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

Chiaki Hara*

March 6, 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, Tadashi Sekiguchi, 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 $\varphi_1$ and $\varphi_2$ 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 $\varphi_1$ to be more concave than $\varphi_2$. But this statement makes sense only if $\varphi_1$ and $\varphi_2$ 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 $\psi_1 : \boldsymbol{R}_{++} \to \boldsymbol{R}_{++}$ and $\psi_2 : \boldsymbol{R}_{++} \to \boldsymbol{R}_{++}$ by $v_1' = \psi_1 \circ u_1'$ and $v_2' = \psi_2 \circ u_2'$ and compare $\psi_1$ and $\psi_2$. These functions $\psi_1$ and $\psi_2$ 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 $\psi_1$ and $\psi_2$ in terms of their elasticities rather than the curvature (which is used when comparing $\varphi_1$ and $\varphi_2$). 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 tranformation function $\varphi$ 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 $\varphi$ (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 $\varphi$ 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 $\boldsymbol{R}$ and $u : I \to \boldsymbol{R}$. Assume that $u$ is thrice continuously differentiable and that $u'' < 0 < u'$. We also impose the Inada condition, that is, $u'(x) \to 0$ as $x \to \sup I$, and $u'(x) \to \infty$ as $x \to \inf I$. We call these conditions the basic conditions.

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

For a utility function $u : I \to \boldsymbol{R}$, we define the Arrow-Pratt measure of absolute risk aversion $a(\cdot, u) : I \to \boldsymbol{R}_{++}$ 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 $I = \boldsymbol{R}$ and, with $\gamma > 0$,

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

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

[1] A 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.

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

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

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

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

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

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

1. *If there is a* $b > 0$ *such that* $r(x, u) \leq 1$ *for every* $x \geq b$*, then* $\operatorname{Ran} u$ *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* $\operatorname{Ran} u$ *is bounded from below.*

3. *If there is a* $b > 0$ *such that* $r(x, u) \geq 1$ *for every* $x \leq b$*, then* $\operatorname{Ran} u$ *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* $\operatorname{Ran} u$ *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* $u : I \to \boldsymbol{R}$ *and* $v : I \to \boldsymbol{R}$ *satisfy the basic conditions. Define* $\psi : \boldsymbol{R}_{++} \to \boldsymbol{R}_{++}$ *by* $\psi = v' \circ (u')^{-1}$*. Then, for every* $x \in I$*,*

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

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

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

Then (3) can be rewritten as

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

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 $u_1 : I_1 \to \boldsymbol{R}$, $v_1 : I_1 \to \boldsymbol{R}$, $u_2 : I_2 \to \boldsymbol{R}$, and $v_2 : I_2 \to \boldsymbol{R}$ 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

$$e(y_1, \psi_1) \geq e(y_2, \psi_2) \tag{5}$$

   for every $y_1 \in \boldsymbol{R}_{++}$ and $y_2 \in \boldsymbol{R}_{++}$. 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

$$e(y_1, \psi_1) > e(y_2, \psi_2) \tag{6}$$

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

In the definition, the domains $I_1$ and $I_2$ may be different, and the ranges $\operatorname{Ran} u_1$, $\operatorname{Ran} v_1$, $\operatorname{Ran} u_2$, $\operatorname{Ran} v_2$ 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

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

for every $y \in \boldsymbol{R}_{++}$, and we may say that $\psi_1$ is more elastic as $\psi_2$. If, in addition, we followed the terminology of the strongly-more-risk-averse-than relation of Ross (1981), we could say that that $\psi_1$ is strongly more elastic as $\psi_2$. By (3), (6) is equivalent to the conditions that

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

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.

In concluding this section, we point out that the definition covers the case where $e(y_n, \psi_n) < 1$, that is, the change from $u_n$ to $v_n$, in fact, reduces risk aversion. No mathematical argument needs to be modified in the subsequent analysis. Thus, the expression, "more risk averse to a greater extent", should be considered as a simplifying terminology that covers the case where $u_n$ is more risk averse than $v_n$ as well as the case where $u_n$ is less risk averse than $v_n$.

# 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 $\psi_n$ 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 $u_1 : I_1 \to \boldsymbol{R}$, $v_1 : I_1 \to \boldsymbol{R}$,

[2]For 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 $e(\cdot, \psi_1)$ take the same constant value. Thus, the symmetric part of $\blacktriangleright$ corresponds to the pair $(\psi_1, \psi_2)$ of identical transformations that have a constant elasticity.

$u_2 : I_2 \to \boldsymbol{R}$, and $v_2 : I_2 \to \boldsymbol{R}$ have constant coefficients $\gamma_1$, $\eta_1$, $\gamma_2$, and $\eta_2$ of absolute risk aversion (1). Then

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

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

$$e(\cdot, \psi_n) = \eta_1/\gamma_n \tag{10}$$

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 $(u_1, v_1) \vartriangleright (u_2, v_2)$ 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 $u_1 : I_1 \to \boldsymbol{R}$, $v_1 : I_1 \to \boldsymbol{R}$, $u_2 : I_2 \to \boldsymbol{R}$, and $v_2 : I_2 \to \boldsymbol{R}$ have constant coefficients $\gamma_1$, $\eta_1$, $\gamma_2$, and $\eta_2$ 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$(u_1, v_1) \vartriangleright (u_2, v_2)$ 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 $\psi_n$ that transforms $u_n'$ to $v_n'$, rather than the function $\varphi_n$ that transforms $u_n$ to $v_n$ $(n = 1, 2)$. Indeed, if we used the latter, the domain of $\varphi_n$ coincides with $\operatorname{Ran} u_n$, which may be either $\boldsymbol{R}_{++}$ or $-\boldsymbol{R}_{++}$, depending on whether $\gamma_n$ is smaller or greater than one. For example, if $\gamma_1 < 1 < \gamma_2$, then the domain of $\varphi_1$ coincides with $\boldsymbol{R}_{++}$, while the domain of $\varphi_2$ coincides with $-\boldsymbol{R}_{++}$. 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 $u_1 : I_1 \to \boldsymbol{R}$ and $v_1 : I_1 \to \boldsymbol{R}$ have constant coefficients $\gamma_1$ and $\eta_1$ of absolute risk aversion (1), and two expected utility functions $u_2 : I_2 \to \boldsymbol{R}$ and $v_2 : I_2 \to \boldsymbol{R}$ have constant coefficients $\gamma_2$ and $\eta_2$ 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$(u_1, v_1) \vartriangleright (u_2, v_2)$ 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 $\varphi_n$ that transforms $u_n$ to $v_n$. Since $I_1 = \boldsymbol{R}$ and $I_2 \in \boldsymbol{R}_{++}$ or $I_2 = -\boldsymbol{R}_{++}$, $I_1 \neq I_2$. 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 $b_n \in \boldsymbol{R}$ and the four expected utility functions $u_1 : (b_1, \infty) \to \boldsymbol{R}$, $v_1 : (b_1, \infty) \to \boldsymbol{R}$, $u_2 : (b_2, \infty) \to \boldsymbol{R}$, and $v_2 : (b_2, \infty) \to \boldsymbol{R}$ have hyperbolic absolute risk aversion with the cautiousness parameters $\gamma_n$ and $\eta_n$:[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 $(u_1, v_1) \rhd (u_2, v_2)$ 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 $b_n \in \boldsymbol{R}$ and the four expected utility functions $u_1 : (-\infty, b_1) \to \boldsymbol{R}$, $v_1 : (-\infty, b_1) \to \boldsymbol{R}$, $u_2 : (-\infty, b_2) \to \boldsymbol{R}$, and $v_2 : (-\infty, b_2) \to \boldsymbol{R}$ have hyperbolic absolute risk aversion with the cautiousness parameters $\gamma_n$ and $\eta_n$:

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

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 $(u_1, v_1) \rhd (u_2, v_2)$ 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 $\blacktriangleright$.

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

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.

**Example 6** Suppose that two expected utility functions $u_1 : I_1 \to \boldsymbol{R}$ and $u_2 : I_2 \to \boldsymbol{R}$ have constant coefficients $\gamma_1$ and $\gamma_2$ of relative risk aversion (2). For each $n = 1, 2$, let $\alpha_n > 0$ and assume that $\varphi_n$ has the same functional form as the expected utility function of constant absolute risk aversion (1), with the parameter $\gamma$ replaced by $\alpha_n$. Define $v_n = \varphi_n \circ u_n$. Then

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

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

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

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 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

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

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 $\boldsymbol{R}_{++}$ as we vary $y_1$ or $y_2$. Thus, there is no value of the $\alpha_n$'s and the $\gamma_n$'s such that $(u_1, v_1) \blacktriangleright (u_2, v_2)$.

In this example, since $I_1 = I_2 = \boldsymbol{R}_{++}$, we can take $y_n = u_n'(x)$ for each $n$ with a

[4]In 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.

common consumption level $x \in \boldsymbol{R}_{++}$ in the above example. Then (12) can be reduced to

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

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 $\gamma_n$ is greater or smaller than one. Indeed, by (11),

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

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 $\varphi_n$ 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 $\varphi_n$ 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 $\varphi_n$ 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 $\gamma_n$ of relative risk aversion greater than one (because, then, $\operatorname{Ran} u_n = -\boldsymbol{R}_{++}$). This problem can be circumvented if we specify the function $\psi_n$ that transforms $u_n'$ to $v_n'$ to be any plausible form, such as (9), because $\operatorname{Ran} u_n' = \boldsymbol{R}_{++}$ regardless of the values of $\gamma_n$.

# 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. For each $x \in \boldsymbol{R}$ and $\delta \in \boldsymbol{R}_{++}$, denote by $\mathscr{F}(x, \delta)$ the set of all cumulative distribution functions of which the mean is equal to $x$, the variance is strictly positive, and the support is included in $[x - \delta, x + \delta]$.

**Definition 2** Suppose that four expected utility functions $u_1 : I_1 \to \boldsymbol{R}$, $v_1 : I_1 \to \boldsymbol{R}$, $u_2 : I_2 \to \boldsymbol{R}$, and $v_2 : I_2 \to \boldsymbol{R}$ satisfy the basic conditions. We say that $v_1$ *is more risk averse than* $u_1$ *in choice at least to the same extent as* $v_2$ *is more risk averse than* $u_2$, if for each $n$ and for every $x_n \in I_n$, there are a $\delta_n > 0$ and a $\tau_n > 0$ such that $\tau_1 \geq \tau_2$ and, for every $F_n \in \mathscr{F}(x_n, \delta_n)$,

$$
\begin{aligned}
u_1(x_1 - q_1) &\leq \int u_1(z_1)\, \mathrm{d}F_1(z_1), \\
v_1(x_1 - \tau_1 q_1) &\geq \int v_1(z_1)\, \mathrm{d}F_1(z_1), \\
u_2(x_2 - q_2) &\geq \int u_2(z_2)\, \mathrm{d}F_2(z_2), \\
v_2(x_2 - \tau_2 q_2) &\leq \int v_2(z_2)\, \mathrm{d}F_2(z_2).
\end{aligned}
$$

We then write $(u_1, v_1) \overset{\circ}{\blacktriangleright} (u_2, v_2)$. If, in addition, $\tau_1 > \tau_2$, then we say that $v_1$ *is more risk averse than* $u_1$ *in choice to a greater extent than* $v_2$ *is more risk averse than* $u_2$ and write $(u_1, v_1) \overset{\circ}{\triangleright} (u_2, v_2)$.

The first inequality in the definition of $\overset{\circ}{\blacktriangleright}$ 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 $\tau_1 q_1$. Hence, the proportional change in the certainty premium induced by the change from $u_1$ to $v_1$ is greater than or equal to $\tau_1$. 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 $\tau_2 q_2$. Hence, the proportional change in the certainty premium induced by the change from $u_2$ to $v_2$ is smaller than or equal to $\tau_2$. 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 $\overset{\circ}{\blacktriangleright}$ 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 $\overset{\circ}{\triangleright}$ are different from those in the definition of $\overset{\circ}{\blacktriangleright}$ only in that the weak inequality $\tau_1 \geq \tau_2$ is replaced by the strict inequality. By taking $\tau_1$ smaller, $\tau_2$ larger, $q_1$ larger, and $q_2$ smaller if necessary, we can define the relation $\overset{\circ}{\triangleright}$ equivalently by replacing, in Definition 2, the four weak inequalities on expected utility levels by the corresponding strict inequalities. Whenever necessary, we shall refer to both the binary relations $\overset{\circ}{\triangleright}$ and

$\mathring{\blacktriangleright}$ indistinguishably as the behavioral even-more-risk-averse-than relation.

The following theorem establishes the relationship among the four binary relations $\blacktriangleright$, $\triangleright$, $\mathring{\blacktriangleright}$, and $\mathring{\triangleright}$.

**Theorem 1** $\triangleright = \mathring{\triangleright} \subseteq \mathring{\blacktriangleright} \subseteq \blacktriangleright$.

This theorem means that for four expected utility functions $u_1 : I_1 \to \boldsymbol{R}$, $v_1 : I_1 \to \boldsymbol{R}$, $u_2 : I_2 \to \boldsymbol{R}$, and $v_2 : I_2 \to \boldsymbol{R}$ satisfying the basic conditions, $(u_1, v_1) \triangleright (u_2, v_2)$ if and only if $(u_1, v_1) \mathring{\triangleright} (u_2, v_2)$; if $(u_1, v_1) \mathring{\triangleright} (u_2, v_2)$, then $(u_1, v_1) \mathring{\blacktriangleright} (u_2, v_2)$; and if $(u_1, v_1) \mathring{\blacktriangleright} (u_2, v_2)$, 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.

We can conclude, roughly, 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. This latter condition makes sense for preference relations over cumulative distributions functions that may not be represented by expected utility functions. In fact, 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 $I_2$. We can then rewrite the inequalities in the definition of $\mathring{\triangleright}$ as

$$\begin{aligned} F_1 &\succsim_1^1 \mathbf{1}_{[x_1 - q_1, \infty)}, \\ \mathbf{1}_{[x_1 - \tau_1 q_1, \infty)} &\succsim_1^2 F_1, \\ \mathbf{1}_{[x_2 - q_2, \infty)} &\succsim_2^1 F_2, \\ F_2 &\succsim_2^2 \mathbf{1}_{[x_2 - \tau_2 q_2, \infty)}, \end{aligned}$$

where, for every $x$, $\mathbf{1}_{[x,\infty)}$ denotes the (degenerate) cumulative distribution function that is equal to one on $[x, \infty)$ and zero on $(-\infty, x)$. The inequalities in the definition of $\mathring{\blacktriangleright}$ 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 $u : I \to \boldsymbol{R}$ and $v : I \to \boldsymbol{R}$ satisfy the basic conditions. Let $\mu$ be a probability measure on $D$. Define a utility function $W : C \to \boldsymbol{R}$ by letting

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

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

$$W(c) = \int_D \varphi\left(\int_S u(c(s))\,\mathrm{d}\pi(s)\right)\mathrm{d}\mu(\pi). \tag{15}$$

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 $\varphi$ is concave, that is, $v$ is more risk-averse than $u$. The probability measure $\mu$ represents his subjective assessment of this uncertainty, which KMM termed as the second-order belief.

We see in (14) that if $\int_S u(c(s))\,\mathrm{d}\pi(s)$ is independent of $\pi$, then, writing $x = u^{-1}\left(\int_S u(c(s))\,\mathrm{d}\pi(s)\right) \in I$, we obtain $W(c) = u(x)$. This means that if the conditional certainty equivalent of $c$ given a probability measure $\pi \in \operatorname{supp}\mu$ is, in fact, independent of $\pi$, 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 $\pi \in \operatorname{supp}\mu$, but the constant that $c$ takes almost surely depends on $\pi$, then denote the value by $c(\pi)$. Then, $W(c) = \int_D v(c(\pi))\,\mathrm{d}\mu(\pi)$. 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 $\Omega$ is a measurable space and $[0, 1]$ is the closed unit interval endowed with the Lebesgue measure $\lambda$. 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

[5] To 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.

distribution $\lambda$ 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 $\Omega$ by $\pi$, the set of all probability measures on $\Omega$ by $D$, the second-order belief on $D$ by $\mu$. By Fubini's theorem, we can then rewrite (14) as

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

## 6.2 An alternative definition of the 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 $(u_1, v_1, \mu_1)$ and $(u_2, v_2, \mu_2)$ via (16). Denote by $I_n$ the common domain of $u_n$ and $v_n$. Denote by $C_n$ the set of all $c_n : S \to I_n$. 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 $\omega\in\Omega$ for every $\xi\in[0,1]$ 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, in fact, 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 $\pi$ on $\Omega$,[7] where the

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

[7] In 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

family is constructed by indexing the pairs by the supports of second-order beliefs in $D$; 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. 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 $\varphi_1$ and $\varphi_2$. However, Definition 3 provides no sound theoretical foundation for such a comparison unless the estimates of $u_1$ and $u_2$ are the same.

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 $u_1$

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.

[8] Assumption 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).

is an affine transformation of $u_2$. Denote by $\iota$ the function defined on $S$ that constantly takes value one. Our definition can then be stated as follows.

**Definition 4** We say that $W_1$ is *at least as ambiguity-averse as* $W_2$ if for each $n$ and for every $x_n \in I_n$, there are a $\delta_n > 0$ and a $\tau_n > 0$ such that $\tau_1 \geq \tau_2$ and, for 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 they satisfy satisfy the last one.

1. $d_n(\omega, \xi)$ is independent of $\omega \in \Omega$ for every $\xi \in [0,1]$. We thus write it as $d_n(\xi)$ and regard as $d_n : [0,1] \to I_n$;

2. Define $e_n : D \to I_n$ by letting

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

   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 mean is equal to $x_n$ and their support is included in $[x_n - \delta_n, x_n + \delta_n]$.

3. For each $n$, there is a $q_n > 0$ that

$$\begin{aligned} W_1((x_1 - q_1)\iota) &\leq W_1(d_1), \\ W_1((x_1 - \tau_1 q_1)\iota) &\geq W_1(c_1), \\ W_2((x_2 - q_2)\iota) &\geq W_2(d_2), \\ W_2((x_2 - \tau_2 q_2)\iota) &\leq W_2(c_2). \end{aligned}$$

We then write $W_1 \widehat{\blacktriangleright} W_2$. If, in addition, $\tau_1 > \tau_2$, then we say that $W_1$ is *more ambiguity-averse than* $W_2$, and write $W_1 \widehat{\triangleright} W_2$.

Just as in the case of the behavioral even-more-risk-averse-than relation, taking $\tau_1$ smaller, $\tau_2$ larger, $q_1$ larger, and $q_2$ smaller if necessary, we can define the relation $\widehat{\triangleright}$ equivalently by replacing, in Definition 4, the four weak inequalities by the corresponding strict inequalities.

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 $(x_n - \tau_n q_n)\iota$ and a possibly ambiguous consumption plan $c_n$. To be more precise, by (16) and the change-of-variable formula,

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

Thus, 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,

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

Thus, the ranking between $c_n$ and $(x_n - \tau_n q_n)\iota$ can be reduced to the ranking between the distribution $\mu_n \circ e_n^{-1}$ and the deterministic consumption level $x_n - \tau_n q_n$ 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 $e_n$ as the second-order consumption plan 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, $x_n - q_n$, $x_n - \tau_n q_n$, 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 $x_n - \tau_n q_n$, in Definition 4, we need to know the inner utility function $u_n$ and the second-order belief $\mu_n$ because $c_n$ and $d_n$ must satisfy $\lambda \circ d_n^{-1} = \mu_n \circ e_n^{-1}$, where the second-order consumption plan $e_n$ associated with $c_n$ in (17) depends on the inner utility function $u_n$. In this sense, the more-ambiguity-averse-than relation of Definition 4 can more easily be checked when the attitudes towards risk (represented by the inner utility function $u_n$) and the second-order beliefs $\mu_n$ 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. They give an easy way to check whether a KMM utility function is more ambiguity-averse than another.

**Theorem 2** *Define two utility function $W_1 : C_1 \to \boldsymbol{R}$ and $W_2 : C_2 \to \boldsymbol{R}$ on the same state space $S = \Omega \times [0,1]$, with $[0,1]$ endowed with the Lebesgue measure, by two triples $(u_1, v_1, \mu_1)$ and $(u_2, v_2, \mu_2)$ via (16). Then*

1. *If $(u_1, v_1) \mathring{\rhd} (u_2, v_2)$, then $W_1 \widehat{\rhd} W_2$.*

2. *If $(u_1, v_1) \mathring{\blacktriangleright} (u_2, v_2)$, then $W_1 \widehat{\blacktriangleright} W_2$.*

[9]Thus, 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 $\mu_1$ and $\mu_2$ 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 $\tau_n$ 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 $x_n - \tau_n q_n$ 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 second-order consumption plans in experiments to infer $\mu_n$ 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).

This theorem shows, roughly, that the behavioral even-more-risk-averse-than relation implies the more-ambiguity-averse-than relation.

The second theorem is an almost converse of the first. To see why the exact converse cannot be obtained, consider the case where the second-order belief $\mu_n$ is degenerate, that is, concentrated on a single probability measure on $\Omega$. Then, for every $(c_n, d_n)$ satisfying the conditions of Definition 4, the distributions $\mu_n \circ e_n^{-1}$ and $\lambda \circ d_n^{-1}$ are concentrated on a single consumption level and, thus, the inequalities in Condition 3 of Definition 4 cannot have any implication on the utility functions $u_n$ and $v_n$. We thus assume that for both $n$, $\mu_n$ is non-degenerate.

**Theorem 3** *Define two utility function $W_1 : C_1 \to \boldsymbol{R}$ and $W_2 : C_2 \to \boldsymbol{R}$ on the same state space $S = \Omega \times [0, 1]$, with $[0, 1]$ endowed with the Lebesgue measure, by two triples $(u_1, v_1, \mu_1)$ and $(u_2, v_2, \mu_2)$ via* (16)*. Suppose that $\mu_1$ and $\mu_2$ are non-degenerate. Then*

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

Since $\triangleright = \mathring{\triangleright}$, Part 1 of this theorem, along with Part 1 of Theorem 2, implies that $\triangleright$, $\mathring{\triangleright}$, and $\widehat{\triangleright}$ are equivalent. Since $\mathring{\blacktriangleright} \subseteq \blacktriangleright$, Part 2 of this theorem, along with Part 1 of Theorem 2, implies that $\widehat{\blacktriangleright}$ lies, in the order of strength, between $\mathring{\blacktriangleright}$ and $\blacktriangleright$.

## 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 $\eta_n/\gamma_n$ 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 $\gamma_n$ 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 $\gamma_n$ in Example 2, and the coefficients of relative risk aversion of the outer expected utility function $v_n$, which is denoted by $\eta$ 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 $\gamma$ is fixed, say, at 2, increasing $\eta$ 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 $\gamma$ 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 $\gamma$ and with respect to $\eta$, while their "ambiguity premium" is equal to the difference between the certainty two premiums. The latter is more pronounced when the coefficient $\gamma$ 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 $(\gamma, \eta)$ 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 $\gamma$, increasing $\eta$ 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 $\gamma$ 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 $\eta$ 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 $\varphi_n$ that transform $u_n$ to $v_n$ are not comparable, because the domain of $\varphi_n$ is $\boldsymbol{R}$ in the case of constant absolute risk aversion, while it is $\boldsymbol{R}_{++}$ or $-\boldsymbol{R}_{++}$ 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), $\eta_n/\gamma_n$ 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 $\eta_n/\gamma_n$ 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.[10]

# 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 $\psi_1$ and $\psi_2$, we require the elasticity of $\psi_1$ is higher than the elasticity of $\psi_2$, regardless of the choices of the marginal utilities, $y_1$ and $y_2$, of the expected utility functions $u_1$ and $u_2$. 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

[10]It 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.

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.

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 $\alpha$-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) . ///

For every cumulative distribution function on $\boldsymbol{R}$, denote its mean by $E(F)$ and variance by $V(F)$, whenever they exist. Recall that for every $x \in \boldsymbol{R}$ and $\delta \in \boldsymbol{R}_{++}$, $\mathscr{F}(x, \delta)$ is the set of all cumulative distribution functions $F$ such that $E(F) = x$, $V(F) > 0$, and the support of $F$ is included in $[x - \delta, x + \delta]$.

In the subsequent analysis, we often assume, without explicitly stating so, that $F$ is concentrated on the set $I$ of consumption levels for a decision maker. For each expected utility function $u : I \to \boldsymbol{R}$ that satisfies the basic conditions, each $x \in I$, and each cumulative distribution function $F$ with mean $x$, define $p(x, F, u)$ by

$$u(x - p(x, F, u)) = \int_{\boldsymbol{R}} u(z) \, \mathrm{d}F(z). \tag{18}$$

Then $p(x, F, u)$ is the certainty premium of the distribution $F$ of random consumption levels.

**Lemma 1** *Let $u : I \to \boldsymbol{R}$ and $v : I \to \boldsymbol{R}$ be two expected utility functions satisfying the basic conditions. For every $x \in I$ and every $\varepsilon > 0$, there is a $\delta > 0$ such that for every $F \in \mathscr{F}(x, \delta)$,*

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

The following proof is essentially due to Pratt (1964).

**Proof of Lemma 1** Let $x \in I$ and, for a moment, let $\delta > 0$ satisfy $[x - \delta, x + \delta] \subset I$ and $F \in \mathscr{F}(x, \delta)$. We shall later take a smaller $\delta$ when necessary. For each $z$ satisfying $z \in I$, write

$$\begin{aligned} R_2(z, x, u) &= u(z) - (u(x) + u'(x)(z - x)), \\ R_3(z, x, u) &= u(z) - \left( u(x) + u'(x)(z - x) + \frac{u''(x)}{2}(z - x)^2 \right). \end{aligned}$$

Then

$$\begin{aligned} u(x - p(x, F, u)) &= u(x) - u'(x)p(x, F, u) + R_2(x - p(x, F, u), x, u), \\ \int_{\boldsymbol{R}} u(z)\, \mathrm{d}F(z) &= u(x) + \frac{u''(x)}{2} V(F) + \int R_3(z, x, u)\, \mathrm{d}F(z). \end{aligned}$$

By (18),

$$-u'(x)p(x, F, u) + R_2(x - p(x, F, u), x, u) = \frac{u''(x)}{2} V(F) + \int R_3(z, x, u)\, \mathrm{d}F(z).$$

By dividing both sides by $-u'(x)$ and rearranging the terms, we obtain

$$p(x, F, u) - \frac{1}{2} a(x, u) V(F) = \frac{1}{u'(x)} \left( R_2(x - p(x, F, u), x, u) - \int R_3(z, x, u)\, \mathrm{d}F(z) \right).$$

Since $u$ is thrice continuously differentiable, for each $k = 1, 2, 3$, there is an $m_k > 0$ such that $|u^{(k)}(z)| \leq m_k$ for every $z \in [x - \delta, x + \delta]$. Then

$$\begin{aligned} |R_2(z, x, u)| &\leq \frac{m_2}{2} |z|^2, \\ |R_3(z, x, u)| &\leq \frac{m_3}{3!} |z|^3 \end{aligned}$$

for every $z \in [x - \delta, x + \delta]$. Moreover, $0 \leq p(x, F, u) \leq \delta$ and, hence,

$$|R_2(x - p(x, F, u), x, u)| \leq \frac{m_2}{2} (p(x, F, u))^2 \leq \frac{m_2}{2} \delta^2.$$

By the Cauchy-Schwartz inequality,

$$\left|\int R_3(z,x,u)\,\mathrm{d}F(z)\right| = \left|\int \frac{R_3(z,x,u)}{(z-x)^2}(z-x)^2\,\mathrm{d}F(z)\right|$$
$$\leq \left(\int \left(\frac{R_3(z,x,u)}{(z-x)^2}\right)^2 \mathrm{d}F(z)\right)^{1/2} \left(\int \left((z-x)^2\right)^2 \mathrm{d}F(z)\right)^{1/2}. \tag{19}$$

Note that

$$\frac{|R_3(z,x,u)|}{(z-x)^2} \leq \frac{m_3}{3!}\frac{|z-x|^3}{(z-x)^2} = \frac{m_3}{3!}|z-x|$$

for every $z \in [x-\delta, x+\delta]$. Thus,

$$\int \left(\frac{R_3(z,x,u)}{(z-x)^2}\right)^2 \mathrm{d}F(z) \leq \left(\frac{m_3}{3!}\right)^2 \delta^2.$$

Again by the Cauchy-Schwartz inequality,

$$\left(\int \left((z-x)^2\right)^2 \mathrm{d}F(z)\right)^{1/2} \leq \left(\left(\int (z-x)^2 \mathrm{d}F(z)\right)^{1/2}\right)^2 \leq \delta^2.$$

By (19),

$$\left|\int R_3(z,x,u)\,\mathrm{d}F(z)\right| \leq \frac{m_3}{3!}\delta^2.$$

Hence,

$$\left|R_2(x-p(x,F,u),x,u) - \int R_3(z,x,u)\,\mathrm{d}F(z)\right| \leq \left(\frac{m_2}{2}+\frac{m_3}{3!}\right)\delta^2.$$

Thus,

$$\left|p(x,F,u) - \frac{1}{2}a(x,u)V(F)\right| \leq \frac{1}{m_1}\left(\frac{m_2}{2}+\frac{m_3}{3!}\right)\delta^2.$$

An analogous inequality holds for $v$. Thus, for every $\varepsilon > 0$, if $\delta > 0$ is sufficiently small, then

$$\left|\frac{p(x,F,v)}{p(x,F,u)} - \frac{a(x,v)}{a(x,u)}\right| = \left|\frac{p(x,F,v)}{p(x,F,u)} - \frac{(1/2)a(x,v)V(F)}{(1/2)a(x,u)V(F)}\right| < \varepsilon.$$

///

**Proof of Theorem 1** Suppose that $(u_1, v_1) \rhd (u_2, v_2)$. For each $n$, let $x_n \in I_n$. By (8), there are a $\tau_1$ and a $\tau_2$ such that

$$\frac{a(x_1, v_1)}{a(x_1, u_1)} > \tau_1 > \tau_2 > \frac{a(x_2, v_2)}{a(x_2, u_2)}.$$

By Lemma 1, for each $n$, there is a $\delta_n > 0$ such that for every $F_n \in \mathscr{F}(x_n, \delta_n)$,

$$\frac{p(x_1, F_1, v_1)}{p(x_1, F_1, u_1)} > \tau_1 \text{ and } \tau_2 > \frac{p(x_2, F_2, v_2)}{p(x_2, F_2, u_2)}.$$

Then let

$$\begin{aligned} q_1 &\in \left(p(x_1, F_1, u_1), \frac{p(x_1, F_1, v_1)}{\tau_1}\right), \\ q_2 &\in \left(\frac{p(x_2, F_2, v_2)}{\tau_2}, p(x_2, F_2, u_2)\right). \end{aligned}$$

Then,

$$\begin{aligned} q_1 &> p(x_1, F_1, u_1), \\ \tau_1 q_1 &< p(x_1, F_1, v_1), \\ q_2 &< p(x_2, F_2, u_2), \\ \tau_2 q_2 &> p(x_2, F_2, v_2). \end{aligned}$$

By the definition (18), the four inequalities of Condition 2 of Definition 2 are met. Thus $(u_1, v_1)\mathring{\rhd}(u_2, v_2)$. Therefore, $\rhd \subseteq \mathring{\rhd}$.

Suppose that $(u_1, v_1)\mathring{\rhd}(u_2, v_2)$. For each $n$, let $x_n \in I_n$. Let $\tau_n$ and $\delta_n$ be as in Definition 2. Let $\varepsilon = \tau_1 - \tau_2$. By Lemma 1, there is an $F_n \in \mathscr{F}_n(x_n, \delta_n)$ such that

$$\left|\frac{p(x_n, F_n, v_n)}{p(x_n, F_n, u_n)} - \frac{a(x_n, v_n)}{a(x_n, u_n)}\right| < \frac{\varepsilon}{2}.$$

By the four inequalities in Definition 2,

$$\frac{p(x_1, F_1, v_1)}{p(x_1, F_1, u_1)} \geq \tau_1 \text{ and } \tau_2 \geq \frac{p(x_2, F_2, v_2)}{p(x_2, F_2, u_2)}.$$

Thus,

$$\frac{a(x_1, v_1)}{a(x_1, u_1)} > \tau_1 - \frac{\varepsilon}{2} = \tau_2 + \frac{\varepsilon}{2} > \frac{a(x_2, v_2)}{a(x_2, u_2)}.$$

Hence, $(u_1, v_1) \rhd (u_2, v_2)$. Thus, $\mathring{\rhd} \subseteq \rhd$. Since $\rhd \subseteq \mathring{\rhd}$, we have shown that $\rhd = \mathring{\rhd}$.

It follows immediately from the definition that $\mathring{\rhd} \subseteq \mathring{\blacktriangleright}$. We prove by contraposition that $\mathring{\blacktriangleright} \subseteq \blacktriangleright$. That is, we show that if $(u_1, v_1) \not\blacktriangleright (u_2, v_2)$, then $(u_1, v_1) \not\mathring{\blacktriangleright}(u_2, v_2)$. To do so, note that if $(u_1, v_1) \mathring{\blacktriangleright} (u_2, v_2)$, then for every $F_n \in \mathscr{F}(x_n, \delta_n)$,

$$\begin{aligned} q_1 &\geq p(x_1, F_1, u_1), \\ \tau_1 q_1 &\leq p(x_1, F_1, v_1), \\ q_2 &\leq p(x_2, F_2, u_2), \\ \tau_2 q_2 &\geq p(x_2, F_2, v_2). \end{aligned}$$

Since $\tau_1 \geq \tau_2$,

$$\frac{p(x_1, F_1, v_1)}{p(x_1, F_1, u_1)} \geq \frac{p(x_2, F_2, v_2)}{p(x_2, F_2, u_2)}.$$

To show that $\mathring{\blacktriangleright} \subseteq \blacktriangleright$, therefore, it suffices to prove that if $(u_1, v_1) \not\blacktriangleright (u_2, v_2)$, then, for each $n$, there is an $x_n \in I_n$ such that for every $\delta_n > 0$ and $\tau_n > 0$, there is an $F_n \in \mathscr{F}(x_n, \delta_n)$ such that

$$\frac{p(x_1, F_1, v_1)}{p(x_1, F_1, u_1)} < \frac{p(x_2, F_2, v_2)}{p(x_2, F_2, u_2)}. \tag{20}$$

In fact, if $(u_1, v_1) \not\blacktriangleright (u_2, v_2)$, then, for each $n$, there is an $x_n$ such that

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

By Lemma 1, for every sufficiently small $\delta_n > 0$ and every $F_n \in \mathscr{F}_n(x_n, \delta_n)$, (20) holds. ///

**Proof of Theorem 2** Suppose that $(u_1, v_1) \mathring{\triangleright} (u_2, v_2)$. For each $n$, let $x_n \in I_n$. Let $\tau_n$ and $\delta_n$ be as in Definition 2. Let $d_n$, $c_n$, and $e_n$ satisfy the first two conditions of Definition 4. Denote by $F_n$ the cumulative distribution function of $\lambda \circ d_n^{-1}$ and $\mu_n \circ e_n^{-1}$. Then $F_n \in \mathscr{F}(x_n, \delta_n)$ and

$$u_1(x_1 - q_1) \leq \int u_1(z_1)\, \mathrm{d}F_1(z_1), \tag{21}$$

$$v_1(x_1 - \tau_1 q_1) \geq \int v_1(z_1)\, \mathrm{d}F_1(z_1), \tag{22}$$

$$u_2(x_2 - q_2) \geq \int u_2(z_2)\, \mathrm{d}F_2(z_2), \tag{23}$$

$$v_2(x_2 - \tau_2 q_2) \leq \int v_2(z_2)\, \mathrm{d}F_2(z_2). \tag{24}$$

Note that

$$W_n((x_n - q_n)\iota) = v_n(x_n - q_n), \tag{25}$$
$$W_n((x_n - \tau_n q_n)\iota) = v_n(x_n - \tau_n q_n). \tag{26}$$

By (16),

$$W_n(d_n) = v_n\left(u_n^{-1}\left(\int_{[0,1]} u_n\left(d_n(\xi)\right)\mathrm{d}\lambda(\xi)\right)\right) = v_n\left(u_n^{-1}\left(\int_{\boldsymbol{R}} u_n\left(z_n\right)\mathrm{d}F_n(z_n)\right)\right) \tag{27}$$

By (17),

$$W_n(c_n) = \int_D v_n\left(e_n(\pi)\right)\mathrm{d}\mu_n(\pi) = \int_{\boldsymbol{R}} v_n(z_n)\,\mathrm{d}F_n(z_n). \tag{28}$$

By applying $v_1 \circ u_1^{-1}$ to both sides of (21), we obtain

$$v_1(x_1 - q_1) \leq v_1\left(u_1^{-1}\left(\int u_1(z_1)\,\mathrm{d}F_1(z_1)\right)\right). \tag{29}$$

By (22), (26), and (28), $W_1((x_1 - \tau_1 q_1)\iota) \geq W_1(c_1)$. By (25), (27), and (29), $W_1((x_1 - q_1)\iota) \leq W_1(d_1)$. We can analogously show that $W_2((x_2 - q_2)\iota) \geq W_2(d_2)$ and $W_2((x_2 - \tau_2 q_2)\iota) \leq W_2(c_2)$. Thus the third condition of Definition 4 is met. Thus, $W_1 \widehat{\triangleright} W_2$. This proves that if $(u_1, v_1)\mathring{\triangleright}(u_2, v_2)$, then $W_1 \widehat{\triangleright} W_2$.

Since $\mathring{\blacktriangleright}$ differs from $\mathring{\triangleright}$ only in $\tau_1 > \tau_2$ and $\tau_1 \geq \tau_2$, the above proof can be used to show that if $(u_1, v_1)\mathring{\blacktriangleright}(u_2, v_2)$, then $W_1 \widehat{\blacktriangleright} W_2$. ///

To prove Theorem 3, we need a lemma on second-order consumption plans. Assume that $\mu$ is not degenerate. This is equivalent to saying that there is a measurable subset $\Upsilon$ of $\Omega$ such that the function $\pi \mapsto \pi(T)$ of $D$ into $[0, 1]$ is not constant on any set to which $\mu$ gives measure one. We keep such an $\Upsilon$ fixed until the end of the next lemma. For each $\underline{x} \in \boldsymbol{R}$ and $\overline{x} \in \boldsymbol{R}$, define $c^{(\underline{x},\overline{x})} \in C$ by letting, for each $(\omega, \xi) \in \Omega \times [0, 1]$,

$$c^{(\underline{x},\overline{x})}(\omega, \xi) = \begin{cases} \overline{x} & \text{if } \omega \in \Upsilon, \\ \underline{x} & \text{otherwise.} \end{cases} \tag{30}$$

Note that $c^{(\underline{x},\overline{x})}(\omega, \xi)$ is independent of $\xi \in [0, 1]$ for every $\omega \in \Omega$. By an abuse of notation, we also write $c^{(\underline{x},\overline{x})}(\omega)$ in place of $c^{(\underline{x},\overline{x})}(\omega, \xi)$.

**Lemma 2** *Let $u : I \to \boldsymbol{R}$ be an expected utility function satisfying the basic condition. For each $x \in I$ and $\delta > 0$, there are an $\underline{x} \in \boldsymbol{R}$ and an $\overline{x} \in \boldsymbol{R}$ such that* $0 < x - \underline{x} \leq \delta$,

$0 < \overline{x} - x \leq \delta$, *and*

$$\int_D u^{-1}\left(\int_\Omega u\left(c^{(\underline{x},\overline{x})}(\omega)\right) \mathrm{d}\pi(\omega)\right) \mathrm{d}\mu(\pi) = x. \tag{31}$$

To grasp a useful implication of this lemma, let $e : D \to I$ be the second-order consumption plan associated with $c^{(\underline{x},\overline{x})}$, which is defined in (17), and denote by $F$ the cumulative distribution function of $\mu \circ e^{-1}$. Since $\pi \mapsto \pi(\Upsilon)$ is not constant on any set to which $\mu$ gives measure one, $V(F) > 0$. By (31),

$$x = \int_D e(\pi) \, \mathrm{d}\mu(\pi) = \int_I z \, \mathrm{d}(\mu \circ e^{-1}) = E(F).$$

Thus, $F \in \mathscr{F}(x, \delta)$.

**Proof of Lemma 2** By the definition of $c^{(\underline{x},\overline{x})}$,

$$u^{-1}\left(\int_\Omega u\left(c^{(\underline{x},\overline{x})}(\omega)\right) \mathrm{d}\pi(\omega)\right) = u^{-1}\left(u(\underline{x}) + \pi(\Upsilon)(u(\overline{x}) - u(\underline{x}))\right).$$

We can assume without loss of generality that $x - \delta \in I$ and $x + \delta \in I$. Since it is not true that $\pi(\Upsilon) = 1$ for $\mu$-almost every $\pi$,

$$\int_D u^{-1}\left(u(x-\delta) + \pi(\Upsilon)(u(x) - u(x-\delta))\right) \mathrm{d}\mu(\pi) < x.$$

By the bounded convergence theorem, for every $\overline{x} > x$ sufficiently close to $x$,

$$\int_D u^{-1}\left(u(x-\delta) + \pi(\Upsilon)(u(\overline{x}) - u(x-\delta))\right) \mathrm{d}\mu(\pi) < x. \tag{32}$$

Thus, if there is no $\overline{x} \in (x, x+\delta]$ such that

$$\int_D u^{-1}\left(u(\underline{x}) + \pi(\Upsilon)(u(\overline{x}) - u(\underline{x}))\right) \mathrm{d}\mu(\pi) = x,$$

then, by the intermediate value theorem, (32) holds when $\overline{x} = x + \delta$, that is,

$$\int_D u^{-1}\left(u(x-\delta) + \pi(\Upsilon)(u(x+\delta) - u(x-\delta))\right) \mathrm{d}\mu(\pi) < x.$$

Since it is not true that $\mu(\Upsilon) = 0$ for $\mu$-almost every $\mu$,

$$\int_D u^{-1}\left(u(x) + \pi(\Upsilon)(u(x+\delta) - u(x))\right) \mathrm{d}\mu(\pi) > x.$$

By the bounded convergence theorem, for every $\underline{x} < x$ sufficiently close to $x$,

$$\int_D u^{-1}\left(u(\underline{x}) + \pi(\Upsilon)(u(x+\delta) - u(\underline{x}))\right) \mathrm{d}\mu(\pi) > x.$$

Again, by the intermediate value theorem, there is a $\underline{x} \in (x-\delta, x)$ such that

$$\int_D u^{-1}\left(u(\underline{x}) + \pi(\Upsilon)(u(x+\delta) - u(\underline{x}))\right) \mathrm{d}\mu(\pi) = x.$$

Thus (31) holds. ///

**Proof of Theorem 3** Suppose that $W_1 \widehat{\rhd} W_2$. For each $n$, let $x_n \in I_n$. Let $\tau_n$ and $\delta_n$ be as in Definition 4. Then $\tau_1 > \tau_2$. Let $\varepsilon = \tau_1 - \tau_2 > 0$. By taking $\delta_n > 0$ smaller if necessary, we can assume that the conclusion of Lemma 1 holds for $\varepsilon/2$ and $\delta_n$ as well.

Apply Lemma 2 to $x_n$ and $\delta_n$ to obtain a consumption plan $c_n^{(\underline{x}_n, \overline{x}_n)}$, which is defined in (30). Let $e_n$ be the second-order consumption plan associated with $c_n^{(\underline{x}_n, \overline{x}_n)}$, which is defined in (17). Denote by $F_n$ the cumulative distribution function of $\mu_n \circ e_n^{-1}$. Then $F_n \in \mathscr{F}(x_n, \delta_n)$. Let $d_n : [0,1] \to I_n$ 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 $F_n$ and $\lambda \circ d_n^{-1} = \mu_n \circ e_n^{-1}$. Thus $c_n^{(\underline{x}_n, \overline{x}_n)}$ and $d_n$ satisfy the first two conditions of Definition 4. Thus the four inequalities in the last condition hold. Thus, by reverting the argument in the proof of Theorem 2, we can show that

$$\begin{aligned}
q_1 &\geq p(x_1, F_1, u_1), \\
\tau_1 q_1 &\leq p(x_1, F_1, v_1), \\
q_2 &\leq p(x_2, F_2, u_2), \\
\tau_2 q_2 &\geq p(x_2, F_2, v_2).
\end{aligned}$$

Hence

$$\frac{p(x_1, F_1, v_1)}{p(x_1, F_1, u_1)} \geq \tau_1 > \tau_2 \geq \frac{p(x_2, F_2, v_2)}{p(x_2, F_2, u_2)},$$

Thus, by Lemma 1,

$$\frac{a(x_1, v_1)}{a(x_1, u_1)} > \tau_1 - \frac{\varepsilon}{2} = \tau_2 + \frac{\varepsilon}{2} > \frac{a(x_2, v_2)}{a(x_2, u_2)}.$$

Hence, $(u_1, v_1) \rhd (u_2, v_2)$.

We shall next prove that if $(u_1, v_1) \widehat{\blacktriangleright} (u_2, v_2)$, then $(u_1, v_1) \blacktriangleright (u_2, v_2)$. We do so by contraposition. Suppose that $(u_1, v_1) \not\blacktriangleright (u_2, v_2)$. We need to show that $(u_1, v_1) \not\widehat{\blacktriangleright} (u_2, v_2)$, that is, for each $n$, there is an $x_n \in I_n$ such that for every $\delta_n > 0$ and every $\tau_n > 0$, there are a $c_n$ and a $d_n$ that satisfy the first two conditions of Definition 4 but do not satisfy the last one when $\tau_1 \geq \tau_2$. As we showed when proving that $\widehat{\triangleright} \subseteq \triangleright$, if $\tau_1 \geq \tau_2$ and if $c_n$ and $d_n$ satisfy the last condition of Definition 4, then

$$\frac{p(x_1, F_1, v_1)}{p(x_1, F_1, u_1)} \geq \frac{p(x_2, F_2, v_2)}{p(x_2, F_2, u_2)},$$

where $F_n$ is the cumulative distribution function of $\mu_n \circ e_n^{-1}$ and $\lambda \circ d_n^{-1}$. In the following, therefore, it suffices to show that for each $n$, there is an $x_n \in I_n$ such that for every $\delta_n > 0$, there are a $c_n$ and a $d_n$ that satisfy the first two conditions of Definition 4 and, yet, they also satisfy

$$\frac{p(x_1, F_1, v_1)}{p(x_1, F_1, u_1)} < \frac{p(x_2, F_2, v_2)}{p(x_2, F_2, u_2)}. \tag{33}$$

Since $(u_1, v_1) \not\blacktriangleright (u_2, v_2)$, for each $n$, there is an $x_n \in I_n$ such that

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

Write

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

Let $\delta_n$ be any positive number smaller than the $\delta_n$ that is obtained in Lemma 1 for $\varepsilon/2$. Apply Lemma 2 to $x_n$ and $\delta_n$ to obtain a consumption plan $c_n^{(\underline{x}_n, \overline{x}_n)}$, which is defined in (30). Let $e_n$ be the second-order consumption plan associated with $c_n^{(\underline{x}_n, \overline{x}_n)}$, which is defined in (17). Denote by $F_n$ the cumulative distribution function of $\mu_n \circ e_n^{-1}$. Then $F_n \in \mathscr{F}(x_n, \delta_n)$ and

$$\frac{p(x_1, F_1, v_1)}{p(x_1, F_1, u_1)} < \frac{a(x_1, v_1)}{a(x_1, u_1)} + \frac{\varepsilon}{2} = \frac{a(x_2, v_2)}{a(x_2, u_2)} - \frac{\varepsilon}{2} < \frac{p(x_2, F_2, v_2)}{p(x_2, F_2, u_2)}.$$

This implies (33) and completes the proof. ///

# 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.