

KIER DISCUSSION PAPER SERIES

KYOTO INSTITUTE OF ECONOMIC RESEARCH

Discussion Paper No.1019

“A Ranking over "More Risk Averse Than" Relations
and its Application to the Smooth Ambiguity Model”

Chiaki Hara

January 2020



KYOTO UNIVERSITY

KYOTO, JAPAN

A Ranking over “More Risk Averse Than” Relations and its Application to the Smooth Ambiguity Model

Chiaki Hara*

January 30, 2020

Abstract

Given two pairs of expected utility functions, we formalize the notion that one expected utility function is more risk-averse than the other in the first pair to a greater extent than in the second pair. We do so by assuming that the utility functions are twice continuously differentiable and satisfy the Inada condition, and, in each of the two pairs, using the function that transforms the derivatives of one expected utility function to the derivatives of the other, rather than the function that transforms one expected utility function to the other. This definition allows us to interpret the quantitative results on the ambiguity aversion coefficients of the smooth ambiguity model of Klibanoff, Marinacci, and Mukerji (2005) in some cases not covered by the more-ambiguity-averse-than relation that they conceived.

JEL Classification Codes: C38, D81, G11.

Keywords: Expected utility functions, risk aversion, ambiguity aversion, smooth ambiguity model.

1 Introduction

An expected utility function (also known as a Bernoulli utility function) is said to be more risk-averse than another if the former is a concave transformation of the latter. The purpose of this paper is, when two pairs of expected utility functions are given, to formalize the idea that one expected utility function is more risk-averse than the other in the first pair to a greater extent than in the second pair. In symbols, if v_1 and u_1 constitute the first pair and v_2 and u_2 constitute the second pair, then we wish to give a rigorous and sufficiently general definition to the statement that v_1 is more risk-averse than u_1 to a greater extent than v_2 is more risk-averse than u_2 . In other words, based on

*I thank Shoko Negishi, Jonathan Newton, and, especially, Sujoy Mukerji for helpful comments. This research is funded by the Open Research Area (ORA) for the Social Sciences “Ambiguity in Dynamic Environments.”

the standard more-risk-averse-than relation, we introduce a new binary relation over the differences in risk aversion between two expected utility functions.

The expected utility function v_1 is more risk-averse than u_1 , and v_2 is more risk-averse than u_2 , if and only if there are two concave functions φ_1 and φ_2 such that $v_1 = \varphi_1 \circ u_1$ and $v_2 = \varphi_2 \circ u_2$. The most natural approach to formalize that statement that v_1 is more risk-averse than u_1 to a greater extent than v_2 is more risk-averse than u_2 , is to require φ_1 to be more concave than φ_2 . But this statement makes sense only if φ_1 and φ_2 have the same domain, that is, u_1 and u_2 have the same range. In many applications, this assumption is violated.

Our approach is, instead, to assume that the utility functions are twice continuously differentiable and satisfy the Inada condition, and use the function that transforms the derivatives of one utility expected function to the derivatives of the other expected utility function. In symbols, we define two functions $\psi_1 : \mathbf{R}_{++} \rightarrow \mathbf{R}_{++}$ and $\psi_2 : \mathbf{R}_{++} \rightarrow \mathbf{R}_{++}$ by $v'_1 = \psi_1 \circ u'_1$ and $v'_2 = \psi_2 \circ u'_2$ and compare ψ_1 and ψ_2 . These functions ψ_1 and ψ_2 have the same domain because the utility functions u_1 and u_2 are assumed to satisfy the Inada condition, and in our definition we rank ψ_1 and ψ_2 in terms of their elasticities rather than the curvature (which is used when comparing φ_1 and φ_2). We will also give necessary and sufficient conditions of this definition in terms of choice behavior between a random and a deterministic consumption plans.

This study is motivated by the smooth ambiguity model of Klibanoff, Marinacci, and Mukerji (2005, hereafter KMM). Their utility functions are defined in the form of nested expected utilities, in which the inner expected utilities, and the associated conditional certainty equivalents, are taken for a utility function u conditional on probability measures on the state space, and the outer expected utility is taken for a utility function v over the conditional certainty equivalents with respect to what they termed as the second-order belief. The decision maker is ambiguity-averse if the outer utility function v is more risk-averse than the inner utility function u . Theorem 2 of KMM proved that the curvature (concavity) of the transformation function φ satisfying $v = \varphi \circ u$ measures his ambiguity aversion, much in the same way as the Arrow-Pratt measure of absolute risk aversion measures risk aversion.

As emphasized by KMM themselves, a caveat on their more-ambiguity-averse-than relation is in order. The theorem implies that whenever one utility function is more ambiguity-averse than another, they share essentially the same inner utility function u . But it is a common practice in empirical studies to estimate or calibrate the curvature of the transformation function φ (which is the KMM measure of ambiguity aversion) or of the outer utility function v , without fixing the inner utility function u a priori. Thus, for two ambiguity-averse utility functions having two different inner utility functions, we cannot conclude that one is more ambiguity-averse than the other even when the transformation function φ of the former is more concave than the latter. This significantly

limits the scope within which we can interpret and compare various quantitative results. Our definition, on the other hand, is tailored to the need for a wider scope of numerical comparison. It can be applied to two pairs of which the inner utility functions are different and, in addition, has a clear equivalent condition in terms of the decision makers' choices. Thus, it allows the researcher to make a quantitative assessment on the KMM measure of ambiguity aversion with no reference to the associated Arrow-Pratt measure for pure risks.¹

This paper is organized as follows. Section 2 lays out the setup of the paper and gives some preliminary results. Section 3 presents a new relation between two pairs of expected utility functions. Section 4 gives examples of the new relation for the case of constant absolute or relative risk aversion. Section 5 provides an essentially equivalent necessary and sufficient condition for the new relation in terms of the decision makers' choice behavior. Section 6 discusses applications to the utility functions of KMM. Section 7 gives a summary and suggests a couple of directions of future research. All proofs are in the appendix.

2 Setup

Let I be a non-degenerate (containing at least two points) open interval of \mathbf{R} and $u : I \rightarrow \mathbf{R}$. Assume that u is twice continuously differentiable and that $u'' < 0 < u'$. We also impose the Inada condition, that is, $u'(x) \rightarrow 0$ as $x \rightarrow \sup I$, and $u'(x) \rightarrow \infty$ as $x \rightarrow \inf I$. We call these conditions the basic conditions.

Denote by the range of $u : I \rightarrow \mathbf{R}$ by $\text{Ran } u$, that is, $\text{Ran } u = u(I) = \{u(x) \mid x \in I\}$. $\text{Ran } u'$ is analogously defined. Since $u'' < 0$, the Inada condition is equivalent to $\text{Ran } u' = \mathbf{R}_{++}$.

For a utility function $u : I \rightarrow \mathbf{R}$, we define the Arrow-Pratt measure of absolute risk aversion $a(\cdot, u) : I \rightarrow \mathbf{R}_{++}$ by letting $a(x, u) = -u''(x)/u'(x)$ for every $x \in I$. For $x > 0$, we define the Arrow-Pratt measure of relative risk aversion as $r(x, u) = -u''(x)x/u'(x)$.

The utility functions that exhibit constant absolute or relative risk aversion satisfy the basic conditions, but their ranges are different. In fact, let $I = \mathbf{R}$ and, with $\gamma > 0$,

$$u(x) = -\frac{1}{\gamma} \exp(-\gamma x). \quad (1)$$

Then u has the constant coefficient γ of absolute risk aversion, and $\text{Ran } u = -\mathbf{R}_{++}$. Let

¹A similar complication arises in recursive utility as well. For example, in presenting functional forms of recursive utility, Epstein (1992, equalities (4.23) and (4.24)) restricted the constant coefficient of relative risk aversion to be at most one and the intertemporal elasticity of substitution to be at least one. But when it comes to estimating these values in any quantitative work, other functional forms are also necessary.

$I = \mathbf{R}_{++}$ and, with $\gamma > 0$,

$$u(x) = \begin{cases} \ln x & \text{if } \gamma = 1, \\ \frac{x^{1-\gamma}}{1-\gamma} & \text{otherwise.} \end{cases} \quad (2)$$

Then u has the constant coefficient of relative risk aversion, and

$$\text{Ran } u = \begin{cases} \mathbf{R}_{++} & \text{if } \gamma < 1, \\ \mathbf{R} & \text{if } \gamma = 1, \\ -\mathbf{R}_{++} & \text{if } \gamma > 1. \end{cases}$$

The following proposition covers the case where the coefficients of relative risk aversion are not constant.

Proposition 1 *Suppose that an expected utility function $u : \mathbf{R}_{++} \rightarrow \mathbf{R}$ satisfies the basic conditions.*

1. *If there is a $b > 0$ such that $r(x, u) \leq 1$ for every $x \geq b$, then $\text{Ran } u$ is not bounded from above.*
2. *If there are a $b > 0$ and a $g \in (0, 1)$ such that $r(x, u) < g$ for every $x \leq b$, then $\text{Ran } u$ is bounded from below.*
3. *If there is a $b > 0$ such that $r(x, u) \geq 1$ for every $x \leq b$, then $\text{Ran } u$ is not bounded from below.*
4. *If there are a $b > 0$ and a $g \in (1, \infty)$ such that $r(x, u) > g$ for every $x \geq b$, then $\text{Ran } u$ is bounded from above.*

Since these results will not be used in the subsequent analysis and their proofs are elementary, we omit them. The message of the proposition is that the range of a utility functions is closely related to the risk attitude that it represents and, hence, an additional restriction on it may well turn out to be a significant restriction on the risk attitude. Since the range of a utility function is the domain of the function that transforms the utility function to another one, the implication of this proposition for a formal definition of the statement that one utility function is more risk-averse than the other in the first pair to a greater extent than in the second is that the function that transforms one utility function to another should not be used. We will, instead, use the function that transforms the derivative of a utility function to the derivative of another.

3 Definition of the new relation

To understand our definition of the ranking over more-risk-averse-than relations, which we will later give, the following proposition is helpful.

Proposition 2 *Suppose that two expected utility functions $u : I \rightarrow \mathbf{R}$ and $v : I \rightarrow \mathbf{R}$ satisfy the basic conditions. Define $\psi : \mathbf{R}_{++} \rightarrow \mathbf{R}_{++}$ by $\psi = v' \circ (u')^{-1}$. Then, for every $x \in I$,*

$$\frac{a(x, v)}{a(x, u)} = \frac{\psi'(u'(x))u'(x)}{\psi(u'(x))}. \quad (3)$$

It is easy to check that $\psi' > 0$. As $y \rightarrow 0$, $(u')^{-1}(y) \rightarrow \sup I$. Thus, $\psi(y) = v'((u')^{-1}(y)) \rightarrow 0$. Analogously, $\psi(y) \rightarrow \infty$ as $y \rightarrow \infty$. Define the elasticity of the transformation function ψ , $e(\cdot, \psi) : \mathbf{R}_{++} \rightarrow \mathbf{R}_{++}$ by

$$e(y, \psi) = \frac{\psi'(y)y}{\psi(y)}.$$

Then (3) can be rewritten as

$$\frac{a(x, v)}{a(x, u)} = e(u'(x), \psi). \quad (4)$$

Proposition 2 implies that v is at least as risk-averse as u if and only if v' is an elastic transformation (that is, everywhere having elasticity greater than or equal to one) of u' . In particular, a proportional increase in the Arrow-Pratt measure of absolute risk aversion from u to v is equal to the elasticity of the transformation.

Definition 1 *Suppose that four expected utility functions $u_1 : I_1 \rightarrow \mathbf{R}$, $v_1 : I_1 \rightarrow \mathbf{R}$, $u_2 : I_2 \rightarrow \mathbf{R}$, and $v_2 : I_2 \rightarrow \mathbf{R}$ satisfy the basic conditions. Write $\psi_1 = v'_1 \circ (u'_1)^{-1}$ and $\psi_2 = v'_2 \circ (u'_2)^{-1}$.*

1. We say that v_1 is more risk averse than u_1 at least to the same extent as v_2 is more risk averse than u_2 , if

$$e(y_1, \psi_1) \geq e(y_2, \psi_2) \quad (5)$$

for every $y_1 \in \mathbf{R}_{++}$ and $y_2 \in \mathbf{R}_{++}$. We then write $(u_1, v_1) \blacktriangleright (u_2, v_2)$.

2. We say that v_1 is more risk averse than u_1 to a greater extent than v_2 is more risk averse than u_2 , if

$$e(y_1, \psi_1) > e(y_2, \psi_2) \quad (6)$$

for every $y_1 \in \mathbf{R}_{++}$ and $y_2 \in \mathbf{R}_{++}$. We then write $(u_1, v_1) \blacktriangleright (u_2, v_2)$.

In the definition, the domains I_1 and I_2 may be different, and the ranges $\text{Ran } u_1$, $\text{Ran } v_1$, $\text{Ran } u_2$, $\text{Ran } v_2$ may all be different. The two levels of marginal utility, y_1 and y_2

that appear on each of the two sides of (5) and (6) may be taken to be different. If they were taken to be equal, then the conditions would be written as

$$e(y, \psi_1) \geq e(y, \psi_2) \text{ or } e(y, \psi_1) > e(y, \psi_2), \quad (7)$$

for every $y \in \mathbf{R}_{++}$, and we may say that ψ_1 is more elastic as ψ_2 . If, in addition, we followed the terminology of the strongly-more-risk-averse-than relation of Ross (1981), we could say that that ψ_1 is strongly more elastic as ψ_2 . By (3), (6) is equivalent to the conditions that

$$\frac{a(x_1, v_1)}{a(x_1, u_1)} > \frac{a(x_2, v_2)}{a(x_2, u_2)} \quad (8)$$

for every $x_1 \in I_1$ and $x_2 \in I_2$.

Both \blacktriangleright and \triangleright are transitive, \triangleright is irreflexive, but \blacktriangleright is neither reflexive nor irreflexive. Moreover, \triangleright is included in the asymmetric (strict) part of \blacktriangleright (that is, if $(u_1, v_1) \triangleright (u_2, v_2)$, then $(u_1, v_1) \blacktriangleright (u_2, v_2)$ and $(u_2, v_2) \not\blacktriangleright (u_1, v_1)$), and the former is strictly smaller than the latter.²

Instead of saying that u_1 is more risk-averse than v_1 to a greater extent than u_2 is more risk-averse than v_2 , we could say more informally that u_1 is more risk-averse than v_1 , and even more so than u_2 is more risk-averse than v_2 . For brevity, we shall thus refer to the binary relations \triangleright and \blacktriangleright as the even-more-risk-averse-than relation in the rest of this paper.

4 Examples

In this section, we give examples of the even-more-risk-averse-than relation that involve constant absolute or relative risk aversion. These examples involves transformation functions ψ_n from u'_n to v'_n that have constant elasticities. The first, simplest, example deals with constant absolute risk aversion.

Example 1 Suppose that four expected utility functions $u_1 : I_1 \rightarrow \mathbf{R}$, $v_1 : I_1 \rightarrow \mathbf{R}$, $u_2 : I_2 \rightarrow \mathbf{R}$, and $v_2 : I_2 \rightarrow \mathbf{R}$ have constant coefficients $\gamma_1, \eta_1, \gamma_2$, and η_2 of absolute risk aversion (1). Then

$$\psi_n(y) = y^{\eta_n/\gamma_n} \quad (9)$$

for every $n = 1, 2$ and every $y \in \mathbf{R}_{++}$. Hence,

$$e(\cdot, \psi_n) = \eta_n/\gamma_n \quad (10)$$

²For every (u_1, v_2) and every (u_2, v_2) , $(u_1, v_1) \blacktriangleright (u_2, v_2)$ and $(u_2, v_2) \blacktriangleright (u_1, v_1)$ if and only if $e(\cdot, \psi_1)$ and $e(\cdot, \psi_2)$ take the same constant value. Thus, the symmetric part of \blacktriangleright corresponds to the pair (ψ_1, ψ_2) of identical transformations that have a constant elasticity.

for every $n = 1, 2$. Thus, $(u_1, v_1) \blacktriangleright (u_2, v_2)$ if and only if $\eta_1/\gamma_1 \geq \eta_2/\gamma_2$, and $(u_1, v_1) \triangleright (u_2, v_2)$ if and only if $\eta_1/\gamma_1 > \eta_2/\gamma_2$.

The next one deals with constant relative risk aversion.

Example 2 Suppose that four expected utility functions $u_1 : I_1 \rightarrow \mathbf{R}$, $v_1 : I_1 \rightarrow \mathbf{R}$, $u_2 : I_2 \rightarrow \mathbf{R}$, and $v_2 : I_2 \rightarrow \mathbf{R}$ have constant coefficients γ_1 , η_1 , γ_2 , and η_2 of relative risk aversion (2). Then, (9) and (10) hold. Thus, $(u_1, v_1) \blacktriangleright (u_2, v_2)$ if and only if $\eta_1/\gamma_1 \geq \eta_2/\gamma_2$, and $(u_1, v_1) \triangleright (u_2, v_2)$ if and only if $\eta_1/\gamma_1 > \eta_2/\gamma_2$.

Example 2 looks much the same as Example 1, but it illuminates what can be brought about by our use of the function ψ_n that transforms u'_n to v'_n , rather than the function φ_n that transforms u_n to v_n ($n = 1, 2$). Indeed, if we used the latter, the domain of φ_n coincides with $\text{Ran } u_n$, which may be either \mathbf{R}_{++} or $-\mathbf{R}_{++}$, depending on whether γ_n is smaller or greater than one. For example, if $\gamma_1 < 1 < \gamma_2$, then the domain of φ_1 coincides with \mathbf{R}_{++} , while the domain of φ_2 coincides with $-\mathbf{R}_{++}$. Hence, it does not make sense to say that one of them is more concave than the other, and we cannot conclude that v_1 is more risk averse than u_1 to a greater extent than v_2 is more risk averse than u_2 or the other way around.

Example 3 Suppose that two expected utility functions $u_1 : I_1 \rightarrow \mathbf{R}$ and $v_1 : I_1 \rightarrow \mathbf{R}$ have constant coefficients γ_1 and η_1 of absolute risk aversion (1), and two expected utility functions $u_2 : I_2 \rightarrow \mathbf{R}$ and $v_2 : I_2 \rightarrow \mathbf{R}$ have constant coefficients γ_2 and η_2 of relative risk aversion (2). Then (9) and (10) hold. Thus, $(u_1, v_1) \blacktriangleright (u_2, v_2)$ if and only if $\eta_1/\gamma_1 \geq \eta_2/\gamma_2$, and $(u_1, v_1) \triangleright (u_2, v_2)$ if and only if $\eta_1/\gamma_1 > \eta_2/\gamma_2$.

This example is an immediate consequence of the first two, but it would have been impossible to compare a pair of expected utility functions of constant absolute risk aversion and a pair of expected utility functions of constant relative risk aversion, if we had stuck to the comparison by means of the function φ_n that transforms u_n to v_n . Since $I_1 = \mathbf{R}$ and $I_2 \in \mathbf{R}_{++}$ or $I_2 = -\mathbf{R}_{++}$, $I_1 \neq I_2$. Thus, the example also shows that the comparison of the more-risk-averse than relation is possible even when the domains are different.

The following example is a generalization of the previous one, in that the expected utility functions have decreasing hyperbolic absolute risk aversion.

Example 4 For each $n = 1, 2$, let $b_n \in \mathbf{R}$ and the four expected utility functions $u_1 : (b_1, \infty) \rightarrow \mathbf{R}$, $v_1 : (b_1, \infty) \rightarrow \mathbf{R}$, $u_2 : (b_2, \infty) \rightarrow \mathbf{R}$, and $v_2 : (b_2, \infty) \rightarrow \mathbf{R}$ have hyperbolic

absolute risk aversion with the cautiousness parameters γ_n and η_n :³

$$a(x_n, u_n) = \frac{1}{\gamma_n(x_n - b_n)},$$

$$a(x_n, v_n) = \frac{1}{\eta_n(x_n - b_n)},$$

Then,

$$\frac{a(x_n, v_n)}{a(x_n, u_n)} = \frac{\gamma_n}{\eta_n}$$

for every $x_n \in (b_n, \infty)$. Thus, $(u_1, v_1) \blacktriangleright (u_2, v_2)$ if and only if $\eta_1/\gamma_1 \leq \eta_2/\gamma_2$, and $(u_1, v_1) \triangleright (u_2, v_2)$ if and only if $\eta_1/\gamma_1 < \eta_2/\gamma_2$. That is, the even-more-risk-averse-than relation can be characterized as a larger proportional decrease in the cautiousness.

In the following example the expected utility functions have increasing, rather than decreasing, hyperbolic absolute risk aversion. It covers the case of quadratic expected utility functions.

Example 5 For each $n = 1, 2$, let $b_n \in \mathbf{R}$ and the four expected utility functions $u_1 : (-\infty, b_1) \rightarrow \mathbf{R}$, $v_1 : (-\infty, b_1) \rightarrow \mathbf{R}$, $u_2 : (-\infty, b_2) \rightarrow \mathbf{R}$, and $v_2 : (-\infty, b_2) \rightarrow \mathbf{R}$ have hyperbolic absolute risk aversion with the cautiousness parameters γ_n and η_n :

$$a(x_n, u_n) = \frac{1}{\gamma_n(b_n - x_n)},$$

$$a(x_n, v_n) = \frac{1}{\eta_n(b_n - x_n)},$$

Then,

$$\frac{a(x_n, v_n)}{a(x_n, u_n)} = \frac{\gamma_n}{\eta_n}$$

for every $x_n \in (-\infty, b_n)$. Thus, $(u_1, v_1) \blacktriangleright (u_2, v_2)$ if and only if $\eta_1/\gamma_1 \leq \eta_2/\gamma_2$, and $(u_1, v_1) \triangleright (u_2, v_2)$ if and only if $\eta_1/\gamma_1 < \eta_2/\gamma_2$. That is, the even-more-risk-averse-than relation can be characterized as a larger proportional decrease in the absolute values of the cautiousness.

The above five examples all involve pairs of expected utility functions for which the ratio of the coefficients of absolute risk aversion, (2), is constant. Thus, every pair in these examples can be compared with every other pair in the examples with respect to \blacktriangleright .

The following example is due to Collard, Mukerji, Sheppard, and Tallon (2018). It is different from the previous ones in that there may be no ranking with respect to the even-more-risk-averse-than relation.

³The cautiousness is defined as the derivative of the reciprocal of the coefficients of absolute risk aversion. This terminology is due to Wilson (1968).

Example 6 Suppose that two expected utility functions $u_1 : I_1 \rightarrow \mathbf{R}$ and $u_2 : I_2 \rightarrow \mathbf{R}$ have constant coefficients γ_1 and γ_2 of relative risk aversion (2). For each $n = 1, 2$, let $\alpha_n > 0$ and assume that φ_n has the same functional form as the expected utility function of constant absolute risk aversion (1), with the parameter γ replaced by α_n . Define $v_n = \varphi_n \circ u_n$. Then

$$\begin{aligned} v_n(x) &= -\frac{1}{\alpha_n} \exp\left(-\frac{\alpha_n}{1-\gamma_n} x^{1-\gamma_n}\right), \\ v'_n(x) &= x^{-\gamma_n} \exp\left(-\frac{\alpha_n}{1-\gamma_n} x^{1-\gamma_n}\right), \end{aligned} \quad (11)$$

and the basic conditions are met. Define $\psi_n = v'_n \circ (u'_n)^{-1}$, then

$$\begin{aligned} \psi_n(y_n) &= (y_n^{-1/\gamma_n})^{-\gamma_n} \exp\left(-\frac{\alpha_n}{1-\gamma_n} (y_n^{-1/\gamma_n})^{1-\gamma_n}\right) = y_n \exp\left(-\frac{\alpha_n}{1-\gamma_n} y_n^{1-1/\gamma_n}\right) \\ \psi'_n(y_n) &= \exp\left(-\frac{\alpha_n}{1-\gamma_n} y_n^{1-1/\gamma_n}\right) + y_n \exp\left(-\frac{\alpha_n}{1-\gamma_n} y_n^{1-1/\gamma_n}\right) \left(-\frac{\alpha_n}{1-\gamma_n} \left(1 - \frac{1}{\gamma_n}\right) y_n^{-1/\gamma_n}\right) \\ &= \left(1 + \frac{\alpha_n}{\gamma_n} y_n^{1-1/\gamma_n}\right) \exp\left(-\frac{\alpha_n}{1-\gamma_n} y_n^{1-1/\gamma_n}\right). \end{aligned}$$

Thus,

$$e(y_n, \psi_n) = \frac{\psi'_n(y_n) y_n}{\psi_n(y_n)} = 1 + \frac{\alpha_n}{\gamma_n} y_n^{1-1/\gamma_n}.$$

Thus, $(u_1, v_1) \blacktriangleright (u_2, v_2)$ if and only if

$$1 + \frac{\alpha_1}{\gamma_1} y_1^{1-1/\gamma_1} \geq 1 + \frac{\alpha_2}{\gamma_2} y_2^{1-1/\gamma_2},$$

which is equivalent to

$$\frac{\alpha_1 \gamma_2}{\alpha_2 \gamma_1} \geq y_1^{1/\gamma_1 - 1} y_2^{1-1/\gamma_2} \quad (12)$$

for every y_1 and every y_2 . If $\gamma_1 = \gamma_2 = 1$, then the right-hand side is equal to one and the inequality holds if and only if $\alpha_1 \geq \alpha_2$.⁴ Otherwise, the right-hand side can take any value in \mathbf{R}_{++} as we vary y_1 or y_2 . Thus, there is no value of the α_n 's and the γ_n 's such that $(u_1, v_1) \blacktriangleright (u_2, v_2)$.

In this example, since $I_1 = I_2 = \mathbf{R}_{++}$, we can take $y_n = u'_n(x)$ for each n with a common consumption level $x \in \mathbf{R}_{++}$ in the above example. Then (12) can be reduced to

$$\frac{\alpha_1 \gamma_2}{\alpha_2 \gamma_1} \geq x^{\gamma_1 - \gamma_2},$$

⁴In this case, u_n has constant coefficient 1 of relative risk aversion and v_n has a constant coefficient $1 + \alpha_n$ of relative risk aversion. This will be shown by (13). In this case, (u_1, v_1) and (u_2, v_2) can be compared by the more-ambiguity-averse-than relation of KMM.

which holds for every x if and only if $\gamma_1 = \gamma_2$ and $\alpha_1 \geq \alpha_2$. That is, (8) holds whenever $x_1 = x_2$ if and only if $\gamma_1 = \gamma_2$ and $\alpha_1 \geq \alpha_2$. This highlights a difference between our definition of the even-more-risk-averse-than relation and the ambiguity measure of KMM, to be defined in Section 6. In our definition, we require (8) to hold for all choices of y_1 and y_2 , while the ambiguity measure of KMM is equivalent to requiring it to hold only when $I_1 = I_2$ and there is an x such that $u'_n(x) = y_n$ for each n . To compare two marginal utilities at a common consumption level, it is, of course, necessary that $I_1 = I_2$, but our definition of the even-more-risk-averse-than relation is applicable even when this condition is not met.

In this example, v_n has decreasing or increasing relative risk aversion, depending on whether γ_n is greater or smaller than one. Indeed, by (11),

$$r(x, v_n) = \gamma_n + \alpha_n x^{1-\gamma_n}. \quad (13)$$

Thus, if $\gamma_n > 1$, then $r(\cdot, v_n)$ is strictly decreasing, while if $\gamma_n < 1$, then it is strictly increasing. This is a rather unexpected consequence of introducing ambiguity aversion by way of φ_n of the form (1). On the one hand, the decision maker's constant coefficient of relative risk aversion over purely risky consumption plans can be measured, say, by the fraction of the total wealth he invests into the asset with purely risky returns. On the other hand, whether he exhibits increase or decreasing relative risk aversion over purely ambiguous consumption plans (second-order acts, according to the terminology of KMM) can be determined, say, by whether he would increase the fraction of the total wealth he invests into the assets with purely ambiguous returns as the total wealth increases. These two attitudes towards risk and ambiguity should better be disentangled in models of any quantitative analysis, but, in this specification, a restriction on one automatically implies a restriction on the other.

The use of functions (1) of constant absolute risk aversion as the function φ_n that transforms u_n to v_n was also suggested by Ju and Miao (2012, pages 566–567). The justification for this is that if we take φ_n to be a function (2) of constant relative risk aversion, then $v_n = \varphi_n \circ u_n$ is not well defined when u_n has constant coefficient γ_n of relative risk aversion greater than one (because, then, $\text{Ran } u_n = -\mathbf{R}_{++}$). This problem can be circumvented if we specify the function ψ_n that transforms u'_n to v'_n to be any plausible form, such as (9), because $\text{Ran } u'_n = \mathbf{R}_{++}$ regardless of the values of γ_n .

5 Behavioral conditions

In this section, we obtain an equivalent behavioral condition of the even-more-risk-averse-than relation. We start with some definitions.

Definition 2 Suppose that four expected utility functions $u_1 : I_1 \rightarrow \mathbf{R}$, $v_1 : I_1 \rightarrow \mathbf{R}$, $u_2 : I_2 \rightarrow \mathbf{R}$, and $v_2 : I_2 \rightarrow \mathbf{R}$ satisfy the basic conditions. We say that v_1 is more risk averse than u_1 in behavior at least to the same extent as v_2 is more risk averse than u_2 , if for each n and for every $x_n \in I_n$, every cumulative distribution functions F_n , and every sufficiently small $\varepsilon_n > 0$, if the support of F_n is included in $[x_n - \varepsilon_n, x_n + \varepsilon_n]$ and its mean is equal to x_n , then there are a $q_n > 0$ and a $\tau > 0$ such that

$$\begin{aligned} u_1(x_1 - q_1) &\leq \int u_1(z) dF_1(z), \\ v_1(x_1 - \tau q_1) &\geq \int v_1(z) dF_1(z), \\ u_2(x_2 - q_2) &\geq \int u_2(z) dF_2(z), \\ v_2(x_2 - \tau q_2) &\leq \int v_2(z) dF_2(z). \end{aligned}$$

We then write $(u_1, v_1) \blacktriangleright^\circ (u_2, v_2)$. If, in addition, these four inequalities hold as strict inequalities, then we say that v_1 is more risk averse than u_1 in behavior to a greater extent than v_2 is more risk averse than u_2 and write $(u_1, v_1) \blacktriangleright^\circ (u_2, v_2)$.

The first inequality in the definition of $\blacktriangleright^\circ$ tells us that the certainty premium under u_1 is smaller than or equal to q_1 , while the second inequality tells us that the certainty premium under v_1 is greater than or equal to τq_1 . Hence, the proportional change in the certainty premium induced by the change from u_1 to v_1 is greater than or equal to τ . The third inequality tells us that the certainty premium under u_2 is greater than or equal to q_2 , while the fourth inequality tells us that the certainty premium under v_2 is smaller than or equal to τq_2 . Hence, the proportional change in the certainty premium induced by the change from u_2 to v_2 is smaller than or equal to τ . Thus, the proportional change in certainty premium is greater or equal when changing from u_1 to v_1 than when changing from u_2 to v_2 . The definition of $\blacktriangleright^\circ$ requires that this be true for every small risk, regardless of the consumption levels at which the certainty premiums are measured. The inequalities in the definition of $\blacktriangleright^\circ$ are different from those in the definition of \blacktriangleright only in that the strict inequalities are all replaced by the weak inequalities. Whenever necessary, we shall refer to the binary relations $\blacktriangleright^\circ$ and \blacktriangleright as the behavioral even-more-risk-averse-than relation.

The following theorem is concerned with the four binary relations \blacktriangleright , \triangleright , $\blacktriangleright^\circ$, and \triangleright° .

Theorem 1 $\triangleright \subseteq \triangleright^\circ \subseteq \blacktriangleright^\circ \subseteq \blacktriangleright$.

This theorem means that for four expected utility functions $u_1 : I_1 \rightarrow \mathbf{R}$, $v_1 : I_1 \rightarrow \mathbf{R}$, $u_2 : I_2 \rightarrow \mathbf{R}$, and $v_2 : I_2 \rightarrow \mathbf{R}$ satisfying the basic conditions. if $(u_1, v_1) \triangleright (u_2, v_2)$, then $(u_1, v_1) \triangleright^\circ (u_2, v_2)$; if $(u_1, v_1) \triangleright^\circ (u_2, v_2)$, then $(u_1, v_1) \blacktriangleright^\circ (u_2, v_2)$; and if $(u_1, v_1) \blacktriangleright^\circ (u_2, v_2)$, then $(u_1, v_1) \blacktriangleright (u_2, v_2)$. It shows that the even-more-risk-averse-than relation is concerned with

the proportional change in the certainty premiums (the difference between the mean of the random prospect and its certainty equivalent) caused by a change in expected utility functions.

Putting the four conditions together and ignoring the difference between strict and weak relations, we can conclude that the even-more-risk-averse-than relation can be detected by a reversal of choices between the deterministic and random consumption plans by some common proportional change in the certainty premiums. Note that the conditions in this theorem are also applicable to define the even-more-risk-averse-than relation between two pairs of preference relations over cumulative distributions functions that may not be represented by expected utility functions. To formalize this last statement, let \succsim_1^1 and \succsim_1^2 be preference relations defined on a set of cumulative distribution functions on I_1 , and \succsim_2^1 and \succsim_2^2 be preference relations defined on a set of cumulative distribution functions on I_2 . Imagining that \succsim_1^1 , \succsim_1^2 , \succsim_2^1 , and \succsim_2^2 are represented by the expected utility functions u_1 , v_1 , u_2 , and v_2 and letting \succ_1^1 , \succ_1^2 , \succ_2^1 , and \succ_2^2 be their asymmetric (strict) parts, we can rewrite the inequalities in the definition of $\overset{\circ}{\triangleright}$ as

$$\begin{aligned} F_1 &\succ_1^1 \mathbf{1}_{[x_1-q_1, \infty)}, \\ \mathbf{1}_{[x_1-\tau q_1, \infty)} &\succ_1^2 F_1, \\ \mathbf{1}_{[x_2-q_2, \infty)} &\succ_2^1 F_2, \\ F_2 &\succ_2^2 \mathbf{1}_{[x_2-\tau q_2, \infty)}, \end{aligned}$$

where, for every x , $\mathbf{1}_{[x, \infty)}$ denotes the (degenerate) cumulative distribution function taking value 1 on $[x, \infty)$ and 0 on $(-\infty, x)$. The inequalities in the definition of $\overset{\circ}{\triangleright}$ can be obtained by replacing the \succ_n^i by the \succsim_n^i . These conditions can be used to as the definition of the statement that \succsim_1^2 is more risk-averse than \succsim_1^1 to a greater extent than (or at least to the same extent as) \succsim_2^2 is more risk-averse than \succsim_2^1 , even for preference relations that cannot be represented by expected utility functions.

6 Application to the utility functions of KMM

As we stated in the introduction, this study is motivated by the smooth ambiguity model of KMM. In this section, we show how our definition can be used to compare two ambiguity-averse utility functions in the model.

6.1 Setup

Let S be the state space, which represents the uncertainty that the decision maker is faced with. Denote by D the set of all probability measures S . Denote by C the set of

all functions of S into I .⁵ Suppose that two expected utility functions $u : I \rightarrow \mathbf{R}$ and $v : I \rightarrow \mathbf{R}$ satisfy the basic conditions. Let μ be a probability measure on D . Define a utility function $W : C \rightarrow \mathbf{R}$ by letting

$$W(c) = \int_D v \left(u^{-1} \left(\int_S u(c(s)) d\pi(s) \right) \right) d\mu(\pi) \quad (14)$$

for every $c \in C$. This nested expected utility function is the functional form that KMM axiomatized. Write $\varphi = v \circ u^{-1}$, then

$$W(c) = \int_D \varphi \left(\int_S u(c(s)) d\pi(s) \right) d\mu(\pi). \quad (15)$$

This shows that the decision maker is averse to the uncertainty that he perceives in the expected utilities calculated by various probability measures $\pi \in D$ if and only if φ is concave, that is, v is more risk-averse than u . The probability measure μ represents his subjective assessment of this uncertainty, which KMM termed as the second-order belief.

We see in (14) that if $\int_S u(c(s)) d\pi(s)$ is independent of π , then, writing $x = u^{-1} \left(\int_S u(c(s)) d\pi(s) \right) \in I$, we obtain $W(c) = u(x)$. This means that if the conditional certainty equivalent of c given a probability measure $\pi \in \text{supp } \mu$ is, in fact, independent of π , then the utility function W is determined by the inner utility function u (as it determines the conditional certainty equivalents), and the outer expected utility function v is irrelevant as it only monotonically transforms the certainty equivalents. Thus, u can be interpreted as representing the attitudes towards pure risk. On the other hand, suppose that c is constant almost surely with respect to every $\pi \in \text{supp } \mu$, but the constant that c takes almost surely depends on π , then denote the value by $c(\pi)$. Then, $W(c) = \int_D v(c(\pi)) d\mu(\pi)$. This means that $W(c)$ is determined only by the outer utility function v , and the inner utility u is irrelevant as we take the certainty equivalents in the calculation for $W(c)$. Thus, v can be interpreted as representing the attitudes towards the uncertainty that the decision maker perceives in the expected utility levels.

To give a new definition of the more-ambiguity-averse-than relation for the utility functions of KMM and compare it with the definition KMM gave (Definition 5), we impose the same restrictions on the state space as they did. Let $S = \Omega \times [0, 1]$, where Ω is a measurable space and $[0, 1]$ is the closed unit interval endowed with the Lebesgue measure λ . It is interpreted as an objective probability, and, as such, all the probabilities that the decision maker may conceive of on the state space S have the common marginal distribution λ on $[0, 1]$.⁶ We assume that S contains at least two elements. By an abuse of notation, we also denote a probability distribution on Ω by π , the set of all probability

⁵To make sure the utility function is indeed well defined, we need to impose some additional conditions on S and C . To simplify the exposition, we omit them.

⁶We have taken $[0, 1]$ and the Lebesgue measure λ as the objective probability measure to guarantee that any distribution of consumption levels can be represented as a random variable on $[0, 1]$.

measures on Ω by D , the second-order belief on D by μ . By Fubini's theorem, we can then rewrite (14) as

$$\begin{aligned}
W(c) &= \int_D v \left(u^{-1} \left(\int_{\Omega \times [0,1]} u(c(\omega, \xi)) d(\pi \otimes \lambda)(\omega, \xi) \right) \right) d\mu(\pi) \\
&= \int_D v \left(u^{-1} \left(\int_{\Omega} \left(\int_{[0,1]} u(c(\omega, \xi)) d\lambda(\xi) \right) d\pi(\omega) \right) \right) d\mu(\pi), \\
&= \int_D v \left(u^{-1} \left(\int_{[0,1]} \left(\int_{\Omega} u(c(\omega, \xi)) d\pi(\omega) \right) d\lambda(\xi) \right) \right) d\mu(\pi).
\end{aligned} \tag{16}$$

6.2 An alternative more-ambiguity-averse-than relation

Let W_1 and W_2 be two KMM utility functions defined on the same state space $S = \Omega \times [0, 1]$ and determined by two triples (u_1, v_1, μ_1) and (u_2, v_2, μ_2) via (16). Denote by I_n the common domain of u_n and v_n . Denote by C_n the set of all $c_n : S \rightarrow I_n$. The following is a simplified version of the more-ambiguity-averse-than relation of KMM.

Definition 3 (KMM) Assume that $I_1 = I_2$ and $\mu_1 = \mu_2$. Write C for C_n . We say that W_1 is *at least as ambiguity-averse as* W_2 if, for every $c \in C$ and every $d \in C$, $W_2(c) \geq W_2(d)$ whenever $d(\omega, \xi)$ is independent of $\omega \in \Omega$ and $W_1(c) \geq W_1(d)$.

In the first part of this definition, we assume that the two utility functions share the same domain of consumption levels and the same second-order belief. The assumption of common domain is needed as this definition is concerned with the rankings by W_1 and W_2 over common consumption plans c and d . The assumption of the common second-order belief is imposed to exclude the possibility that the difference in ambiguity attitudes arises from a difference in second-order beliefs. The integral part of their definition is in the second part of this definition. It requires that for two consumption plans c and d , if d is unambiguous and it is at most as desirable as another, possibly ambiguous, consumption plan c for W_1 , then d should also be at most as desirable as c for W_2 . This definition formalizes the idea, putting the discrepancy between weak and strict preferences aside, that if the unambiguous consumption plan is inferior for the more ambiguity-averse utility function W_1 , it should also be so for the less ambiguity-averse utility function W_2 .

The original definition by KMM is more intricate than Definition 3. They gave the more-ambiguity-averse-than relation over the family of pairs of preference relation on the set C and preference relations on the set of fictitious consumption plans (termed by KMM as second-order acts) contingent on probability measures π on Ω ,⁷ where the family is constructed by indexing the pairs by the supports of second-order beliefs in

⁷In fact, KMM axiomatized the functional form (14) in terms of a pair of a preference relation on C and a preference relation on the set of second-order acts, rather than just in terms of a preference relation on C . Seo (2009) axiomatized the functional form that extends (14) by dispensing with the preference relation on the set of second-order acts and introducing three-stage, rather than two-stage, lotteries.

D ; and they defined one pair as being more ambiguity-averse than another if the same rankings between an unambiguous consumption plan d and a possibly ambiguous plan c holds as in Definition 3 for all supports of second-order beliefs. In contrast, Definition 3 does not involve any preference relation on the set of second-order acts, and deals with a single preference relation rather than a family of preference relations. KMM's fully-fledged definition is important, especially when we interpret numerical results on KMM utility functions, because it makes explicit the otherwise implicit assumption that a decision maker's attitudes towards risk (represented by the inner utility function u) and ambiguity (represented by the outer utility function v) should travel with him across different settings (represented by the supports of second-order beliefs).⁸ Yet, in the subsequent analysis, we use Definition 3 because this simplified version is sufficient to illustrate the difference in the definition of a more-ambiguity-averse-than relation between KMM and this paper.

Theorem 2 of KMM shows that for W_1 and W_2 defined through (16) with $I_1 = I_2$ and $\mu_1 = \mu_2$, W_1 is at least as ambiguity averse as W_2 if and only if u_1 is an affine transformation of u_2 and v_1 is a concave transformation of v_2 . The affinity between u_1 and u_2 follows from the fact that in Definition 3 (a simplified version of the definition of KMM), the consumption plan c may be unambiguous as well. In fact, by restricting the definition to the case where both d and c are unambiguous, we can see that W_1 and W_2 must agree on the ranking between unambiguous consumption plans whenever one is more ambiguity-averse than the other in the sense of Definition 3. But this is equivalent to saying that u_1 is an affine transformation of u_2 .

This consequence of the more-ambiguity-averse-than relation is somewhat unfortunate, because it significantly limits the scope within which we can compare various quantitative results on ambiguity attitudes in the literature. To see this point, imagine that given a set of data on portfolio choices, we have estimated ambiguity-averse utility functions W_1 and W_2 for two groups of investors that are formed on the basis of some observable characteristics, such as age, gender, and occupation. We would then like to know to what extent the difference in ambiguity attitudes can account for the difference in portfolio choices between the two groups. The natural course of action would be to compare the estimated φ_1 and φ_2 . However, Definition 3 would provide no sound theoretical foundation for such a comparison if the estimates of u_1 and u_2 were different.

Our definition of a more-ambiguity-averse-than relation does not suffer from this deficiency. Unlike Definition 3, our definition neither assume that $\mu_1 = \mu_2$ nor imply that u_1 is an affine transformation of u_2 . Denote by ι the function defined on S that constantly takes value one. Our definition can then be stated as follows.

⁸Assumption 4 of KMM requires these attitudes to be separable from the settings. But whether such a separation is possible is a contentious issue, as can be seen in the discussions of Epstein (2010) and Klibanoff, Marinacci, and Mukerji (2012).

Definition 4 We say that W_1 is *at least as ambiguity-averse as* W_2 if for each n and for every $x_n \in I_n$, every sufficiently small $\varepsilon_n > 0$, every $c_n \in C_n$, and every $d_n \in C_n$, if they satisfy the first two of the following three conditions, then there are a $q_n > 0$ and a $\tau > 0$ that satisfy the last one.

1. $d_n(\omega, \xi)$ does not depend on $\omega \in \Omega$. We thus write $d_n(\xi)$ and regard $d_n : [0, 1] \rightarrow I_n$;
2. Define $e_n : D \rightarrow I_n$ by letting

$$e_n(\pi) = u_n^{-1} \left(\int_{\Omega} u_n(c_n(\omega, \xi)) d(\pi \otimes \lambda)(\omega, \xi) \right) \quad (17)$$

for every $\pi \in D$. Then the distribution of e_n , $\mu_n \circ e_n^{-1}$, coincides with the distribution of d_n , $\lambda \circ d_n^{-1}$. Moreover, their support is included in $[x_n - \varepsilon_n, x_n + \varepsilon_n]$ and their mean is equal to x_n ,

- 3.

$$\begin{aligned} W_1((x_1 - q_1)\iota) &\leq W_1(d_1), \\ W_1((x_1 - \tau q_1)\iota) &\geq W_1(c_1), \\ W_2((x_2 - q_2)\iota) &\geq W_2(d_2), \\ W_2((x_2 - \tau q_2)\iota) &\leq W_2(c_2). \end{aligned}$$

We then write $W_1 \widehat{\triangleright} W_2$. If, in addition, the inequalities in the last condition hold as strict inequalities, then we say that W_1 is *more ambiguity-averse than* W_2 , and write $W_1 \widehat{\triangleright} W_2$.

This definition compares the preference between a deterministic consumption plan $(x_n - q_n)\iota$ and an unambiguous consumption plan d_n , with the preference between a deterministic consumption plan $(x_n - \tau q_n)\iota$ and a purely ambiguous consumption plan c_n . To be more precise, by (16) and the change-of-variable formula,

$$W_n(d_n) = v_n \left(u_n^{-1} \left(\int_{[0,1]} u_n(d_n(\xi)) d\lambda(\xi) \right) \right) = v_n \left(u_n^{-1} \left(\int_{I_n} u_n(x) d(\lambda \circ d_n^{-1})(x) \right) \right).$$

Since v_n is monotone, the ranking between d_n and $(x_n - q_n)\iota$ can be reduced to the ranking between the distribution $\lambda \circ d_n^{-1}$ and the deterministic consumption level $x_n - q_n$ by the inner expected utility function u_n . By (17) and the change-of-variable formula,

$$W_n(c_n) = \int_D v_n(e_n(\pi)) d\mu_n(\pi) = \int_{I_n} v_n(x) d(\mu_n \circ e_n^{-1})(x).$$

Thus, the ranking between c_n and $(x_n - \tau q_n)\iota$ can be reduced to the ranking between the distribution $\mu_n \circ e_n^{-1}$ and the deterministic consumption level $x_n - \tau q_n$ by the outer

expected utility function v_n . Since $\lambda \circ d_n^{-1} = \mu_n \circ e_n^{-1}$, the two rankings differ only in the expected utility function with respect to which the two (random and deterministic) consumption plans are ranked. Following the terminology of KMM, we shall refer to e as the second-order act associated with c_n .

This sort of comparison between two rankings was envisaged by Definition 2, but the comparison in Definition 4 is different from it in an important respect: While the alternatives, $x_n - q_n$, $x_n - \tau q_n$, and F_n , in Definition 2 can be set up without knowing the utility functions (u_n and v_n), to set up the alternatives, c_n , d_n , $x_n - q_n$, and $x_n - \tau q_n$, in Definition 4, we need to know the inner utility function u_n and the second-order belief μ_n because c_n and d_n must satisfy $\lambda \circ d_n^{-1} = \mu_n \circ e_n^{-1}$, where the fictitious consumption plan e_n in (17) depends on the inner utility function u_n . In this sense, the more-ambiguity-averse-than relation of Definition 4 is better suited to the case where the attitudes towards risk (represented by the inner utility function u_n) and the second-order beliefs μ_n are already known.⁹

The next two theorems show how our more-ambiguity-averse-than relation is related to the even-more-risk-averse-than relation in behavior, which we defined in Definition 2. They give an easy way to check whether a KMM utility function is more ambiguity-averse than another. The first one is simple.

Theorem 2 *Define two utility function $W_1 : C_1 \rightarrow \mathbf{R}$ and $W_2 : C_2 \rightarrow \mathbf{R}$ on the same state space $S = \Omega \times [0, 1]$, with $[0, 1]$ endowed with the Lebesgue measure, by two triples (u_1, v_1, μ_1) and (u_2, v_2, μ_2) via (16). Then*

1. *If $(u_1, v_1) \overset{\circ}{\triangleright} (u_2, v_2)$, then $W_1 \widehat{\triangleright} W_2$.*
2. *If $(u_1, v_1) \overset{\bullet}{\triangleright} (u_2, v_2)$, then $W_1 \widehat{\triangleright} W_2$.*

This theorem shows, roughly, that the even-more-risk-averse-than relation, in behavior, implies the more-ambiguity-averse-than relation. The second theorem is a partial, but not the full, converse of the first. To see why the full converse cannot be obtained, consider the case where the second-order belief μ_n is concentrated on a single probability measure on Ω , then, for every (c_n, d_n) satisfying the conditions of Definition 4, the distribution $\mu_n \circ e_n^{-1}$ is degenerate on a single consumption level and, thus, the inequalities

⁹Thus, if we were to conduct experiments to infer and compare two KMM utility functions W_1 and W_2 , we should do so in two stages under the assumption that we know that the two second-order beliefs μ_1 and μ_2 are the same and, in addition, what the common second-order belief is. In the first stage of experiments, we only use unambiguous consumption plans to infer the inner utility functions u_n . In the second stage, based on the inner utility function u_n inferred in the first stage and the common second-order belief posited at the beginning, we set up c_n and d_n to satisfy condition 2 of Definition 4, and choose x_n , q_n , and τ see if it is possible to generate a preference reversals between the two utility functions W_1 and W_2 when c_n and d_n are compared to $x_n - \tau q_n$ and $x_n - q_n$. The assumption that the two second-order beliefs are known and identical would be unnecessary if it were possible to set up fictitious consumption plans contingent on π (second-order acts) in experiments to infer μ_n and v_n . KMM argued that it may well be possible to do so to justify their Assumption 2, which is one of the axioms for the functional form (14).

in Condition 3 of Definition 4 cannot have any implication on the outer utility function v_n . We thus impose the following condition on each KMM utility function W defined on the state space $S = \Omega \times [0, 1]$ and determined by a triple (u, v, μ) via (16).¹⁰

Definition 5 We say that the second-order belief μ is *rich* if for every Borel probability measure ν on the set I of consumption levels with bounded support, there is a function $c : \Omega \times [0, 1] \rightarrow I$ such that $\nu = \mu \circ e^{-1}$, where $e : D \rightarrow I$ is defined by

$$e(\pi) = u^{-1} \left(\int_{\Omega} u(c(\omega, \xi)) d(\pi \otimes \lambda)(\omega, \xi) \right) \quad (18)$$

for every $\pi \in D$.

The aim of this condition is self-evident. It is to guarantee that every distribution on I with a bounded support can be attained as a distribution of some associated second-order act. We can then compare the risk aversion of the inner utility function u_n and of the outer utility function v_n through the distribution of associated second-order acts. The converse of Theorem 2 under the richness condition can be stated as follows.

Theorem 3 Define two utility function $W_1 : C_1 \rightarrow \mathbf{R}$ and $W_2 : C_2 \rightarrow \mathbf{R}$ on the same state space $S = \Omega \times [0, 1]$, with $[0, 1]$ endowed with the Lebesgue measure, by two triples (u_1, v_1, μ_1) and (u_2, v_2, μ_2) via (16). Suppose that μ_1 and μ_2 are rich. Then

1. If $W_1 \widehat{\triangleright} W_2$, then $(u_1, v_1) \overset{\circ}{\triangleright} (u_2, v_2)$.
2. If $W_1 \widehat{\blacktriangleright} W_2$, then $(u_1, v_1) \overset{\circ}{\blacktriangleright} (u_2, v_2)$.

A drawback of this theorem is that the validity of the richness condition may be difficult to check. Below, we give a couple of examples that does and does not satisfy the condition to illustrate the fine line between rich and non-rich second-order beliefs.¹¹

Example 7 Let $\Omega = [0, 1]$ and take the support of the second-order belief to be $\{\delta_{\omega} \mid \omega \in \Omega\}$, where δ_{ω} is the (degenerate) probability measure concentrated on ω . With the parametrization of the (first-order) beliefs $\omega \mapsto \delta_{\omega}$ on Ω , take the second-order belief to be the uniform distribution on Ω .

For every $c : \Omega \rightarrow I$, if we define e by (18), then

$$e(\delta_{\omega}) = u^{-1} \left(\int_{\Omega} u(c(\sigma)) d\delta_{\omega}(\sigma) \right) = u^{-1}(u(c(\omega))) = c(\omega)$$

¹⁰The richness condition can be stated in the more general form (14) of KMM utility functions.

¹¹The apparent difference between the two examples lies in the supports of the first-order beliefs in the support of the second-order belief. In the first example, where the richness condition is satisfied, the supports of the first-order beliefs in the support of the second-order belief are mutually disjoint. In the second example, where the richness condition is not satisfied, these supports are identical almost surely with respect to the second-order belief. The importance of disjoint supports of first-order beliefs is mentioned in Footnote 5 of KMM and used the proof of their Proposition 1.

for every $\omega \in \Omega$. Thus, for every Borel probability measure ν on I , if we denote its cumulative distribution function by F and let c be its generalized inverse in the sense of Embrecht and Hofert (2014), then $\nu = \mu \circ e^{-1}$. Thus, the uniform distribution on the parameterized family $(\delta_\omega)_{\omega \in \Omega}$ of probabilities on Ω satisfies the richness condition.

Example 8 Let $\Omega = [0, 1]$. Let π^0 and π^1 be two probability measures on Ω . For each $\theta \in [0, 1]$, define $\pi^\theta = (1 - \theta)\pi^0 + \theta\pi^1$. Then define the second-order belief μ as the uniform distribution on the parameterized family $(\pi^\theta)_{\theta \in [0, 1]}$. Thus, just like Example 7, the second-order belief is a family of probability measures parameterized by $[0, 1]$. We will, however, see that the richness condition is violated in this example.

For every $c : \Omega \times [0, 1] \rightarrow I$, if we define e by (18), then

$$\begin{aligned} e(\pi^\theta) &= u^{-1} \left(\int_{\Omega \times [0, 1]} u(c(\omega, \xi)) \, d(\pi^\theta \otimes \lambda)(\omega, \xi) \right) \\ &= u^{-1} \left((1 - \theta) \int_{\Omega \times [0, 1]} u(c(\omega, \xi)) \, d(\pi^0 \otimes \lambda)(\omega, \xi) + \theta \int_{\Omega \times [0, 1]} u(c(\omega, \xi)) \, d(\pi^1 \otimes \lambda)(\omega, \xi) \right). \end{aligned}$$

for every $\theta \in [0, 1]$. Write $\hat{e}(\theta) = e(\pi^\theta)$ and $y^\theta = \int_{\Omega \times [0, 1]} u(c(\omega, \xi)) \, d(\pi^\theta \otimes \lambda)(\omega, \xi)$, then the above equality can be more succinctly written as

$$\hat{e}(\theta) = u^{-1}((1 - \theta)y^0 + \theta y^1).$$

If $y^0 = y^1$, then \hat{e} is a constant function and $\mu \circ c^{-1}$ is a degenerate probability measure on I . Suppose that $y^0 < y^1$. Then \hat{e} is a strictly increasing function that maps $[0, 1]$ onto $[u^{-1}(y^0), u^{-1}(y^1)]$. Let $F : [u^{-1}(y^0), u^{-1}(y^1)] \rightarrow [0, 1]$ be the cumulative distribution function of $\mu \circ e^{-1}$. Then,

$$F(z) = \frac{u(z) - y^0}{y^1 - y^0} \tag{19}$$

for every $z \in [u^{-1}(y^0), u^{-1}(y^1)]$. Indeed, since \hat{e} is strictly increasing and continuous and the second-order belief μ as the uniform distribution on the parameterized family $(\pi^\theta)_{\theta \in [0, 1]}$, F coincides with the inverse $\hat{e}^{-1} : [0, 1] \rightarrow [u^{-1}(y^0), u^{-1}(y^1)]$. But \hat{e}^{-1} coincides with the right-hand side of (19).

The equality (19) implies that every probability measure $\mu \circ e^{-1}$ that can be attained as in Condition 2 of Definition 4 has a cumulative distribution function that coincides with a truncated and normalized inner utility function u . Therefore, the second-order belief μ is not rich.

6.3 Relevance of the alternative definition to the literature

The ambiguity aversion coefficients in the smooth model were inferred or estimated from experimental evidence or asset market data, borrowed from earlier works, or quoted

as a consensus in the profession, by Halevy (2007), Ju and Miao (2011), Chen, Ju, and Miao (2014), Jahan-Parvar and Liu (2014), Thimme and Vöckert (2015), Gallant, Jahan-Parvar, and Liu (forthcoming), Altug, Collard, Çakmakli, Mukerji, and Özsöylev (2018), and Hara and Honda (2018). These studies used or obtained different (constant) coefficients of ambiguity aversion, which corresponds to η_n/γ_n in Examples 1, 2, and 3. It is impossible to conclude that the decision maker with a higher estimated coefficient of ambiguity aversion is more ambiguity-averse in the sense of KMM (Definition 3 of this paper), because these studies involve different risk aversion coefficients (which correspond to γ_n in Examples 1, 2, and 3).

To see how our more-ambiguity-averse-than relation can be fit in these studies, let's take up Chen, Ju, and Miao (2014), who studied the optimal portfolio choice problem of an investor who has a utility function of Hayashi and Miao (2011), which not only extends utility functions of KMM to a dynamic setting but also generalizes recursive utility functions of Epstein and Zin (1989), thereby allowing for the three-way separation between risk aversion, ambiguity aversion, and intertemporal elasticity of substitution. Tables 1 and 3 of their paper list up various configurations of the coefficients of relative risk aversion of the inner expected utility function u_n , which is denoted by γ_n in Example 2, and the coefficients of relative risk aversion of the outer expected utility function v_n , which is denoted by η in Example 2. In Table 1, for each pair $(\gamma, \eta) \in \{0.5, 2, 5, 10, 15\} \times \{40, 50, 60, 70, 80, 90, 100, 110\}$, they presented the ambiguity premium, defined as the difference between the certainty equivalents of a purely risky act and a purely ambiguous (second-order) act. When γ is fixed, say, at 2, increasing η from 50 to 100 leads to a more ambiguity-averse investor in the sense of KMM. But, the investor is not more ambiguity-averse when $(\gamma, \eta) = (2, 40)$ than when $(\gamma, \eta) = (5, 90)$ or the other way around, because the coefficients γ of relative risk aversion are different between the two pairs. Yet, according to our definition, the investor is deemed as more ambiguity-averse when $(\gamma, \eta) = (2, 40)$ than when $(\gamma, \eta) = (5, 90)$, because $40/2 = 20 > 18 = 90/5$. Nonetheless, the ambiguity premium is lower when $(\gamma, \eta) = (2, 40)$ than when $(\gamma, \eta) = (5, 90)$. This is due to the difference in the way the premiums are defined. In this paper, the even-more-risk-averse-than relation is defined according to the ratio of the certainty premiums (the differences between the expected reward and the certainty equivalents of a lottery) with respect to γ and with respect to η , while their “ambiguity premium” is equal to the difference between the certainty two premiums. The latter is more pronounced when the coefficient γ of relative risk aversion is larger.

In Table 3, Chen, Ju, and Miao (2014) presented the optimal fraction of investment into the stock (the other asset being riskless in their model). The pairs (γ, η) that they

used are

$$\begin{aligned} &(2, 2), (2, 60), (2, 80), (2, 100), \\ &(5, 5), (5, 60), (5, 80), (5, 100), \\ &(10, 10), (10, 60), (10, 80), (10, 100). \end{aligned}$$

They observed that for a fixed γ , increasing η leads to a lower fraction of investment into the stock. The definition of KMM covers this case, but does not tell us whether the investor is more ambiguity-averse when $(\gamma, \eta) = (5, 60)$ than when $(\gamma, \eta) = (10, 100)$. According to our more-ambiguity-averse-than relation, the investor is deemed as more ambiguity-averse when $(\gamma, \eta) = (5, 60)$ than when $(\gamma, \eta) = (10, 100)$, because $60/5 = 12 > 10 = 100/10$. They found that the optimal fraction of investment in the stock is higher when $(\gamma, \eta) = (5, 60)$ than when $(\gamma, \eta) = (10, 100)$. This is consistent with their observation that the coefficient γ of relative risk aversion for the inner expected utility function u has larger effects on the optimal fraction of investment into the stock than the coefficient η of relative risk aversion for the inner expected utility function v .

Another instance in which the scope of comparison of ambiguity aversion is enhanced by our definition of ambiguity aversion is Hara and Honda (2018) versus the other contributions mentioned at the beginning of this subsection. Hara and Honda (2018) assumed constant absolute risk aversion as in Example 1, and the others assumed constant relative risk aversion as in Example 2. The two are not comparable according the more-ambiguity-averse-than relation of KMM. Moreover, the concavity of the functions φ_n that transform u_n to v_n are not comparable, because the domain of φ_n is \mathbf{R} in the case of constant absolute risk aversion, while it is \mathbf{R}_{++} or $-\mathbf{R}_{++}$ in the case of constant relative risk aversion. Yet, as mentioned right after Example 3, our definition of the more-ambiguity-averse-than relation allows us to compare the ambiguity aversion between the two cases on sound economic ground.

Hara and Honda (2018) found that for the representative consumer, who holds the stock market index (a proxy of the market portfolio), η_n/γ_n must be at least 9.25 and may well be higher. This figure is much higher than the figures obtained in many other works for the representative consumer. For example, Ju and Miao's (2012) calibration shows that η_n/γ_n is around 4.43. It is worthwhile to attempt to explain where the difference is from, but without our definition of the more-ambiguity-averse-than relation, this question would have been ill-posed.¹²

¹²It is tempting to speculate that the difference arises from the difference in settings, because Hara and Honda (2018) considered a static model with multiple risky assets, while the others considered a dynamic model with a single risky asset. But such a speculation may not be consistent with the basic tenet of KMM utility functions, explained in Footnote 8 as well, whereby the second-order belief may depend on settings but the ambiguity aversion must not.

7 Conclusion

Given two pairs of expected utility functions, we have formalized the statement that one expected utility function is more risk-averse than the other in the first pair to a greater extent than in the second pair. To do so, we used the elasticity of the function that transforms the derivatives of one expected utility function to the derivatives of the other.

As was seen in (5), (6), and (8), when we compared the elasticities of ψ_1 and ψ_2 , we require the elasticity of ψ_1 is higher than the elasticity of ψ_2 , regardless of the choices of the marginal utilities, y_1 and y_2 , of the expected utility functions. This makes our definition of the even-more-risk-averse-than relation rather stringent, and two pairs of expected utility function may not be comparable according to the relation. One might be led to think that it would be more practical to define the even-more-risk-averse-than relation by choosing the marginal utilities, y_1 and y_2 , to be equal. There are two reasons why this attempt is unlikely to be successful. First, since the level of marginal utilities may be changed by a scalar multiplication to an expected utility function (which does not change the risk attitudes it represents), choosing the same level of marginal utilities for two expected utility functions has, in general, no economic meaning. Second, as we did in our explanation after Example 6, it might make sense to take y_1 and y_2 to be the marginal utilities at a common consumption level. This is possible, however, only if the expected utility functions of the two pairs have the same domain. This would restrict the applicability of our definition, as it would exclude cases such as Example 3. Yet, when the domains are the same, it might be possible to give a less stringent, more practical definition of the even-more-risk-averse-than relation. Exploring the implication of this alteration can be a direction of future research.

Another direction of future research is to find a narrower class of cumulative distribution functions F_n that can be used in the definitions of the behavioral even-more-risk-averse-than relation (Definition 2) and the more-ambiguity-averse-than relation (Definition 4). In these definitions, arbitrary cumulative distribution functions with sufficiently small supports are used, but because of this arbitrariness, we needed to impose a rather stringent richness condition (Definition 5) on second-order beliefs to establish the equivalence between the behavioral even-more-risk-averse-than relation and the more-ambiguity-averse-than relation. It might be sufficient to consider just cumulative distribution functions for binary distributions, but a careful analysis is in order.

Finally, the most important direction of future research is to extend the more-ambiguity-averse-than relation (Definition 4) to other types of ambiguity-averse utility functions. As explained in Footnote 8 of KMM, two utility functions that are comparable with respect to the more-ambiguity-averse-than relations employed for other classes of ambiguity-averse utility functions, such as α -MEU functions, must also exhibit the same preference over purely risky consumption plans. This property, again, significantly narrows down the

scope of comparison of ambiguity attitudes. Finding a general definition of the more-ambiguity-averse-than relation that covers these classes is imperative to increase the usefulness of ambiguity-averse utility functions in numerical and empirical analysis.

A Lemmas and Proofs

Proof of Proposition 2 By differentiating both sides of $v'(x) = \psi(u'(x))$ with respect to x , we obtain $v''(x) = \psi'(u'(x))u''(x)$. By dividing both sides of this equality by both sides of the previous one, we obtain

$$a(x, v) = -\frac{\psi'(u'(x))}{\psi(u'(x))}u''(x).$$

By substituting $u''(x)$ by $-u'(x)a(x, u)$, we obtain (3) .

///

We say that a cumulative distribution function on \mathbf{R} satisfies the basic conditions its mean is equal to zero and its variance is finite and greater than zero.

For each distribution function F with zero mean, write $V(F) = \int_{\mathbf{R}} z^2 dz$. It is equal to the variance of any random variable whose distribution coincides with F . The following fact is well known. We omit the proof.

Lemma 1 *Let $u : I \rightarrow \mathbf{R}$ be an expected utility functions satisfying the basic conditions. For each $x \in I$ and each cumulative distribution function F satisfying the basic conditions, define $p(x, F, u)$ by*

$$u(x - p(x, F, u)) = \int_{\mathbf{R}} u(x + z) dF(z), \quad (20)$$

and $r(x, F, u)$ by

$$r(x, F, u) = p(x, F, u) - \frac{a(x, u)}{2}V(F). \quad (21)$$

Then, for every $x \in I$ and every $\varepsilon > 0$, there is a $\delta > 0$ such that for every cumulative distribution function F satisfying the basic conditions, if the support of F is included in $[-\delta, \delta]$, then

$$\frac{|r(x, F, u)|}{V(F)} < \varepsilon.$$

Lemma 2 *Let $u : I \rightarrow \mathbf{R}$ and $v : I \rightarrow \mathbf{R}$ be two expected utility functions satisfying the basic conditions. For each $x \in I$ and each cumulative distribution function F satisfying the basic conditions, define $p(x, F, u)$ by*

$$u(x - p(x, F, u)) = \int_{\mathbf{R}} u(x + z) dF(z), \quad (22)$$

and analogously for $p(x, F, v)$. Then, for every $x \in I$ and every $\varepsilon > 0$, there is a $\delta > 0$ such that for every cumulative distribution function F satisfying the basic conditions, if

the support of F is included in $[-\delta, \delta]$, then

$$\left| \frac{p(x, F, v)}{p(x, F, u)} - \frac{a(x, v)}{a(x, u)} \right| < \varepsilon.$$

Proof of Lemma 2 By the definitions of $r(x, F, u)$ and $r(x, F, v)$,

$$\frac{p(x, F, v)}{p(x, F, u)} = \frac{\frac{V(F)}{2}a(x, v) + r(x, F, v)}{\frac{V(F)}{2}a(x, u) + r(x, F, u)} = \frac{a(x, v) + 2\frac{r(x, F, v)}{V(F)}}{a(x, u) + 2\frac{r(x, F, u)}{V(F)}}.$$

By Lemma 1, by taking a sufficiently small $\delta > 0$, for every cumulative distribution function F whose support is included in $[-\delta, \delta]$, the second terms of the numerator and the denominator can be made arbitrarily close to zero. The proof is thus completed. ///

Proof of Theorem 1 Suppose that $(u_1, v_1) \triangleright (u_2, v_2)$. Let x_n, F_n , and ε_n be as in Definition 2. Then (8) holds. Write \widehat{F}_n be the cumulative distribution function defined by $\widehat{F}_n(z) = F_n(z + x)$. Then \widehat{F}_n has mean zero and its support is included in $[-\varepsilon_n, \varepsilon_n]$. By Lemma 2, ε_n can be so small that

$$\frac{p(x_1, \widehat{F}_1, v_1)}{p(x_1, \widehat{F}_1, u_1)} > \frac{p(x_2, \widehat{F}_2, v_2)}{p(x_2, \widehat{F}_2, u_2)}.$$

Let

$$\begin{aligned} \tau &\in \left(\frac{p(x_2, \widehat{F}_2, v_2)}{p(x_2, \widehat{F}_2, u_2)}, \frac{p(x_1, \widehat{F}_1, v_1)}{p(x_1, \widehat{F}_1, u_1)} \right), \\ q_1 &\in \left(p(x_1, \widehat{F}_1, u_1), \frac{p(x_1, \widehat{F}_1, v_1)}{\tau} \right), \\ q_2 &\in \left(\frac{p(x_2, \widehat{F}_2, v_2)}{\tau}, p(x_2, \widehat{F}_2, u_2) \right). \end{aligned}$$

Then

$$\begin{aligned} q_1 &> p(x_1, \widehat{F}_1, u_1), \\ \tau q_1 &< p(x_1, \widehat{F}_1, v_1), \\ q_2 &< p(x_2, \widehat{F}_2, u_2), \\ \tau q_2 &> p(x_2, \widehat{F}_2, v_2). \end{aligned}$$

Thus, the four inequalities of Condition 2 of Definition 2 are met.

The inclusion $\triangleright \subseteq \blacktriangleright$ follows from the definitions of \triangleright and \blacktriangleright .

Suppose that $(u_1, v_1) \blacktriangleright (u_2, v_2)$. Let $x_n \in I_n$. For each small $\varepsilon_n > 0$ and each cumulative distribution function that have zero mean and support in $[-\varepsilon_n, \varepsilon_n]$, there are q_n and τ such that the four inequalities of Condition 1 of Definition 2 hold. Thus

$$\begin{aligned} q_1 &\geq p(x_1, \widehat{F}_1, u_1), \\ \tau q_1 &\leq p(x_1, \widehat{F}_1, v_1), \\ q_2 &\leq p(x_2, \widehat{F}_2, u_2), \\ \tau q_2 &\geq p(x_2, \widehat{F}_2, v_2). \end{aligned}$$

Thus,

$$\frac{p(x_2, \widehat{F}_2, v_2)}{p(x_2, \widehat{F}_2, u_2)} \leq \tau \leq \frac{p(x_1, \widehat{F}_1, v_1)}{p(x_1, \widehat{F}_1, u_1)}.$$

By Lemma 1, by taking $\varepsilon_n > 0$ sufficiently small, we can make $p(x_n, \widehat{F}_n, v_n)/p(x_n, \widehat{F}_n, u_n)$ arbitrarily close to $a(x_n, v_n)/a(x_n, u_n)$. Thus,

$$\frac{a(x_2, v_2)}{a(x_2, u_2)} \leq \frac{a(x_1, v_1)}{a(x_1, u_1)}.$$

Thus $(u_1, v_1) \widehat{\blacktriangleright} (u_2, v_2)$. ///

Proof of Theorem 2 Suppose that $(u_1, v_1) \blacktriangleright (u_2, v_2)$, then $W_1 \widehat{\blacktriangleright} W_2$. Let x_n, ε_n, c_n , and d_n be as in Definition 4. Let e_n be as in Condition 2 of the definition. Denote by F_n the cumulative distribution function of $\lambda \circ d_n^{-1}$ (which coincides with $\mu_n \circ e_n^{-1}$). Then F_n has mean x_n and its support is included in $[x_n - \varepsilon_n, x_n + \varepsilon_n]$. By the definition of $\widehat{\blacktriangleright}$, there are a q_n and τ such that

$$u_1(x_1 - q_1) < \int u_1(z) dF_1(z), \quad (23)$$

$$v_1(x_1 - \tau q_1) > \int v_1(z) dF_1(z), \quad (24)$$

$$u_2(x_2 - q_2) > \int u_2(z) dF_2(z), \quad (25)$$

$$v_2(x_2 - \tau q_2) < \int v_2(z) dF_2(z). \quad (26)$$

Note that

$$W_n((x_n - q_n)\iota) = v_n(x_1 - q_n), \quad (27)$$

$$W_n((x_n - \tau q_n)\iota) = v_n(x_n - \tau q_n). \quad (28)$$

By (16),

$$W_n(c_n) = \int_D v_n(e_n(\pi)) d\mu_n(\pi) = \int_{I_n} v_n(z) dF_n(z), \quad (29)$$

$$W_n(d_n) = v_n \left(u_n^{-1} \left(\int_{[0,1]} u_n(d_n(\xi)) d\lambda(\xi) \right) \right) = v_n \left(u_n^{-1} \left(\int_{I_n} u_n(z) dF_n(z) \right) \right) \quad (30)$$

By applying $v_1 \circ u_1^{-1}$ to both sides of (23), we obtain

$$v_1(x_1 - q_1) < v_1 \left(u_1^{-1} \left(\int u_1(z) dF_1(z) \right) \right). \quad (31)$$

By (27), (30), and (31), $W_1((x_1 - q_1)\iota) < W_1(d_1)$. By (24), (28), and (29), $W_1((x_1 - \tau q_1)\iota) > W_1(c_1)$. We can analogously show that $W_2((x_2 - q_2)\iota) > W_2(d_2)$ and $W_2((x_2 - \tau q_2)\iota) < W_2(c_2)$. Thus Condition 3 of Definition 4 is met. Thus, $W_1 \widehat{\triangleright} W_2$. This proves that if $(u_1, v_1) \overset{\circ}{\triangleright} (u_2, v_2)$, then $W_1 \widehat{\triangleright} W_2$.

We can similarly show that if $(u_1, v_1) \blacktriangleright (u_2, v_2)$, then $W_1 \blacktriangleright W_2$. ///

Proof of Theorem 3 Suppose that $W_1 \widehat{\triangleright} W_2$, then $(u_1, v_1) \overset{\circ}{\triangleright} (u_2, v_2)$. To do so, for each $n = 1, 2$ and let $x_n \in I_n$, $\varepsilon_n > 0$, and F_n be as in Definition 2. Let d_n be the generalized inverse of F_n , in the sense of Embrecht and Hofert (2014), then the cumulative distribution function of d_n coincides with F_n . Since μ_n is rich, there is a $c_n : \Omega \times [0, 1] \rightarrow I$ such that the cumulative distribution function of $\mu_n \circ e_n^{-1}$ coincides with F_n , where e_n is defined by (17). Then $\lambda \circ d_n^{-1} = \mu_n \circ e_n^{-1}$, and they have mean x_n and supports included in $[x_n - \varepsilon_n, x_n + \varepsilon_n]$. By the definition of $\widehat{\triangleright}$, the four inequalities in Condition 3 of Definition 4 hold as strict inequalities. By reverting the argument of showing that that if $(u_1, v_1) \triangleright (u_2, v_2)$ then $W_1 \widehat{\triangleright} W_2$, we can show that these four inequalities imply the four inequalities in Condition 3 of Theorem 1. This proves that if $W_1 \widehat{\triangleright} W_2$, then $(u_1, v_1) \overset{\circ}{\triangleright} (u_2, v_2)$.

We can similarly show that if μ_1 and μ_2 are rich and $W_1 \blacktriangleright W_2$, then $(u_1, v_1) \blacktriangleright (u_2, v_2)$. ///

References

- [1] Chen, H., Ju, N., and J. Miao, 2014, Dynamic asset allocation with ambiguous return predictability. *Review of Economic Dynamics* 17, 799–823.
- [2] Collard, F., S. Mukerji, K. Sheppard, and J. Tallon, 2018, Ambiguity and the historical equity premium, *Quantitative Economics* 945–993.
- [3] Paul Embrechts and Marius Hofert, 2014, A note on generalized inverses, manuscript.

- [4] Larry Epstein, 1995, Behavior under risk: recent developments in theory and applications, in Chapter 1 of *Advances in Economic Theory: Sixth World Congress* Vol. II, edited by Jean-Jacques Laffont, Cambridge University Press.
- [5] Larry Epstein, 2010, A paradox for the “smooth ambiguity” model of preference, *Econometrica* 78, 2085–2099.
- [6] Epstein, L. G., and S. E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57: 937–969.
- [7] Gallant, A. R., M. Jahan-Parvar, and H. Liu, forthcoming, Does smooth ambiguity matter for asset pricing? *Review of Financial Studies*.
- [8] Yoram Halevy, 2007, Ellsberg revisited: an experimental study, *Econometrica* 75, 503–536.
- [9] Chiaki Hara and Toshiki Honda, 2018, Implied ambiguity: Mean-variance efficiency and pricing errors, KIER Discussion Paper Series 1004.
- [10] Takashi Hayashi and Jianjun Miao, 2011, Intertemporal substitution and recursive smooth ambiguity preferences, *Theoretical Economics* 6, 423–472.
- [11] Jahan-Parvar, M. R., and H. Liu, 2014, Ambiguity aversion and asset prices in production economies, *Review of Financial Studies* 27, 3060–3097.
- [12] Ju, N., and J. Miao, 2012, Ambiguity, learning, and asset returns. *Econometrica* 80, 559–591.
- [13] Peter Klibanoff, Massimo Marinacci, and Sujoy Mukerji, 2005, A smooth model of decision making under ambiguity, *Econometrica* 73, 1849–1892.
- [14] Peter Klibanoff, Massimo Marinacci, and Sujoy Mukerji. 2012. On the smooth ambiguity model: A reply, *Econometrica* 80, 1303–1321.
- [15] John W. Pratt, 1964, Risk aversion in the small and in the large, *Econometrica* 32, 122–136.
- [16] Stephen A. Ross, 1981, Some stronger measures of risk aversion in the small and the large with applications, *Econometrica* 49, 621–638.
- [17] Kyoungwong Seo, 2009, Ambiguity and second-order belief, *Econometrica* 77, 1575–1605.

- [18] Thimme, J., and C. Völkert, 2015, Ambiguity in the cross-section of expected returns: An empirical assessment, *Journal of Business and Economic Statistics* 33, 418–429.
- [19] Robert Wilson, 1968, The theory of syndicates, *Econometrica* 36, 119–132.