.

Neilson, W. S. (1993): “Ambiguity Aversion: An Axiomatic Approach Using Second Order Probabilities,” *mimeo*.

Neyman, J. and E. Scott (1948): “Consistent Estimates Based on Partially Consistent Observations,” *Econometrica*, 16, 1–32.

Sarin, R. and P. P. Wakker (2000): “Cumulative dominance and probabilistic sophistication,” *Mathematical Social Sciences*, 40, 191–196.

Savage, L. (1972): *The Foundations of Statistics*, Dover.

Schmeidler, D. (1989): “Subjective Probability and Expected Utility without Additivity,” *Econometrica*, 57, 571–587.

Skiadas, C. (2003): "Robust control and recursive utility," *Finance and Stochastics*, 7, 475 – 489.

Tobin, J. (1958): "Liquidity Preference as Behavior Toward Risk," *Review of Economic Studies*, 25, 65–86.

Villegas, C. (1964): "On Qualitative Probability σ-Algebras," *The Annals of Mathematical Statistics*, 35, 1787–1796.

Whittle, P. (1981): "Risk-Sensitive Linear/Quadratic/Gaussian Control," *Advances in Applied Probability*, 13, 764–777.

Woodford, M. (2006): "An Example of Robustly Optimal Monetary Policy with Near-Rational Expectations," *Journal of the European Economic Association*, 4, 386–395.