KIER DISCUSSION PAPER SERIES

KYOTO INSTITUTE
OF
ECONOMIC RESEARCH

Discussion Paper No.1097

"Arrow-Pratt-Type Measure of Ambiguity Aversion"

Chiaki Hara

September 2023

KYOTO UNIVERSITY

KYOTO, JAPAN

Arrow-Pratt-Type Measure of Ambiguity Aversion

Chiaki Hara*†

September 8, 2023

Abstract

We define a measure of ambiguity aversion for ambiguity-averse utility functions in a way analogous to the Arrow-Pratt measure of risk aversion. The measure is determined by the second Peano derivative, which exists even for non-differentiable functions, such as maximin and Choquet expected utility functions. Unlike the standard notion of comparative ambiguity aversion, it allows us to compare ambiguity aversion between two utility functions exhibiting different risk attitudes. We introduce a notion of ambiguity premium and show that our measure is related to the second-order, as opposed to the first-order, ambiguity premium. We also show that it is related to the first-order impact on matching probabilities of the size of prizes.

JEL Classification Codes: C38, D81, G11.

Keywords: Expected utility functions, risk aversion, ambiguity aversion, ambiguity premium, matching probabilities, Peano derivative.

1 Introduction

A decision maker in the face of uncertainty is said to perceive ambiguity in an event if he cannot attach any single reliable probability to the event, and, thus,

*Institute of Economic Research, Kyoto University. Email: hara.chiaki.7x@kyoto-u.ac.jp

†Some of the contents of this paper were included in the papers circulated under the title "Comparative Ambiguity Aversion for Smooth Utility Functions," "Measure of Ambiguity Aversion," and so on. The research is funded by the Open Research Area (ORA) for the Social Sciences "Ambiguity in Dynamic Environments," the JSPS KAKENHI Grant Number 17K18561, the Trust Fund Foundation of Japan, Nomura Foundation, and the Joint Usage/Joint Research Program at the Institute of Economic Research of Kyoto University. I am grateful to Pierpaolo Battigali, Simone Cerreia-Vioglio, Mira Frick, Yuhki Hosoya, Ryota Iijma, Tomoki Inoue, Peter Klibanoff, Fabio Maccheroni, Massimo Marinacci, Sujoy Mukerji, Jonathan Newton, Yusuke Osaki, Frank Riedel, Takashi Ui, Koji Shirai, Norio Takeoka, Jan Werner, Michael Zierhut, and Peter Wakker for their insightful comments.

needs to take multiple probabilities into consideration for decision making. If the decision maker tends to prefer state-contingent consequences whose distributions are independent of the choice of these relevant probabilities to those whose distributions are not, then he is said to be ambiguity-averse. Ambiguity aversion has been well documented ever since the experiments by Ellsberg (1961). It has been given a theoretical foundation in the form of axiomatization by Schmeidler (1989), Gilboa and Schmeidler (1989), and their followers. Its implications on portfolio choice and asset pricing have been explored by Dow and Werlang (1992), Epstein and Wang (1994), and many others. More recent studies, such as Ju and Miao (2011), Chen, Ju, and Miao (2014), Jahan-Parvar and Liu (2014), and Gallant, Jahan-Parvar, and Liu (2019), and Collard, Mukerji, Sheppard, and Tallon (2018) have shown not only that introducing ambiguity aversion solves otherwise unexplainable puzzles, such as the home bias puzzle, but also how much ambiguity aversion is needed to solve such puzzles. A typical quantitative exercise there is to fix a form of the decision maker's ambiguity-averse utility function and, then, infer values of ambiguity aversion parameters necessary to solve the puzzle based on experimental evidence or market data.

While these quantitative exercises undoubtedly help us grasp a better understanding of the role of ambiguity aversion in decision making, their approach has a problem. To see what it is, note that inference of the ambiguity aversion parameter often comes with that of the risk aversion parameter. If two pairs of inferred or calibrated values of risk and ambiguity aversion parameters, say (θ_1, γ_1) and (θ_2, γ_2), involve two different values of risk aversion parameter, $\theta_1 \neq \theta_2$, then, even when one ambiguity aversion parameter is greater than the other, $\gamma_1 > \gamma_2$, we cannot conclude that the decision maker (θ_1, γ_1) is more ambiguity-averse than the decision maker (θ_2, γ_2). This is because the notion of comparative ambiguity aversion employed by Epstein (1999, Section 2.3), Ghirardato and Marinacci (2002, Definition 7), and the subsequent literature is restricted the case where the two decision makers share the same risk attitudes, with a couple of exceptions to which we will come back later. Let us illustrate the limitation that the restriction imposes on the analysis in the setting of the home bias puzzle.

Imagine that there are two Japanese investors of equal wealth, say 10 million yen, who invest in a Japanese stock, an American stock, and the bond. Suppose that they perceive no ambiguity in the return on the Japanese stock, but they perceive ambiguity in the return on the American stock. Their portfolios are presented in Table 1. That is, investor 1 puts just a half of the money into the American stock (1 million yen) that he invests into the Japanese stock (2 million yen), while investor 2 allocates the same amount of money in the two stocks (4 million yen each). Then it seems reasonable to think of investor 1 as being more ambiguity-averse than investor 2. However, since the total amount that he invests in the two stock is 3 million yen, which is much less than the total amount (8 million yen) that investor 2 invests in the two stocks, it seems equally reasonable to think of investor 1 as being more risk-averse than investor 2.[1] But, if this is indeed the case, then we cannot say that investor 1 is more ambiguity-averse than investor 2, because we could do so only when they are equally risk-averse according to the standard notion of comparative ambiguity aversion.[2] In other words, we have no theoretical foundation to attribute the difference in the allocation of wealth between the two stocks to the difference in the degree of the investors' ambiguity aversion.

Table 1: Two Japanese investors with equal wealth invest into a Japanese stock, an American stock, and the bond. The figures are in million yen.

	Investor 1	Investor 2
Japanese stock	2	4
American stock	1	4
bond	7	2
total	10	10

The purpose of this paper is to introduce a measure of ambiguity aversion along the lines of Arrow-Pratt measure of risk aversion[3] under the assumption that the risk attitude can be represented by an expected utility function. It allows us to compare ambiguity attitudes of two decision makers who do not have the same risk attitudes. In Table 1, for example, the measure offers the scope for concluding that investor 1 is more risk-averse and ambiguity-averse than investor 2. It also allows us to compare ambiguity attitudes of two decision makers whose

[1] It was also found in the experiments by Bossaerts et al (2010) that those who invest more in stocks (Arrow securities) invest proportionally more in stocks whose returns are ambiguous than in stocks whose returns are unambiguous.

[2] Bossaerts et al (2010) postulated that the investors have α-maximin expected utility functions of Ghirardato, Maccheroni, and Marinacci (2004) and used the inferred values of α alone to compare their degrees of ambiguity aversion with no reference to risk aversion.

[3] Sujoy Mukerji kindly suggested the title of the paper.

ambiguity attitudes are represented by utility functions of different forms. To see this second point, recall that the literature on ambiguity has endeavored to clarify the relationship between particular forms of utility functions and particular (less demanding) variants of the independence axiom in the expected utility theory. Within a class of ambiguity-averse utility functions of the same form, equivalent conditions have been given, in term of parameters in the common functional form, for one decision maker to be more ambiguity-averse than another.[4] Since our measure of ambiguity aversion can be defined without relying on any particular functional forms, it can also be used to compare two decision makers who are ambiguity-averse but satisfy, say, different variants of the independence axiom.

A technical contribution of this paper is to use second Peano derivatives[5] to accommodate some important classes of ambiguity-averse utility functions, most notably Choquet expected utilities of Schmeidler (1989) and maximin expected utilities of Gilboa and Schmeidler (1989), which are not even differentiable in the standard sense. The second Peano derivative is defined as twice the coefficient of the second-order term in the Taylor approximation up to the second order. It does exist under fairly weak conditions and is sufficient for our purpose, because we are interested not in the second derivative of the utility function per se (which is defined as the derivative of derivatives) but in the second-order approximation that it provides.

While our measure of ambiguity aversion is defined in terms of utility functions, it can be characterized in terms of choice behavior, such as matching probabilities and ambiguity premiums. The matching probability is defined as follows. Let x be a baseline consumption level and ε be the value of the prize of a bet on an event A. Then, the decision maker consumes $x+\varepsilon$ on A and x outside A. The matching probability $\rho(A)$ is defined so that the decision maker is indifferent between this binary act and the lottery in which $x+\varepsilon$ is obtained with probability $\rho(A)$ and x is obtained with probability $1-\rho(A)$. In general, the matching probabilities depend on the choice of x and ε. We will show (in Section 5) that the marginal change in matching probabilities due to a marginal increase in prizes ε is approximately proportional to our measure of ambiguity aversion. The ambiguity premium is

[4]The α-maximin expected utility functions mentioned in Footnote 2 constitute such an example.

[5]Mukhopadhyay (2012) is a textbook giving some definitions and facts on Peano derivatives. Evans and Weil (1981-82) is an accessible survey. They are both concerned only with the univariate case. Massimo Marinacci kindly pointed out to me that the notion of second right-derivatives that I introduced in an earlier version was, in fact, a multivariate version of the Peano derivative.

defined as the maximum consumption level that the decision maker is willing to give up in exchange for the assurance that the true probability coincides with his benchmark probability in the sense of Ghirardato and Marinacci (2002). Define the risk premium as the maximum consumption level the decision maker is willing to give up in exchange for full insurance under the benchmark probability. We will show (in Section 6) that our measure of ambiguity aversion is approximately equal to the ratio of the ambiguity premium to the risk premium.

Let us now turn to some earlier works that are most relevant to this one. Dimmock, Kouwenberg, and Wakker (2016, Section 3) defined a measure of ambiguity aversion as the shortfall of the matching probabilities $\rho(A)$ defined for various events A from the subjective probabilities that he would attach to unambiguous acts.[6] They assumed that for preferences in the source method, $\rho(A)$ is uniquely determined by A, independently of the choice of the baseline consumption level x and the prize ε.[7] This uniqueness property holds also for all biseparable preferences in the sense of Ghirardato and Marinacci (2002, Definition 2), including those represented by Choquet expected utilities and α-maximin expected utilities of Ghirardato, Maccheroni, and Marinacci (2004).

Wang (2019) used an ingenious idea of restricting preference comparison on a set of "aligned" acts to give a notion of a decision maker as being more ambiguity-averse than another that neither requires nor implies common risk attitudes. His notion relies on matching probabilities. But, for non-biseparable preferences, such as the smooth ambiguity model of Klibanoff, Marinacci, and Mukerji (2005, hereafter KMM), the matching probabilities depend on the choice of the baseline consumption level x and the prize ε; and so does Wang's definition (Definition 5 of his paper) of the more-ambiguity-averse-than relation. This fact suggests that for general preferences, the matching probability, as well as the measure of ambiguity aversion, should be amended by specifying the baseline consumption level x and

[6]In their setting, the decision maker may use a probability weighting function to calculate expected utility levels. When he does, the subjective probabilities here should be replaced by the values of the probability weighting function taken at the subjective probabilities.

[7]If the subjective probability of an event A is denoted by $P(A)$, then the shortfall is equal to $P(A) - \rho(A)$. In the case of differentiable utility functions, P is defined as (a scalar multiple of) the gradient of the utility function at a constant act. In some experiments, as in the case of Ellsberg's two-color urn, it is defined based on some type of symmetry consideration (exchangeability). For natural, non-experimental events, Abdellaoui, Baillon, Placido, and Wakker (2011, Section IV) offered a way to identify it. Denoting the complement of A by A^c, the shortfall of $\rho(A) + \rho(A^c)$ from one can also be taken as a measure of ambiguity aversion; and it can be identified even when P is unknown. This method of eliciting ambiguity aversion without identifying P was explored by Baillon, Huang, Selim, and Wakker (2018) and Baillon, Bleichrodt, Li, and Wakker (2021)

the prize ε. This is precisely what we will do in this paper.

Cerreia-Vioglio, Maccheroni, and Marinacci (2022) gave a quadratic approximation of the uncertainty premium for ambiguity-averse utility functions. They assumed twice continuous Gateaux differentiability, thereby excluding maximin and Choquet expected utility functions. They characterized the quadratic approximation of the uncertainty premium with a possibly infinite state space, but did not discuss its relation to matching probabilities or other characteristics of preferences, such as the curvatures of indifference curves.

The rest of this paper is organized as follows. Section 2 gives the setup of the paper. Section 3 defines a measure of ambiguity aversion. Section 4 characterizes the measure of ambiguity aversion in terms of utility functions and the curvatures of indifference curves. Section 5 relates it to matching probabilities. Section 6 relates it to ambiguity premiums. Section 7 shows how it can be rewritten for the smooth ambiguity model of KMM, and presents a work-out example of the home bias puzzle. All the results so far are obtained under the assumption of twice continuous differentiability in the standard sense. Section 8 introduces the notion of Peano differentiability, extends the measure of ambiguity aversion to non-differentiable but twice Peano differentiable utility functions, and generalizes the results on ambiguity premiums and matching probabilities in the preceding sections. Section 9 suggests directions of future research. The appendix gathers proofs for the main results in the twice continuously differentiable case. The online appendix consists of three sections. The first one contains proofs for the application to the smooth ambiguity model. The second one presents an extension of the measure of ambiguity aversion and its relation to matching probabilities and ambiguity premiums to the case where the decision maker's risk attitude cannot be represented by any expected utility function.[8] The last one gathers lemmas and and the extension to the non-differentiable case.

2 Setup

The state space is a finite set S. By an abuse of notation, we also write its cardinality as S. As it is finite, the σ-field is the power set and every random variable can be regarded as a vector in $\boldsymbol{R}^S$. Denote by e the vector in $\boldsymbol{R}^S$ whose coordinates are all equal to one. It represents a profile of state-independent consumption or utility

[8] Peter Wakker encouraged me to work on this extension.

levels. Write $\Delta = \{p \in \boldsymbol{R}^S_+ \mid p \cdot e = 1\}$. Then, Δ is the set of all probabilities on the state space. Each $g \in \boldsymbol{R}^S$ can be regarded as a random variable on the state space, and its expectation under $p \in \Delta$, $E^p[g]$ is equal to the dot product $p \cdot g$.

Let T be a nonempty open interval of $\boldsymbol{R}$. It is the set of all possible consumption levels. We assume that the decision maker has an expected utility function for pure risk (without ambiguity) and denote by $v : T \to \boldsymbol{R}$ his Bernoulli utility function, following the terminology of Mas-Colell, Whinston, and Green (1995). Let $I : v(T)^S \to \boldsymbol{R}$ and assume that I is increasing and normalized, that is, $I(ye) = y$ for every $y \in v(T)^S$. We refer to I as the aggregator.

Let $\Pi(T)$ be the set of all (Borel) probability measures on T. For each $f : S \to \Pi(T)$, or $f \in \Pi(T)^S$, denote by $v \circ f$ the vector in $v(T)^S$ whose s-th coordinate is equal to the expected utility level $\int_T v(x)\,\mathrm{d}(f(s))(x)$ given by the lottery $f(s) \in \Pi(T)$. Define $V : \Pi(T)^S \to \boldsymbol{R}$ by letting $V(f) = I(v \circ f)$ for every $f \in \Pi(T)^S$. This is the form of utility functions that we study in this paper. We then say that V is defined by (S, T, v, I), or, simply, by (v, I). By Proposition 1 of Cerreia-Vioglio, Ghirardato, Maccheroni, Marinacci, and Siniscalchi (2011), such an I exists for each transitive, complete, continuous, and increasing preference relation that satisfy the independence axiom on the set of lottery acts, to be defined shortly.

An element of $\Pi(T)^S$, which assigns a lottery to each state, is known as an Anscombe-Aumann act (named after Anscombe and Aumann (1963)). An Anscombe-Aumann act f for which $f(s)$ is independent of s is called a lottery act, and the set of lottery acts can be identified with $\Pi(T)$. Since the aggregator I is normalized, the preference restricted on $\Pi(T)$ is represented by an expected utility function with a Bernoulli utility function v. The subjective expected utility is the case where I is the expectation operator under some probability on S. In contrast, an element of T^S assigns a (deterministic) consumption level to each state and is termed a monetary act by Marinacci, Maccheroni, and Rustichini (2006, Section 3.5). We call an act a constant act if it a lottery act and, at the same time, a monetary act. That is, a constant act gives a common deterministic consumption level over all states. The set of constant acts can thus be identified with T. The measure of ambiguity aversion we introduce in this paper is a local measure that represents ambiguity attitudes around constant acts.

While the use of Anscombe-Aumann acts in the theory of choice may be questionable, as voiced by Kreps (1988, Chapter 7) and Epstein (1999, Section 5), we have opted for the domain of Anscombe-Aumann acts, rather than that of monetary acts, because, then, V specifies risk attitudes by ranking lottery acts. Once the Bernoulli utility function v is known, the subsequent analysis is valid (and simpler) even when the utility function V is restricted to the set of monetary acts. Bear in mind, though, that the separation between the Bernoulli utility function v and the aggregator I is incomplete on the set of monetary acts: It is possible that a pair (v_1, I_1) defines V_1, another, different, pair (v_2, I_2) defines V_2, and, yet, $V_1 = V_2$ on T^S. Since our measure of ambiguity aversion depends on how V is decomposed into v and I, we have chosen the set of Anscombe-Acts as the domain of the utility function V to circumvent the problems arising from this inseparability.

We assume throughout this paper that v is twice continuously differentiable and satisfies $v'' \leq 0 < v'$. Then $v(T)$ is a nonempty open interval of $\boldsymbol{R}$. We also assume that $v'' < 0$, except in Section 5, where we explore the relationship of our measure of ambiguity aversion to matching probabilities. In addition, we assume that I is twice continuously differentiable until Section 8, where we deal with the case of non-differentiable aggregators. The assumption of twice continuous differentiability is violated by the maximin expected utility of Gilboa and Schmeidler (1989) and Choquet expected utility of Schmeidler (1989), but satisfied by the smooth ambiguity-averse utility functions of KMM and the relative entropy used by Hansen and Sargent (2001), among others.

3 Measure of ambiguity aversion

In Section 2, we defined a utility function $V : \Pi(T)^S \to \boldsymbol{R}$ based on a Bernoulli utility function $v : T \to \boldsymbol{R}$ and a aggregator $I : v(T)^S \to \boldsymbol{R}$. In this section, we define a measure of ambiguity aversion that is analogous to the Arrow-Pratt measure of risk aversion for Bernoulli utility functions.

Let $x \in T$ and write $p = \nabla I(v(x)e)$. Since I is non-decreasing and normalized, $p \in \Delta$. For each $z \in \boldsymbol{R}^S$, denote the variance of z under p by $\mathrm{Var}^p[z]$. It is equal to $\sum_s p(s)z(s)^2 - \left(\sum_s p(s)z(s)\right)^2$ and, thus, to $E^p[z^2] - (E^p[z])^2$, where $z^2 = ((z(s))^2)_s \in \boldsymbol{R}^S$.

Definition 1 Let $x \in T$ and $z \in \boldsymbol{R}^S$. Write $p = \nabla I(v(x)e)$ and suppose that

$\mathrm{Var}^p[z] > 0$. Then, we define

$$H^x(z) = \frac{-\dfrac{z^\top \left(v'(x)\nabla^2 I(v(x)e)\right) z}{\mathrm{Var}^p[z]}}{-\dfrac{v''(x)}{v'(x)}}. \tag{1}$$

The denominator of the right-hand side of (1) is nothing but the Arrow-Pratt measure of risk aversion of the Bernoulli utility function v at x. The numerator is determined, in part, by the Hessian of the aggregator I at $v(x)e$. In the case of a subjective expected utility function, I is linear, the Hessian is zero, and our measure is also equal to zero. It can, thus, be considered as representing how much the utility function V is different from the subjective expected utility functions.

The denominator of the right-hand side of (1) is invariant to the affine transformations of v. By Proposition 1 of Cerreia-Vioglio, Ghirardato, Maccheroni, Marinacci, and Siniscalchi (2011), the numerator is also invariant as long as I is transformed along with an affine transformation of v so that I is normalized and V represents the same preference. Thus, the value of $H^x(z)$ is uniquely determined by the risk attitudes and the preference relation on T^S. The measure $H^x(z)$ is invariant to scalar multiplications and addition of constant acts to z. Hence, to identify the values of $H^x(z)$ for all z, it suffices to know the values of $H^x(z)$ for the z with zero mean and unit variance.

A curious aspect of the definition (1) is that while the ambiguity aversion seems fully captured by the numerator of the right-hand side, the definition also involves the denominator, which is determined solely by the risk aversion. There are two reasons for employing this definition. First, as we will see in Sections 5, 6, and 8, this definition allows us to obtain tractable asymptotic results on ambiguity premiums and matching probabilities, even when I is not differentiable. Second, in many applications of ambiguity-averse utility functions, most notably in the home bias puzzle, which we will take up in Section 7.2, the ambiguity aversion in excess of the risk aversion is more important than the ambiguity aversion itself. In any case, we will give equivalent expressions of the numerator in terms of derivatives of ambiguity premiums and matching probabilities in the subsequent analysis. They may well be helpful for the potential use of the numerator, on its own, as a measure of ambiguity aversion.

The measure of ambiguity aversion can be best interpreted in terms of benchmark preferences and benchmark measures in the sense of Ghirardato and Marinacci (2002, Section 2.2). All subjective expected utility functions are ambiguity-neutral, and suppose that there is one that is less ambiguity-averse than V according to their Definition 7. Such a subjective expected utility function is called a benchmark preference of V, and the subjective probability is called a benchmark measure. As discussed in the introduction, a benchmark preference exhibits the same risk attitudes as V, which is represented by v. Since the aggregator I is (twice continuously) differentiable, the benchmark measure, which we will also refer to as the benchmark probability, is unique and coincides with $\nabla I(v(x)e)$ for any $x \in T$.[9] Hence, the measure of ambiguity aversion is equal to the quadratic form of the Hessian of I divided by the Arrow-Pratt measure of risk aversion and the variance of deviations from a constant act relative to its benchmark preference. In the rest of this paper, we sometimes refer to $\nabla I(v(x)e)$ as the benchmark probability, even when there is no benchmark preference, as a shorthand for the subjective probability implicit at the constant act xe.

4 Utility functions and indifference curves over monetary acts

In this section, we give an equivalent expression of the measure of ambiguity aversion, (1), in terms of the utility function V. When its domain is restricted to the set of monetary acts, since it is a subset of $\boldsymbol{R}^S$, we can apply the classical demand theory to find its illustrative properties.

Let $x \in T$ and $p = \nabla I(v(x)e)$. Let $\bar{I}$ be the expectation operator under p and $\bar{V}$ be defined by $(v, \bar{I})$. Define $V^* = v^{-1} \circ V$ and $\bar{V}^* = v^{-1} \circ \bar{V}$. Then, V^* represents the same preference as V, and $\bar{V}^*$ represents the same preference as $\bar{V}$. Moreover, if V^* is concave, then it is a least concave utility function in the sense of Debreu (1976),[10] and the same can be said for $\bar{V}^*$. In the following theorem, the utility

[9] When there is a benchmark preference, $\nabla I(v(x)e)$ is independent of the choice of $x \in T$. This property is equivalent to Axiom 7, Translation Invariance at Certainty, of Rigotti, Shannon, and Strzalecki (2008), as their Proposition 8 showed.

[10] We need to assume here that V^* is concave because it need not be so under the set of assumptions given so far. Additional assumptions on v necessary to guarantee concavity are given in Hardy, Littlewood, and Polya (1952, Section 3.16). In fact, then, a property stronger than least concavity can be obtained: Every concave function that represents the same preference as V is a concave transformation of V^*. This follows from the fact that V^* is linear on the set of constant acts and least concave utility functions are affine transformations of each other.

functions V^* and $\bar{V}^*$ should be thought as being defined on the set T^S of monetary acts and their Hessians as $S \times S$ matrices.

Theorem 1 *For all $x \in T$ and $z \in \boldsymbol{R}^S$,*

$$z^\top \nabla^2 V^*(xe) z = z^\top \left(v'(x) \nabla^2 I(v(x)e) \right) z + \frac{v''(x)}{v'(x)} \mathrm{Var}^p[z]. \tag{2}$$

Since $\bar{V}$ is a subjective expected utility function under p, a well known result of Pratt (1964) implies that

$$z^\top \nabla^2 \bar{V}^*(xe) z = \frac{v''(x)}{v'(x)} \mathrm{Var}^p[z]. \tag{3}$$

Thus, this theorem implies the following equivalent expression of the measure of ambiguity aversion.

Corollary 1 *For all $x \in T$ and $z \in \boldsymbol{R}^S$ with* $\mathrm{Var}^p[z] > 0$*,*

$$H^x(z) = \frac{z^\top \nabla^2 V^*(xe) z - z^\top \nabla^2 \bar{V}^*(xe) z}{z^\top \nabla^2 \bar{V}^*(xe) z}. \tag{4}$$

That is, $H^x(z)$ can be obtained by dividing the difference between two quadratic forms $z^\top \nabla^2 V^*(xe) z$ and $z^\top \nabla^2 \bar{V}^*(xe) z$ by the latter. Thus, the measure of ambiguity aversion is the ratio of the difference in the second-order effects on least concave utility levels between the preferences represented by V and $\bar{V}$ to that of the preference represented by $\bar{V}$.

A simple but useful insight can be obtained from Corollary 1 by applying the second-order Taylor approximation. Note that $\nabla V^*(xe) z = \nabla \bar{V}^*(xe) z = E^p[z]$ and assume that $E^p[z] = 0$. Then,

$$\frac{(V^*(xe) - V^*(xe + \varepsilon z)) - (\bar{V}^*(xe) - \bar{V}^*(xe + \varepsilon z))}{\bar{V}^*(xe) - \bar{V}^*(xe + \varepsilon z)} \to H^x(z)$$

as $\varepsilon \to 0$. Since $V^*(xe + \varepsilon z) < V^*(xe)$ and $\bar{V}^*(xe + \varepsilon z) < \bar{V}^*(xe)$, the asymptotic result says that the reduction is larger for V^* than for $\bar{V}^*$, by the factor approximately equal to $H^x(z)$.

We now give yet another equivalent expression of the measure of ambiguity aversion in terms of indifference curves, which admits a particularly illuminating graphical representation. For each $z \in \boldsymbol{R}^S$ with $E^p[z] = 0$ and $\mathrm{Var}^p[z] > 0$, let

$L(z)$ be the plane spanned by e and z. Then, $T^S \cap L(z)$ is an open subset of $L(z)$ that contains xe for every $x \in T$. Then the intersection of $L(z)$ with the indifference hypersurfaces that contain xe, $\{f \in T^S \cap L(z) \mid V(f) = V(xe)\}$ and $\{f \in T^S \cap L(z) \mid \bar{V}(f) = \bar{V}(xe)\}$, are twice continuously differentiable curves on $L(z)$. Denote their curvatures by $c^x(z)$ and $\bar{c}^x(z)$.

Theorem 2 *For every $x \in T$ and every $z \in \boldsymbol{R}^S$ with $E^p[z] = 0$ and $\mathrm{Var}^p[z] > 0$,*

$$c^x(z) = -S^{1/2}\left(\frac{z^\top \left(v'(x)\nabla^2 I(v(x)e)\right) z}{\mathrm{Var}^p[z]} + \frac{v''(x)}{v'(x)}\right),$$
$$\bar{c}^x(z) = -S^{1/2}\frac{v''(x)}{v'(x)}$$

This theorem immediately gives the following equivalent expression for $H^x(z)$.

Corollary 2 *For every $x \in T$ and every $z \in \boldsymbol{R}^S$ with $E^p[z] = 0$ and $\mathrm{Var}^p[z] > 0$,*

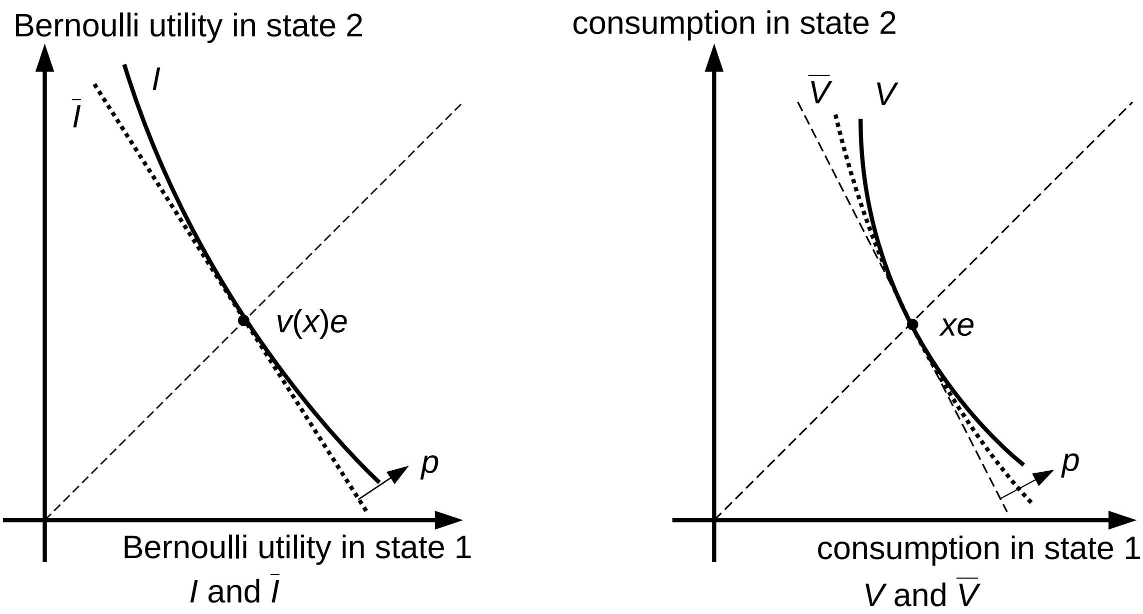
$$H^x(z) = \frac{c^x(z) - \bar{c}^x(z)}{\bar{c}^x(z)}.$$

Since the curvature of a curve at a point measures how much curved it is around the point, this corollary that the measure of ambiguity aversion represents how much more curved, in proportion, the indifference curves of V is than that of $\bar{V}$ in the direction z of deviation at the constant act xe.

We can now illustrate the measure of ambiguity aversion in terms of the indifference curves of the utility function V on the set of monetary acts along the lines of Yaari (1969). Let $S = 2$. Let $x \in T$ and consider the constant act xe. The corresponding vector of Bernoulli utility profiles, $v \circ (xe)$, coincides with $v(x)e$. The indifference curves of I and $\bar{I}$ that go through $v(x)e$ are shown on the left panel of Figure 1. The indifference curve of I is convex to the origin, which represents ambiguity aversion, while that of $\bar{I}$ is a straight line with normal vector $p = \nabla I(v(x)e)$. These two indifference curves correspond to the two indifference curves of V and $\bar{V}$ that go through the constant act xe on the right panel. Since v is concave, both are convex to the origin, but the indifference curve of V is more convex to the origin than the indifference curve of $\bar{V}$ due to the concavity of I. How much more convex the former is than the latter is represented by their curvatures; and Corollary 1 shows that the measure of ambiguity aversion represents how much more curved, in proportion, the indifference curve of V is than that of

$\bar{V}$.

Figure 1: Measure of ambiguity aversion represented on the indifference curves of V and $\bar{V}$.



There are only two states ($S = 2$). On the left panel, the solid curve is the indifference curve of I that goes through the (Bernoulli) utility profile $v(x)e$ of the constant act xe, and the dotted line is the indifference curve of $\bar{I}$ that goes through the same profile. Since v is concave, these indifference curves are transformed into more concave ones on the set (plane) of monetary acts on the right panel, but they are still tangent to the line with normal vector $p = \nabla I(v(x)e)$. The measure of ambiguity aversion quantifies how much more curved, in proportion, the indifference curves of V is at xe than that of $\bar{V}$.

5 Matching probabilities

In Section 3, we defined our measure of ambiguity aversion, (1), in terms of the Hessians of aggregators. In this section, we explore its behavioral characterization. Specifically, we show that it is equal to the ratio of the derivative of the matching probability and its deviation from the derivative of the fictitious matching probability, which is derived from a utility function that coincides with the utility function V on the set T^S of monetary acts but exhibits risk neutrality. We then give an asymptotic result, establishing that it is approximately equal to the ratio of the matching probability and its deviation from the fictitious matching probability when acts are almost constant.

Let $A \subset S$ and $x \in T$. Denote by e_A the indicator function of A, that is, $e_A(s) = 1$ for every $s \in A$ and $e_A(s) = 0$ otherwise. For each $\varepsilon > 0$, εe_A represents a bet on the event A with the prize ε, and $xe + \varepsilon e_A$ is a binary act that represents the consumption plan for a decision maker who holds a benchmark consumption level x and places the bet. For each $r \in [0, 1]$, denote by $f^x(\varepsilon, r) \in \Pi(T)$ the lottery that assigns probability r to $x + \varepsilon$ and $1 - r$ to x. Then, regarding $f^x(\varepsilon, r)$ as a lottery act, define $\rho^x(\varepsilon, A)$ by

$$V\left(f^x(\varepsilon, \rho^x(\varepsilon, A))\right) = V(xe + \varepsilon e_A). \tag{5}$$

Thus, (5) means that the decision maker is indifferent between the lottery with the winning probability $\rho^x(\varepsilon, A)$ and the bet on A. The winning probability $\rho^x(\varepsilon, A)$ is called the matching probability of the event A at the consumption level x with the prize ε, because the equality means that the probabilistic assessment the decision maker gives to the winning event A is equal to $\rho^x(\varepsilon, A)$. Since V is an expected utility function on the set of lottery acts, (5) can be rewritten as

$$\rho^x(\varepsilon, A)v(x + \varepsilon) + (1 - \rho^x(\varepsilon, A))v(x) = V(xe + \varepsilon e_A). \tag{6}$$

This shows that $\rho^x(\varepsilon, A)$ is a continuously differentiable function of $\varepsilon > 0$.

Even when $\varepsilon < 0$ (and $x + \varepsilon \in T$), we define $\rho^x(\varepsilon, A)$ as in (5) and (6). In this case, since $xe + \varepsilon e_A = (x + \varepsilon)e - \varepsilon e_{S\setminus A}$, the bet is on the complementary event $S \setminus A$ with the prize $-\varepsilon$, and the decision maker's initial deterministic consumption level is $x + \varepsilon$. It can then be shown that

$$\rho^x(\varepsilon, A) = 1 - \rho^{x+\varepsilon}(-\varepsilon, S \setminus A), \tag{7}$$

where, since $-\varepsilon > 0$, the right-hand side was defined as in the previous paragraph. This shows that $\rho^x(\varepsilon, A)$ is a continuously differentiable function of $\varepsilon < 0$.

For $\varepsilon = 0$, we set $\rho^x(0, A) = p(A)$, where $p(A) = \nabla I(v(x)e)e_A$. Thus $p(A)$ is the probability of A under the subjective probability implicit at the constant act xe. We have thus defined the function $\varepsilon \mapsto \rho^x(\varepsilon, A)$ on an open interval containing 0. If V represents a subjective expected expected utility function, then $\rho^x(\varepsilon, A) = 0$ for every ε and $\partial\rho^x(0,A)/\partial\varepsilon = 0$. The next theorem asserts, more generally, that this function is differentiable at 0 and the derivative is equal to half the denominator of $H^x(e_A)$.

Theorem 3 *Let $A \subset S$ and $x \in T$. Then the function $\varepsilon \mapsto \rho^x(\varepsilon, A)$ is differentiable at 0 and*

$$\frac{\partial \rho^x}{\partial \varepsilon}(0,A) = \frac{1}{2} e_A^\top \left(v'(x)\nabla^2 I(v(x)e)\right) e_A.$$

To obtain an expression of the numerator of $H^x(e_A)$ in terms of matching probabilities, write $\underline{I} = v^{-1} \circ V$. Then, $\underline{I}$ gives the certainty equivalents in terms of constant acts. It is defined on $\Pi(T)^S$ but think of its domain as restricted on the set T^S of monetary acts for the subsequent analysis. Then, it is continuous, increasing, and normalized. Let $\underline{v}$ be the identity function on T and define a utility function $\underline{V}$ by $(\underline{v}, \underline{I})$.[11] Since $\underline{v} \circ f = f$ and, hence, $\underline{V}(f) = \underline{I}(\underline{v} \circ f) = v^{-1}(V(f))$ for every $f \in T^S$, $\underline{V}$ represents the same preference as V on the set of monetary acts. But, since $\underline{V}$ is risk-neutral, they differ from each other outside the set of monetary acts. Moreover, any aversion to non-constant acts exhibited by $\underline{V}$ is attributed entirely to ambiguity aversion. Thus, any difference in ambiguity aversion, and hence in matching probabilities, between V and $\underline{V}$ is attributed to the risk aversion represented by the Bernoulli function v. Define $\underline{\rho}^x(\varepsilon, A)$ as in (5) for $\varepsilon \neq 0$ using $\underline{V}$ in place of V, and let $\underline{\rho}^x(0,A) = \nabla \underline{I}(\underline{v}(x))e_A$. Since $\nabla \underline{I}(\underline{v}(x))e_A = \nabla I(v(x))e_A = p(A)$, $\underline{\rho}^x(0,A)$ is equal to $p(A)$. Theorem 3 holds for $(\underline{v}, \underline{I})$. In particular, $\varepsilon \mapsto \underline{\rho}^x(\varepsilon, A)$ is differentiable at 0. By combining $\underline{\rho}^x(\varepsilon, A)$ with $\rho^x(\varepsilon, A)$, we obtain the following expression for $H^x(z)$.

Corollary 3 *Let $A \subset S$ and $x \in T$ and suppose that $0 < p(A) < 1$, then*

$$H^x(e_A) = \frac{\dfrac{\partial \underline{\rho}^x}{\partial \varepsilon}(0,A)}{\dfrac{\partial \underline{\rho}^x}{\partial \varepsilon}(0,A) - \dfrac{\partial \rho^x}{\partial \varepsilon}(0,A)}. \tag{8}$$

The first-order Taylor approximation shows that

$$\frac{p(A) - \rho^x(\varepsilon, A)}{\left(p(A) - \underline{\rho}^x(\varepsilon, A)\right) - \left(p(A) - \rho^x(\varepsilon, A)\right)} \to H^x(z) \tag{9}$$

[11] The Bernoulli utility function $\underline{v}$ violates the assumption of negative second derivatives, stated in Section 2. This, however, will not cause any problem for the analysis in this section.

as $\varepsilon \to 0$. This asymptotic result can be explained as follows. The difference $p(A) - \rho^x(\varepsilon, A)$ is the reduction in the matching probabilities of the event A caused by an increase in prizes of bets on the event. The difference $p(A) - \underline{\rho}^x(\varepsilon, A)$ is also the reduction in the matching probabilities of the event A but measured for the fictitious utility function $\underline{V}$. Since V and $\underline{V}$ represent the same preference for monetary acts but the latter is risk-neutral, the reduction in the matching probabilities of A for $\underline{V}$ is due not only to the ambiguity aversion represented by I but also, erroneously, to the risk aversion represented by v (as it is missed out by $\underline{v}$). Thus the difference $\left(p(A) - \underline{\rho}^x(\varepsilon, A)\right) - (p(A) - \rho^x(\varepsilon, A))$ in the reductions of matching probabilities for $\underline{V}$ and V is due solely to the risk aversion of v. Thus, it can approximate the Arrow-Pratt measure of risk aversion of v. Corollary 3 shows that the ratio of the reduction in the matching probabilities due to ambiguity aversion and the reduction in the fictitious matching probabilities due to risk aversion is approximately equal to our measure of ambiguity aversion.

As any statement on derivatives is hard to test when there are only finitely many observations (such as investors' portfolio choices in asset markets and subjects' responses in laboratory experiments), this asymptotic result may be useful when identifying the decision maker's ambiguity attitude from his choice behavior. The availability of such an asymptotic result is one of the reasons, stated in Section 3, why we include the Arrow-Pratt measure in the definition (1) of the measure of ambiguity aversion.

6 Ambiguity premiums

In this section, we explore another behavioral characterization of the measure of ambiguity aversion. We show that it is equal to the ratio of the second derivatives of the risk and ambiguity premiums. We then give an asymptotic result, establishing that it is approximately equal to the ratio of the risk and ambiguity premiums when acts are almost constant. Finally, we give a graphical illustration of the ambiguity premium.

Let $x \in T$ and $p = \nabla I(v(x)e)$. Let $z \in T^S$. For each ε close to 0, denote by $f^x(\varepsilon, z) \in \Pi(T)$ the lottery that coincides with the distribution on T of the monetary act $x + \varepsilon z$ under the probability p. Then, regarding $f^x(\varepsilon, z)$ as a lottery act, define $\kappa^x(\varepsilon, z)$ by $V\left(f^{x-\kappa^x(\varepsilon,z)}(\varepsilon, z)\right) = V(xe + \varepsilon z)$. That is, $\kappa^x(\varepsilon, z)$ is the maximum consumption level that the decision maker is willing to give up in exchange for the assurance that the true probability coincides with his benchmark probability. It can, thus, be called the ambiguity premium of the act $xe + \varepsilon z$ with respect to the utility function V.

As in Section 3, let $\bar{I}$ be the expectation operator under p and define $\bar{V}$ by $(v, \bar{I})$. Since V coincides with $\bar{V}$ on the set of lottery acts, $\kappa^x(\varepsilon, z)$ can also be defined by

$$\bar{V}\left((x - \kappa^x(\varepsilon, z))\, e + \varepsilon z\right) = V(xe + \varepsilon z). \tag{10}$$

By applying the implicit function theorem to (10), we can see that $\kappa^x(\varepsilon, z)$ is a twice continuously differentiable function of ε. In particular, if V represents a subjective expected utility function, then $\kappa^x(\varepsilon, z) = 0$ for every ε and $\partial^2 \kappa^x(0, z)/\partial\varepsilon^2 = 0$. The following theorem, more generally, relates the numerator of the measure of ambiguity aversion (1) to the second derivative of the ambiguity premium.

Theorem 4 *For all $x \in T$ and $z \in \boldsymbol{R}^S$,*

$$\begin{aligned}
\kappa^x(0, z) &= 0 \\
\frac{\partial \kappa^x}{\partial \varepsilon}(0, z) &= 0, \\
\frac{\partial^2 \kappa^x}{\partial \varepsilon^2}(0, z) &= -z^\top \left(v'(x) \nabla^2 I(v(x)e\right) z.
\end{aligned}$$

In the terminology analogous to that of Segal and Spivak (1990), the equality $\partial \kappa^x(0, z)/\partial\varepsilon = 0$ means that the utility function V is second-order ambiguity-averse.

The Arrow-Pratt measure of absolute risk aversion, which appears in the denominator of the definition (1) of the measure of ambiguity aversion, can be written in terms of the second derivative of the risk premium. This is a well known result due to Pratt (1964, Section 3), but we reproduce the argument here for the subsequent analysis.

Define $\bar{\kappa}^x(\varepsilon, z)$ by $\bar{V}\left((x - \bar{\kappa}^x(\varepsilon, z))\, e\right) = \bar{V}(xe + \varepsilon z)$. Then $\bar{\kappa}^x(\varepsilon, z)$ is the risk premium, that is, $x - \bar{\kappa}^x(\varepsilon, z)$ is equal to the certainty equivalent of the monetary act $xe + \varepsilon z$ under the probability p. Pratt (1964) showed that if $E^p[z] = 0$, then

$$\frac{\partial \bar{\kappa}^x}{\partial \varepsilon}(0, z) = 0, \tag{11}$$

$$\frac{\partial^2 \bar{\kappa}^x}{\partial \varepsilon^2}(0, z) = -\frac{v''(x)}{v'(x)}\mathrm{Var}^p[z]. \tag{12}$$

Hence, Theorem 4 leads to the following characterization of $H^x(z)$ in terms of ambiguity and risk premiums.

Corollary 4 *For all $x \in T$ and $z \in \boldsymbol{R}^S$ with $E^p[z] = 0$,*

$$H^x(z) = \frac{\dfrac{\partial^2 \kappa^x}{\partial \varepsilon^2}(0, z)}{\dfrac{\partial^2 \bar{\kappa}^x}{\partial \varepsilon^2}(0, z)}.$$

By applying the second-order Taylor approximation, we can show that as $\varepsilon \to 0$,

$$\frac{\kappa^x(\varepsilon, z)}{\bar{\kappa}^x(\varepsilon, z)} \to H^x(z) \tag{13}$$

That is, our measure of ambiguity aversion, $H^x(z)$, is approximately equal to the ratio of the ambiguity premium $\kappa^x(\varepsilon, z)$ to the risk premium $\bar{\kappa}^x(\varepsilon, z)$. This corollary will be useful for the quantitative analysis of ambiguity premiums, just as Corollary 3 will be useful for the quantitative analysis of matching probabilities, because any statement on (second) derivatives, in Theorem 3 or 4, is hard to test when there are only finitely many observations.

Before giving a graphical illustration of the ambiguity premium, let us introduce another type of premium, which can be called the uncertainty premium. Define $\breve{\kappa}^x(\varepsilon, z)$ by $V\left(\left(x - \breve{\kappa}^x(\varepsilon, z)\right) e\right) = V(xe + \varepsilon)$. That is, $x - \breve{\kappa}^x(\varepsilon, z)$ is the certainty equivalent of $xe + \varepsilon z$ with respect to the utility function V, which reflects both risk and ambiguity aversion. We could define the part of the uncertainty aversion that cannot be attributed to risk aversion, $\breve{\kappa}^x(\varepsilon, z) - \bar{\kappa}^x(\varepsilon, z)$, as the ambiguity premium. Although this is, in general, not equal to $\kappa^x(\varepsilon, z)$, they are equal up to the first- and second-order. In particular,

$$\frac{\partial^2 \kappa^x}{\partial \varepsilon^2}(0, z) = \frac{\partial^2 \breve{\kappa}^x}{\partial \varepsilon^2}(0, z) - \frac{\partial^2 \bar{\kappa}^x}{\partial \varepsilon^2}(0, z).$$

Hence, Theorem 4 can be restated using the right-hand side of the above equality.

We can now give a graphical representation of the risk premium and the ambiguity premium. Assume that $S = 2$. Let $x \in T$ and write $p = \nabla I(v(x)e)$. Let z represent a deviation such that $E^p[z] = 0$. Let $\varepsilon > 0$. The constant act xe and the monetary act $xe + \varepsilon z$ are depicted on Figure 2. The solid curve is the indifference curve of the utility function V; the upper dotted curve is the indifference curve of the benchmark utility function $\bar{V}$; and the dashed line is the iso-expectation line, all going through $xe + \varepsilon z$. By the definition of the uncertainty premium, the solid indifference curve intersects the 45-degree line at $(x - \breve{\kappa}^x(\varepsilon, z))\, e$; by the definition of the risk premium, the upper dotted indifference curve intersects the 45-degree line at $(x - \bar{\kappa}^x(\varepsilon, z))\, e$; and since $E^p[z] = 0$, the dashed iso-expectation line goes though xe. Thus, the risk premium $\bar{\kappa}^x(\varepsilon, z)$ is the distance between the upper dotted indifference cure and the dashed iso-expectation line, measured along the 45-degree line. The lower dotted curve is the indifference curve of the benchmark utility function $\bar{V}$ that goes through $(x - \breve{\kappa}^x(\varepsilon, z))\, e$. It is tangent to the solid indifference curve, because they are both tangent to the iso-expectation line. It also goes through $(x - \kappa^x(\varepsilon, z))\, e + \varepsilon z$, because

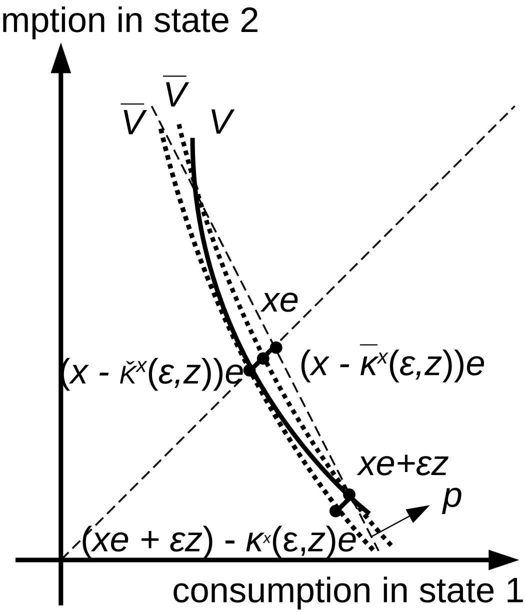
$$\bar{V}\left((x - \breve{\kappa}^x(\varepsilon, z))\, e\right) = V\left((x - \breve{\kappa}^x(\varepsilon, z))\, e\right) = V(xe + \varepsilon z) = \bar{V}\left((x - \kappa^x(\varepsilon, z))\, e + \varepsilon z\right).$$

Thus, the ambiguity premium $\kappa^x(\varepsilon, z)$ is the distance between the upper and lower dotted curves as measured along the 45-degree line going through $xe+\varepsilon z$. Corollary 4 shows that our measure of ambiguity aversion is approximately equal to the ratio of the distance between the dashed iso-expectation line and the upper dotted indifference curve and the distance between the upper and lower dotted indifference curves.

7 Smooth ambiguity model

In this section, we deal with the smooth ambiguity model of KMM, as it is the prime example of twice continuously differentiable aggregators. We show that it admits an insightful equivalent expression of the measure of ambiguity aversion, and, then, give a simple numerical example to illustrate how much ambiguity aversion needs to be introduced to solve the home bias puzzle.

Figure 2: Risk premium, ambiguity premium, and the indifference curves of V and $\bar{V}$.



The solid curve is the indifference curve of the utility function V; the upper dotted curve is the indifference curve of the benchmark utility function $\bar{V}$; and the dashed line is the iso-expectation line, all going through $xe+\varepsilon z$. The measure of ambiguity aversion is approximately equal to the ratio of the distance between the dashed iso-expectation line and the upper dotted indifference curve, measured along the 45-degree line going through the origin versus the distance between the upper and lower dotted indifference curves, measured along the 45-degree line going through $xe + \varepsilon z$.

7.1 Equivalent expression of the measure of ambiguity aversion

Let $v : T \to \boldsymbol{R}$ be a Bernoulli utility function be as in Section 2, $\phi : v(T) \to \boldsymbol{R}$ be a twice continuously differentiable function that satisfy $\phi'' \leq 0 < \phi'$, and μ be a probability measure on Δ such that $\int_\Delta p \, \mathrm{d}\mu(p) \in \boldsymbol{R}^S_{++}$. This is the reduced probability of μ and denoted by p^I. Define $I : v(T)^S \to \boldsymbol{R}$ by letting

$$I(g) = \phi^{-1} \left(\int_\Delta \phi \left(E^p[g] \right) \mathrm{d}\mu(p) \right) \tag{14}$$

for every $g \in v(T)^S$. As in Section 2, define $V : T^S \to \boldsymbol{R}$ by (v, I). Then,

$$\phi(V(f)) = \int_\Delta \phi\left(E^p(v \circ f)\right) \mathrm{d}\mu(p). \tag{15}$$

This is the smooth ambiguity model axiomatized by KMM. Rigotti, Shannon, and Strzalecki (2008, Proposition 5) showed that the reduced probability p^I is its benchmark probability.

Let $w = \phi \circ v$, then (15) can be rewritten as

$$\phi(V(f)) = \int_\Delta w\left(v^{-1}\left(E^p(v \circ f)\right)\right) \mathrm{d}\mu(p).$$

While ϕ is a Bernoulli utility function over expected utility levels, w is a Bernoulli utility function over certainty equivalents. Denote by $E^{\cdot}[z]$ the random variable $p \mapsto E^p[z]$ defined on the probability space $(\Delta, \mathscr{B}(\Delta), \mu)$. Denote by E^μ and Var^μ be the expectation operator and the variance operator for the random variance defined on Δ under the probability μ. In this class of smooth ambiguity-averse utility functions, the measure of ambiguity aversion can be written as follows.

Proposition 1 *If I is defined by* (14)*, then*

$$H^x(z) = \left(\frac{-\dfrac{w''(x)}{w'(x)}}{-\dfrac{v''(x)}{v'(x)}} - 1 \right) \frac{\mathrm{Var}^\mu[E^{\cdot}[z]]}{\mathrm{Var}^{p^I}[z]} \tag{16}$$

for all $x \in T$ and $z \in \boldsymbol{R}^S$ with $\mathrm{Var}^{p^I}[z] > 0$.

This proposition shows that in the case of the smooth ambiguity utility functions, the measure $H^x(z)$ of ambiguity aversion consists of two terms. The first term of (16) represents how much in proportion the decision maker is more averse to the variability in certainty equivalents under different probabilities on the state space than to the variability that is independent of the choice of such probabilities. This term is independent of the second-order belief μ and also of the choice of z. The second term is dependent on the choice of z. Since its numerator is the variance of the random variable $E^{\cdot}[z] : p \mapsto E^p[z]$ under the probability μ on Δ, it measures how much uncertainty is perceived in the distribution of z. The denominator is a normalizing factor, as the law of total variance guarantees that the fraction lies in $[0, 1]$.[12] Since p^I is the reduced probability of the second-order belief, the second term is determined by the second-order belief, and represents how much ambiguity is perceived in the direction the profile z of state-dependent utility levels. Thus, the measure of ambiguity aversion can be decomposed into two parts, the first part determined by the decision maker's inherent aversion to ambiguity, and the second part determined by the decision maker's perception of ambiguity in the environment he is in.[13]

7.2 Application to the home bias puzzle

We now turn to a simple work-out example in which the measure of ambiguity aversion can be inferred from the observed portfolio choice in the home bias puzzle. In the following specification of utility functions, the robust quadratic approximation Maccheroni, Marinacci, and Ruffino (2013) is exact. The general case of an arbitrary number of assets was considered by Hara and Honda (2022).

Three types of assets, the home stock, the foreign stock, and the risk-free bond, can be bought and sold. The state space S coincides with the set of all possible realizations of a bivariate random variable $R = (R_1, R_2)$ of home and foreign stocks returns. The return of the risk-free bond is denoted by R_f.

In the specification (v, w, μ) of a KMM utility function, we assume that v has a constant coefficient θ of absolute risk aversion, and w has a constant coefficient γ of absolute risk aversion, with $\theta \leq \gamma$. Write $\eta = \gamma/\theta - 1$. The second-order belief μ is defined as follows. Let $\sigma > 0$ and $\tau \in (0, 3/4]$, and define a 2×2 matrix $\Sigma_{R|M}$ by

$$\Sigma_{R|M} = \begin{pmatrix} \sigma^2 & (1/2)\sigma^2 \\ (1/2)\sigma^2 & (1-\tau)\sigma^2 \end{pmatrix}.$$

Since $\tau \leq 3/4$, $\Sigma_{R|M}$ is positive semidefinite. Then, for each $m \in \boldsymbol{R}^2$, let the bivariate normal distribution $\mathscr{N}(m, \Sigma_{R|M})$ be a first-order belief, that is, a proba-

[12] This point can be most clearly seen when $z = e_A$, that is, z is the indictor function of an event A. Then, the denominator of the second term is equal to $p^I(A)(1 - p^I(A))$, while the numerator is equal to the variance of $p \mapsto p(A)$ under μ, which does indeed measure the variability of the probability of A due to ambiguity perception. Jewitt and Mukerji (2017) gave a comprehensive analysis on more-ambiguous-than relations between two acts and events.

[13] The phrase, "inherent aversion of ambiguity", has a narrower sense than "ambiguity aversion" in the "measure of ambiguity aversion", as the latter includes the perception of ambiguity but the former does not.

bility measure on the state space.[14] We let the support of the second-order belief be the set of all these bivariate normal distributions. As such, it is parameterized by $m \in \boldsymbol{R}^2$. Let $\bar{R} > R_{\mathrm{f}}$ and write

$$\Sigma_M = \begin{pmatrix} 0 & 0 \\ 0 & \tau\sigma^2 \end{pmatrix}.$$

Let the second-order belief μ in terms of the parameters $m \in \boldsymbol{R}^2$ be the bivariate normal distribution $\mathscr{N}(\bar{R}\mathbf{1}, \Sigma_M)$, where $\mathbf{1} = (1, 1) \in \boldsymbol{R}^2$. Write

$$\Sigma_R = \begin{pmatrix} \sigma^2 & (1/2)\sigma^2 \\ (1/2)\sigma^2 & \sigma^2 \end{pmatrix},$$

then $\Sigma_R = \Sigma_{R|M} + \Sigma_M$ and, by the law of total variance, the joint distribution of $R = (R_1, R_2)$ under the reduced probability of μ, which coincides with the benchmark probability, is $\mathscr{N}\left(\bar{R}\mathbf{1}, \Sigma_R\right)$.

The second-order belief μ has the following features. Under the benchmark probability, the home and foreign stocks share the same expected return $\bar{R}$ and the same variance σ^2, and their correlation coefficient is equal to 1/2. Lewis (1999, Table 2, Panel A), for example, reported that the mean and standard deviation of the annualized monthly returns of the home (US) stock from January 1970 to December 1996 are 11.14% and 15.07%, those of the foreign (Europe, Australia, and Far East, measured in dollars) stocks are 12.12% and 16.85%, and the correlation coefficient between the returns of the two stocks is 0.48. As such, the assumptions that the two stock returns share the same mean $\bar{R}$ and the same variance σ^2 and that their correlation coefficient is 1/2 are not overly unrealistic.

The decision maker perceives no ambiguity in the return of the home stock, but he is unsure of the expected return of the foreign stock. In particular, he perceives that the expected return itself is normally distributed and the variance of the foreign stock return due to this ambiguity has a proportion τ in the (total) variance under the benchmark probability.

Denoting his total wealth by W, we formulate the decision maker's portfolio

[14]This assumption implies that the state space S is infinite and, thus, violates a maintained assumption of this paper. This, however, causes no problem because our analysis is concentrated on the linear subspace spanned by R_1, R_2, and the constant-valued function in the set of all random variables on S.

choice problem as the problem of maximizing

$$\int_{\boldsymbol{R}^2} w \left(v^{-1} \left(\int_{\boldsymbol{R}^2} v(a_0 R_{\mathrm{f}} + a \cdot r) \, \mathrm{d}\mathscr{N}(m, \Sigma_{R|M})(r) \right) \right) \mathrm{d}\mathscr{N}(\bar{R}\mathbf{1}, \Sigma_M)(m), \tag{17}$$

where $r = (r_1, r_2)$, $a = (a_1, a_2)$, and $a \cdot r = a_1 r_1 + a_2 r_2$, by choosing the money amounts invested, $(a_0, a_1, a_2) \in \boldsymbol{R}^3$, subject to the budget constraint $a_0 + a_1 + a_2 \leq W$. We can then derive the following one-to-one relation between the measure of ambiguity aversion and the fraction of the wealth invested in the foreign stock out of the total wealth invested in the two stocks.

Proposition 2 *For every $x \in \boldsymbol{R}$, $H^x(R_1) = 0$ and $H^x(R_2) = \eta\tau$. Moreover, if $(a_0, a_1, a_2) \in \boldsymbol{R}^3$ is the solution to* (17)*, then*

$$\frac{a_2}{a_1 + a_2} = \frac{1}{2\,(1 + \eta\tau)}.$$

The expected consumption level under the benchmark probability is equal to $a_0 R_{\mathrm{f}} + a_1 \bar{R} + a_2 \bar{R}$. The proposition shows that the equalities $H^x(R_1) = 0$ and $H^x(R_2) = \eta\tau$ hold not only for $x = a_0 R_{\mathrm{f}} + a_1 \bar{R} + a_2 \bar{R}$ but also for any other x, and that the measure of ambiguity aversion in the direction of R_2 and the fraction invested in the foreign stock depend on the CARA coefficient γ of the outer Bernoulli utility function v only through $\eta = \gamma/\theta - 1$. This substantiates our claim, made right after the definition (1) of the measure of ambiguity aversion, that the ambiguity aversion in excess of the risk aversion is often more important than the ambiguity aversion itself. The proof shows that of the measure $H^x(R_2) = \eta\tau$, η is equal to the first term, and τ is equal to the second term, on the right-hand side of (16).

The proposition provide us with simple but interesting quantitative implications. First, the fraction of investment in the foreign stock is a decreasing function of the measure of ambiguity aversion, starting from 1/2 in the case of ambiguity neutrality and converging to 0 in the case of unboundedly high ambiguity aversion. Second, by reverting the equality, we can infer the measure of ambiguity aversion from the (observed) fraction of the investment into the foreign stock. Lewis (1999) reported that the fraction was just 8% in the US data, which implies that the representative American's measure of ambiguity aversion is equal to 5.25.

The possibility of inferring the measure of ambiguity aversion from observable choice behavior is of considerable help in any quantitative analysis of ambiguity aversion. Recall, as shown by Corollaries 3 and 4, that the measure of ambiguity aversion is related to the ambiguity premium and the matching probabilities. Thus, we can check, by comparing the inferred values with those derived from laboratory experiments, whether the inferred value is derived solely from ambiguity aversion or contaminated by institutional factors, such as incomplete asset markets and transaction costs in the context of the home bias puzzle.

8 Extension to the non-differentiable case

8.1 Background and motivation

To define the measure of ambiguity aversion in (1), we assumed that the aggregator I is twice continuously differentiable. This assumption excludes, most notably, maximin expected utility of Gilboa and Schmeidler (1989) and Choquet expected utility of Schmeidler (1989). As they have significant behavioral implications, such as the inertia in portfolio choice, we extend the definition of the measure of ambiguity aversion to such non-differentiable aggregators and generalize the results on matching probabilities and ambiguity premiums in Sections 5 and 6. This allows us to compare the ambiguity aversion exhibited by non-differentiable aggregators on the same footing with the ambiguity aversion exhibited by differentiable ones.

The extension to the non-differentiable case involves two additional considerations that were unnecessary for the twice continuously differentiable case. The first one is to give a notion of twice differentiability that can accommodate maximin expected utility and Choquet expected utility. Since these utility functions are not even differentiable, we need to give one that is particularly suited to our purpose. Our definition is based on Peano right-derivatives.

The second consideration we need specifically for the non-differentiable case is the choice of the benchmark probability in the sense of Ghirardato and Marinacci (2002). As is well known, for concave but non-differentiable aggregators, there may be multiple benchmark probabilities. When defining the measure of ambiguity aversion, the choice of the benchmark probability p affects the ambiguity measure $H^x(z)$ through the variance $\mathrm{Var}^p[z]$ of the deviation z from the constant act xe. In our extended definition, we will allow p to depend on the choice of z (even after we fix the deterministic consumption level x), but require p to minimize its mean $E^p[z]$ among all supergradients and, in addition, to maximize its variance $\mathrm{Var}^p[z]$ among all such mean-minimizing supergradients, for a reason that we will elaborate on in the next subsection.

8.2 Definitions

We assume that the aggregator I is continuous, non-decreasing, normalized, concave, but not necessarily twice continuously differentiable. Then, I is locally Lipschitz continuous. It follows from Theorem 3.6(ii) and Theorem 4.8 of Chapter 3 of Delfour (2019) that I is Hadamard right-differentiable at every point in all directions in the sense of Definition 3.1(ii) in Section 3 of Chapter 3 of Delfour (2019). Denote the Gateaux and Hadamard right-derivative of I at g in the direction z by $\mathrm{d}_{\mathrm{G}}I(g;z)$ and $\mathrm{d}_{\mathrm{H}}I(g;z)$. They are positively homogeneous, but not necessarily linear, in z. They coincide whenever the latter exists. They are linear in z if I is differentiable in the standard sense; and we can then write $\mathrm{d}_{\mathrm{G}}I(g;z) = \mathrm{d}_{\mathrm{H}}I(g;z) = \nabla I(g)z$ using the gradient vector $\nabla I(g)$ of I at g.

Below are two versions of twice right-differentiability that we use in this paper. It is given in terms of a general function F defined on an open subset A of an Euclidean space $\boldsymbol{R}^N$ of arbitrary dimension, as it will be used not only for the aggregator I also for other functions such as the ambiguity premium.

Definition 2 Let $a \in A$ and $b \in \boldsymbol{R}^N$.

1. We say that F is *twice Peano-Gateaux right-differentiable at a in the direction b* if F is Gateaux right-differentiable at a in the direction b and there is an $L \in \boldsymbol{R}$ such that

$$\frac{F(a+\varepsilon b) - (F(a) + \varepsilon \mathrm{d}_{\mathrm{G}}F(a;b))}{\varepsilon^2} \to L$$

   as $\varepsilon \to 0+$. We denote $2L$ by $\mathrm{d}^2_{\mathrm{G}}F(a;b)$ and call it the *second Peano-Gateaux right-derivative of F at a in the direction b.*

2. We say that F is *twice Peano-Hadamard right-differentiable at a in the direction b* if F is Hadamard right-differentiable at a in all directions near b and there is an $L \in \boldsymbol{R}$ such that

$$\frac{F(a+\varepsilon z) - (F(a) + \varepsilon \mathrm{d}_{\mathrm{H}}F(a;z))}{\varepsilon^2} \to L$$

as $\varepsilon \to 0+$ and $z \to b$. We denote $2L$ by $\mathrm{d}^2_{\mathrm{G}} F(a;b)$ and call it the *second Peano-Hadamard right-derivative of F at a in the direction b.*

Both of the two definitions require the function F to satisfy the second-order Taylor approximation, but they are different in that the twice Peano-Gateaux right-differentiability is met whenever the direction b of deviation from a is fixed throughout the limiting operation while the twice Peano-Hadamard right-differentiability requires the existence of the limit even when the direction of deviation itself varies around b. The twice Peano-Hadamard right-differentiability implies twice Peano-Gateaux right-differentiability, and, then, the second Peano-Hadamard and Peano-Gateaux right-derivatives coincide. They also coincide whenever $N = 1$, and we then refer to them simply as the second Peano right-derivative. They are positively homogeneous of degree two, but not necessarily quadratic, in b. They are quadratic in b if F is twice differentiable in the standard sense; and we can then write $\mathrm{d}^2_{\mathrm{G}} F(a;b) = \mathrm{d}^2_{\mathrm{H}} F(a;b) = b^\top \nabla^2 F(a) b$ using the Hessian $\nabla^2 F(a)$.

The above definition of twice right-differentiability weaker than other known definitions of twice right-differentiability, such as Definition 3.12 in Section 3 of Chapter 3 of Delfour (2019), Theorem 2.3 of Rockafellar (2000), and Definition 13.1 of Rockafellar and Wets (1998), mainly in that the differentiability of the (first) right-derivative on a neighborhood of the point a is not a prerequisite. For example, Definition 3.12 in Section 3 of Chapter 3 of Delfour (2019) goes as follows: If F is Gateaux right-differentiable at every point near a in all directions, and there is an $L \in \boldsymbol{R}$ such that

$$\frac{\mathrm{d}_{\mathrm{G}} F(a + \varepsilon c; b) - \mathrm{d}_{\mathrm{G}} F(a; b)}{\varepsilon} \to L$$

as $\varepsilon \to 0+$, then F is said to be twice right-differentiable at a in the direction (b, c). Unfortunately, this condition is not satisfied by maximin expected utilities of Gilboa and Schmeidler (1995).

To extend the measure of ambiguity aversion to the non-differentiable case, we give the following notation. For each $y \in v(T)$ and $z \in \boldsymbol{R}^S$, denote by Δ^y_I the set of all supergradients of I at ye, by $\Delta^y_I(z)$ the set of all $p \in \Delta^y_I$ such that $p \cdot z \leq q \cdot z$ (that is, $E^p[z] \leq E^q[z]$) for every $q \in \Delta^y_I$, and by $\Lambda^y_I(z)$ as the set of all $p \in \Delta^y_I(z)$ such that $p \cdot z^2 \geq q \cdot z^2$ (that is, $E^p[z^2] \geq E^q[z^2]$, which is equivalent to $\mathrm{Var}^p[z] \geq \mathrm{Var}^q[z]$) for every $q \in \Delta^y_I(z)$, where $z^2 = ((z(s))^2)_s \in \boldsymbol{R}^S$. These sets are nonempty and compact.

Definition 3 Let $x \in T$ and $z \in \boldsymbol{R}^S$. Suppose that I is twice Peano-Gateaux right-differentiable at $v(x)e$ in the direction z and let $p \in \Lambda_I^{v(x)}(z)$. Then, we define

$$H^x(z) = \frac{-\dfrac{v'(x)\mathrm{d}_{\mathrm{G}}^2 I(v(x)e; z)}{\mathrm{Var}^p[z]}}{-\dfrac{v''(x)}{v'(x)}}. \tag{18}$$

If I is twice continuously differentiable, then $\Lambda_I^{v(x)} = \{\nabla I(v(x)e)\}$ and $\mathrm{d}_{\mathrm{H}}^2 I(v(x)e; z) = z^\top \nabla^2 I(v(x)e)z$. Thus, the above definition of $H^x(z)$ is, indeed, an extension of (1).

The use of a $p \in \Lambda_I^{v(x)}(z)$ in the definition of $H^x(z)$ can be justified as follows. For each $p \in \Delta_I^{v(x)}$, let $\bar{I}_p$ be the expectation operator under p and $\bar{V}_p$ be defined by $\left(v, \bar{I}_p\right)$. Then, $\bar{V}_p$ represents a benchmark preference of (the preference represented by) V, and $\bar{V}_p \geq V$. Thus, the best approximation of V among the $\bar{V}_p$ over $p \in \Delta_I^{v(x)}$ is the one that minimizes the values of $\bar{V}_p$. By the Taylor's theorem,

$$\begin{aligned}\bar{V}_p(xe + \varepsilon z) &\approx \bar{V}_p(xe) + \sum_s p(s)\left(v'(x)\varepsilon z(s) + \frac{1}{2}v''(x)(\varepsilon z(s))^2\right)\\ &= v(x) + v'(x)E^p[z]\varepsilon + \frac{1}{2}v''(x)E^p[z^2]\varepsilon^2.\end{aligned}$$

To minimize the first-order impact on $\bar{V}_p(xe+\varepsilon z)$, we need to minimize the second-term on the far right-hand side, that is, choose a $p \in \Delta_I^{v(x)}(z)$. Once this is done, $\bar{V}_p(xe+\varepsilon z)$ is minimized when the last term is minimized. This is true when $E^p[z^2]$ is maximized, that is, when $p \in \Lambda_I^{v(x)}$. Note here that the risk aversion, $v'' < 0$, is crucial: When choosing a benchmark preference and probability, we should not ignore the second-order impact on V arising from risk aversion. If we did, then $H^x(z)$ would overestimate the ambiguity aversion represented by I.

8.3 Matching probabilities

In this subsection, we generalize the characterization of the measure of ambiguity aversion in terms of matching probabilities (Theorem 3 and Corollary 3). For a sufficiently small $\varepsilon > 0$, define the matching probability $\rho^x(\varepsilon, A)$ just as in (5). For $\varepsilon = 0$, we now let $\rho^x(0, A) = p(A)$, where $p \in \Delta_I^{v(x)}(e_A)$ and $p(A) = p \cdot e_A$. The value of $p(A)$ is independent of the choice of $p \in \Delta_I^{v(x)}(e_A)$.

Theorem 5 *Let $A \subset S$ and $x \in T$. If I is twice Peano-Gateaux right-differentiable at $v(x)e$ in the direction e_A, then $\rho^x(\cdot, A)$ is right-differentiable at 0 and*

$$\mathrm{d_G}\rho^x(0, A) = \frac{1}{2} v'(x)\mathrm{d}^2_{\mathrm{G}} I(v(x)e; e_A). \tag{19}$$

In many cases of interest, a symmetry consideration suggests that there is a "natural" benchmark probability in $\Delta_I^{v(x)}$ that is outside $\Delta_I^{v(x)}(e_A)$. For example, in the case of the two-color Ellsberg urn, the natural probability assigns probability 1/2 to the event of drawing a ball of the winning color, while the probability in $\Delta_I^{v(x)}(e_A)$ assigns the lowest probability, among all benchmark probabilities, to the winning event, which is less than 1/2. Let p^* be the natural benchmark probability and p be in $\Delta_I^{v(x)}(e_A)$, then the difference $p^*(A) - p(A)$ is often taken as a measure of ambiguity aversion. Our measure of ambiguity aversion differs from this one, as it is related to further decrease in matching probabilities $\rho^x(\varepsilon, A)$ as the value of the prize, ε, increases. When coupled with the often-used measure, however, it can improve the approximation of the difference between the natural benchmark probability and the matching probability via

$$p^*(A) - \rho^x(\varepsilon, A) \approx (p^*(A) - p(A)) - \frac{\varepsilon}{2} v'(x)\mathrm{d}^2_{\mathrm{G}} I(v(x)e; e_A).$$

for a small $\varepsilon > 0$, because $\rho^x(0, A) = p(A)$ for any $p \in \Delta_I^{v(x)}(e_A)$.

Define $\underline{I}$, $\underline{v}$, $\underline{V}$, and $\underline{\rho}$ as in Section 5, that is, $\underline{I} = v^{-1} \circ V$, $\underline{v}$ is the identify function, $\underline{V}$ is defined by $(\underline{v}, \underline{I})$, and $\underline{\rho}$ is the matching probability for the utility function $\underline{V}$. Corollary 3 can be extended to the non-differentiable case as follows.

Corollary 5 *Under the assumptions of Theorem 5, $\underline{I}$ also satisfies the assumptions of Theorem 5 and*

$$H^x(e_A) = \frac{\mathrm{d_G}\rho^x(0, A)}{\mathrm{d_G}\underline{\rho}^x(0, A) - \mathrm{d_G}\rho^x(0, A)}. \tag{20}$$

The asymptotic result (9) can also be extended to the non-differentiable case.

Corollary 6 *Under the assumptions of Theorem 5, for every $p \in \Delta_I^{v(x)}(z)$,*

$$\frac{p(A) - \rho^x(\varepsilon, A)}{\big(p(A) - \underline{\rho}^x(\varepsilon, A)\big) - (p(A) - \rho^x(\varepsilon, A))} \to H^x(e_A)$$

as $\varepsilon \to 0$.

8.4 Biseparable preferences

In this subsection, we prove that for the utility functions that represent biseparable preferences, the extended measure of ambiguity aversion is always equal to zero and, then, compare this fact, graphically, with the case of twice continuously differentiable utility functions.

The following definition of biseparability is due to Ghirardato and Marinacci (2001, Section 2.2) and Ghirardato and Marinacci (2002, Section 1).

Definition 4 We say that V is *biseparable* if, for every $A \subset S$, there is a $\rho(A) \in [0,1]$ such that $\rho^x(\varepsilon, A) = \rho(A)$ for every $x \in T$ and every $\varepsilon > 0$.

According to this definition, a utility function is biseparable if the matching probability of any event A does not depend on the consumption levels on A or outside A, as long as the former is higher than the latter. Examples of biseparable preferences include Choquet expected utility functions of Schmeidler (1989) and α-maximin expected utility functions of Ghirardato, Maccheroni, and Marinacci (2004).

Proposition 3 *Suppose that V is defined by (v, I) and biseparable. Then, for all $x \in T$ and $A \subset S$, I is twice Peano-Gateaux right-differentiable at $v(x)e$ in the direction e_A and $\mathrm{d}^2_{\mathrm{G}} I(v(x)e; e_A) = 0$. If $0 < p(A) < 1$ for any $p \in \Lambda_I^{v(x)}(e_A)$, then $H^x(e_A) = 0$.*

It can be easily shown that for Choquet expected utility functions and α-maximin expected utility functions, the aggregator I is, in fact, twice Peano-Hadamard right-differentiable, and $\mathrm{d}^2_{\mathrm{H}} I(v(x)e; e_A) = 0$.

We now turn to the illustration of matching probabilities for biseparable, twice continuously differentiable but ambiguity-averse, and subjective expected utility functions. Let V_1 and V_2 be two utility functions defined by (v_1, I_1) and (v_2, I_2). Suppose that the preference represented by V_2 is a benchmark preference of the preference represented by V_1, with p the benchmark probability. Then, they share the same risk attitudes, by which we can assume that $v_1 = v_2$ and write v in place of v_1 and v_2. Moreover, by Theorem 12 of Ghirardato and Marinacci (2002), $p \in \Delta_{I_1}^{v(x)}$ for every $x \in T$. By the definition of $\rho^x(0, A)$, $\rho^x(0, A) \leq p(A)$ and

$\rho^x(0, S \setminus A) \leq p(S \setminus A)$. Thus, $\rho^x(0, A) + \rho^x(0, S \setminus A) \leq 1$. The argument so far is applicable regardless of whether V_1 is biseparable or smooth.

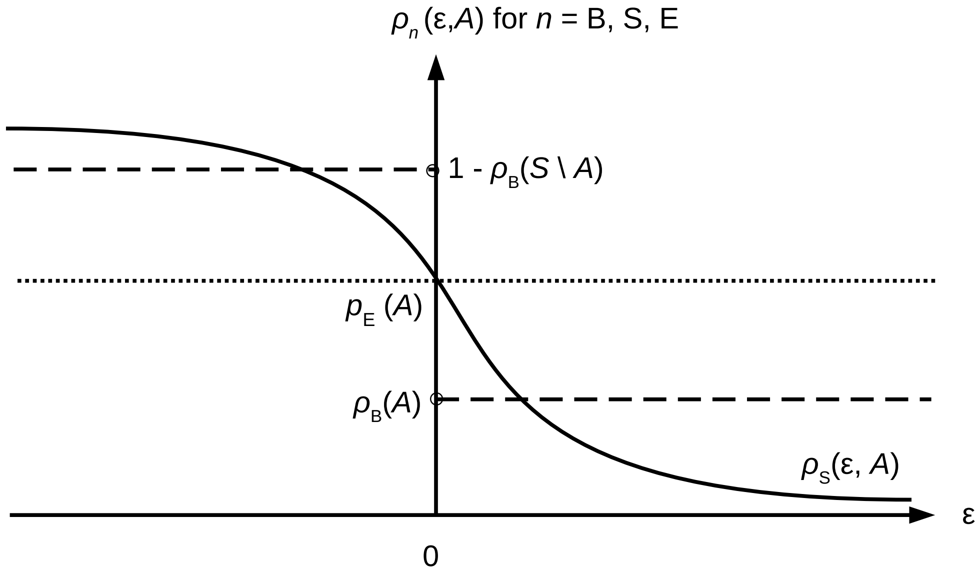
Suppose now that V_1 is biseparable and A is an ambiguous event for the preference in the sense of Ghirardato and Marinacci (2002, Section 4). Then, $\rho^x(\varepsilon, A) = \rho(A)$ for every $\varepsilon > 0$. Moreover, by the definition of $\rho^x(0, A)$, this equality holds even when $\varepsilon = 0$. By Proposition 22 of Ghirardato and Marinacci (2002), $\rho(A) + \rho(S \setminus A) < 1$. Hence, $\rho(A) \leq p(A) \leq 1 - \rho(S \setminus A)$, and at least one of these two weak inequalities holds as a strict inequality. As for the case where $\varepsilon < 0$, by (7), $\rho^x(\varepsilon, A) = 1 - \rho(S \setminus A)$. For a biseparable preference, therefore, the graph of the matching probability, $\varepsilon \mapsto \rho^x(\varepsilon, A)$, jumps down at zero, crossing the benchmark probability $p(A)$, but is constantly equal to $\rho(A)$ on the right of zero, and to $1 - \rho(S \setminus A)$ on the left of zero.

Suppose, instead, that V_1 is twice continuously differentiable and A is an ambiguous event for V_1 in the sense that $H^x(e_A) > 0$. Then $\rho^x(0, A) = p(A)$, where $p = \nabla I_1(v(x)e)$, the unique benchmark probability for V_1. Thus, the graph of the matching probability $\varepsilon \mapsto \rho^x(\varepsilon, A)$ goes through the benchmark probability $p(A)$ at 0. Moreover, by Theorem 3, it is differentiable at $\varepsilon = 0$ and, by (8), the derivative is negative. For a twice continuously differentiable utility function, therefore, the graph of the matching probability, $\varepsilon \mapsto \rho^x(\varepsilon, A)$, is differentiable and downward-sloping around 0, where it is equal to the benchmark probability.

These observations are gathered in Figure 3, in which we consider a biseparable utility function V_{B}, a twice continuously differentiable utility function V_{S}, and an expected utility function V_{E}. Assume that V_{E} represents a benchmark preference for V_{B} and also for V_{S}. Let p be the benchmark probability corresponding to V_{E}. Let $A \subset S$. Assume that A is an ambiguous event for both V_{B} and V_{S}, that is, $\rho_{\mathrm{B}}(A) + \rho_{\mathrm{B}}(S \setminus A) < 1$ and $H^x_{\mathrm{S}}(e_A) > 0$.

Then, the matching probability $\rho_{\mathrm{E}}(\varepsilon, A)$ for the expected utility function V_{E} is equal to $p(A)$ for every ε. Its graph is horizontal and intercepts the vertical axis at $p(A)$. The matching probability $\rho_{\mathrm{S}}(\varepsilon, A)$ for the twice continuously differentiable utility function V_{S} is equal to $p(A)$ at $\varepsilon = 0$ and decreasing around $\varepsilon = 0$. Its graph intercepts the vertical axis at $p(A)$ and slopes downwards around there. The matching probability $\rho_{\mathrm{B}}(\varepsilon, A)$ for the biseparable utility function V_{B} is equal to $\rho_{\mathrm{B}}(A)$ at every $\varepsilon \geq 0$ and to $1 - \rho_{\mathrm{B}}(S \setminus A)$ at every $\varepsilon < 0$. The graph of the matching probability has a downward jump on the vertical axis, passing $p(A)$, although it is horizontal in either side of the vertical axis.

Figure 3: Matching probabilities for a biseparable utility function, a twice continously differentiable utility function, and an expected utility function



The dashed lines, with a gap on the vertical axis, constitute the graph of the matching function for a biseparable utility function. The solid downward-sloping smooth curve is the graph of the matching probability for a twice continuously differentiable utility function. The dotted horizontal line is the graph of the matching function of an expected utility function, which represents a benchmark preference for the two ambiguity-averse utility functions.

The graphs of matching probabilities can help us distinguish the nature of ambiguity aversion of twice continuously differentiable utility functions from that of biseparable preferences. For a biseparable preference, at every constant act, there is a range of probabilities of an ambiguous event that is the decision maker deems possible or relevant. When it comes to betting on an event (which induces a non-constant act), however, his assessment of the likelihood of the event goes, discontinuously, down to the minimum of these probabilities, with no further decrease in the assessment when the prize goes up. In contrast, for a twice continuously differentiable utility function, at every constant act, the decision maker believes in a single probability, which coincides with the benchmark probability. As the prize goes up, his assessment of the likelihood of the event keeps, smoothly, going down. Therefore, for a sufficiently small prize, a biseparable preference is, in terms of matching probabilities, more ambiguity-averse than a twice continuously differentiable utility function; but an increase in prizes may well cause the former less ambiguity-averse than the latter.

8.5 Minimum over twice continuously differentiable functions

We would now like to study ambiguity premiums for the non-differentiability case, but, as we will elaborate on in the next subsection, the analysis is more subtle than that of matching probabilities because the second-order impact of risk aversion, $v'' < 0$, is harder to evaluate than in the analysis of matching probabilities. To ease the analysis, we concentrate on the case where the non-differentiable aggregator I is the minimum of the twice continuously differentiable functions. This class of aggregators I is of particular interest, as it includes maximin expected utility functions of Gilboa and Schmeidler (1989) and more general one for which the matching probability jumps down for any, however small, positive prize and decreases further as the value of the prize increases.

Assumption 1 There is a set $\mathscr{J}$ of twice continuously differentiable (C^2), increasing, normalized, and concave real-valued functions defined on $v(T)^S$ that is compact with respect to the C^2 compact-open topology and satisfies $I(g) = \min_{J \in \mathscr{J}} J(g)$ for every $g \in v(T)^S$.

The set of C^2 real-valued functions defined on $v(T)^S$ is endowed with the weak topology in Hirsch (1976, Section 1 of Chapter 2), which is also referred to as the C^2 compact-open topology in Mas-Colell (1985, Section K.1 of Chapter 1). With respect to this topology, a sequence of functions J_n $(n = 1, 2, \dots)$ converges to a function J if and only if, on every compact subset B of $v(T)^S$, the sequences of the J_n and all of the derivatives up to the second order of the J_n converge, uniformly over B, to J and all of the derivatives up to the second order of J. The compactness assumption guarantees that for every $g \in v(T)^S$, $\min_{J \in \mathscr{J}} J(g)$ indeed exists. The assumption implies that I is continuous, increasing, normalized, and concave.

Assumption 1 requires every element of $\mathscr{J}$ to be normalized. This is intended to make the class of aggregators under consideration a generalization of the class of maximin expected utility functions. This should be contrasted with the dual expression of utility functions by Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011, Theorem 3). They showed that every uncertain-averse aggregator can be written as $I(g) = \min_{p\in\Delta} G(E^p[g], p)$ for some grounded function $G : \boldsymbol{R} \times \Delta \to \boldsymbol{R}$ (that is, $\min_{p\in\Delta} G(y, p) = y$ for every y). But, the (partial) functions $g \mapsto G(E^p[g], p)$ in this expression need not be normalized.

Equally important for the subsequent analysis is that the topology with respect to which $\mathscr{J}$ is required to be compact is the C^2 compact-open topology. If $\mathscr{J}$ were required to be compact only with respect to the C^1 compact-open topology, every increasing and concave I could satisfy this assumption with a set $\mathscr{J}$ that consists of functions with zero Hessians at a given point,[15] which indicates that there is no way to relate the second Peano derivatives of I with the Hessians of the elements of $\mathscr{J}$.

We now give a sufficient condition for I to be twice Hadamard right-differentiable under Assumption 1. Note that $\Delta_I^y = \operatorname{conv}\{\nabla J(ye) \mid J \in \mathscr{J}\}$ for each $y \in v(T)$. For each $y \in v(T)$ and each $z \in \boldsymbol{R}^S$, define $\mathscr{J}(y, z)$ as the set of all $J \in \mathscr{J}$ that satisfy $\nabla J(ye)z \leq \nabla K(ye)z$ for all $K \in \mathscr{J}$. Then, $\mathscr{J}(y, z)$ is nonempty and compact, because $\mathscr{J}$ is compact, and $\Delta_I^y(z) = \operatorname{conv}\{\nabla J(ye) \mid J \in \mathscr{J}(y, z)\}$.

Proposition 4 *Let $y \in v(T)$ and $z \in \boldsymbol{R}^S$. Under Assumption 1, I is Hadamard right-differentiable at ye in the direction z, and $\mathrm{d_H} I(ye; z) = \nabla J(ye)z$ for any $J \in \mathscr{J}(y, z)$. If, in addition, $z^\top \nabla^2 J(ye)z = z^\top \nabla^2 K(ye)z$ for all $J \in \mathscr{J}(y, z)$ and $K \in \mathscr{J}(y, z)$, then I is twice Peano-Hadamard right-differentiable at ye in the direction z, and $\mathrm{d}^2_{\mathrm{H}} I(ye; z) = z^\top \nabla^2 J(ye)z$ for any $J \in \mathscr{J}(y, z)$.*

Since the sufficient condition in this proposition for twice Peaon-Hadamard right-differentiability will be used in the sequal, we put it up as an assumption.

Assumption 2 Let $y \in v(T)$ and $z \in \boldsymbol{R}^S$. The aggregator I satisfies Assumption 1 and $z^\top \nabla^2 J(ye)z = z^\top \nabla^2 K(ye)z$ for all $J \in \mathscr{J}(y, z)$ and $K \in \mathscr{J}(y, z)$

The following two examples satisfy Assumption 2 and, hence, twice Peano-Hadamard right-differentiability.

Example 1

1. Let Λ a strictly convex and compact subset of Δ, where the strict convexity is with respect to the relative topology on Δ. Let $J : \boldsymbol{R}^S \times \Lambda \to \boldsymbol{R}$ be twice continuously differentiable. Assume that for every $p \in \Delta$, the partial function $J(\,\cdot\,, p) : \boldsymbol{R}^S \to \boldsymbol{R}$ is increasing, normalized, and concave, and that $\nabla_g J(ye, p) = p^\top$ for every y, where ∇_g denote the partial derivative

[15]Think of $\mathscr{J}$ as consisting of convolutions of piecewise-linear utility functions that approximate I as in Afriat's (1967) construction of utility functions.

(Jacobian) with respect to the first coordinate. Define $\mathscr{J} = \{J(\,\cdot\,, p) \mid p \in \Lambda\}$.

2. Denote by Π the set of all permutations on S. For each $p \in \Delta$ and $\pi \in \Pi$, denote $(p(\pi(s)))_s \in \Delta$ by $p \circ \pi$. Let $p^* \in \Delta$ and $\Lambda = \{p^* \circ \pi \mid \pi \in \Pi\}$. Let $\gamma > 0$. For each $p \in \Lambda$, let $J_p : \boldsymbol{R}^S \to \boldsymbol{R}$ be a twice continuously differentiable, increasing, normalized, and concave function that satisfy $J_p(g) = E^p[g] - (\gamma/2)\mathrm{Var}^p[g]$ for every g in some open set that includes the diagonal $\{ye \mid y \in \boldsymbol{R}\}$. Let $\mathscr{J} = \{J_p \mid p \in \Lambda\}$.

The first example differs from maximin expected utility functions as it allows the Hessians $\nabla_g^2 J(ye, p)$ to be non-zero and, yet, requires the set Λ of relevant probabilities to be strictly convex. The strict convexity is obtained when Λ is the set of all probabilities whose distance from some reference probability is at most some small threshold, and the distance is a strictly convex function, as in the case of the Euclidean norm and the relative entropy. The second example is, locally, a mean-variance utility function under ambiguity, in the sense that the decision maker has a mean-variance utility function but is unsure of which probability to use to evaluate mean and variance. The set of relevant probabilities, Λ, is symmetric, in the sense that any probability distribution that the decision maker deems relevant is still relevant after swapping the probabilities of any two states.

8.6 Ambiguity premium

In this subsection, we extend the characterization result (Theorem 4) of our ambiguity measure in terms of the ambiguity premium under Assumption 2. Let $x \in T$, $z \in \boldsymbol{R}^S$, and $p \in \Delta_I^{v(x)}$. For each $\varepsilon \geq 0$, denote by $f_p^x(\varepsilon, z) \in \Pi(T)$ the lottery that coincides with the distribution of the monetary act $x + \varepsilon z$ on T under the probability p. Regarding $f_p^x(\varepsilon, z)$ as a lottery act, define the ambiguity premium $\kappa_p^x(\varepsilon, z)$ under p by

$$V\left(f^{x-\kappa_p^x(\varepsilon,z)}(\varepsilon, z)\right) = V(xe + \varepsilon z).$$

Let $\bar{I}_p$ be the expectation operator under p and define $\bar{V}_p$ by $(v, \bar{I}_p)$. Then, $\kappa_p^x(\varepsilon; z)$ can equivalently defined by

$$\bar{V}_p\left((x - \kappa_p^x(\varepsilon; z))e + \varepsilon z\right) = V(xe + \varepsilon z).$$

This definition of the ambiguity premium is different from the original definition (10) in that we specify the benchmark probability p under which the lottery act $f_p^x(\varepsilon, z)$ is defined. We need to do so because there are multiple benchmark probabilities when I is not (twice continuously) differentiable but concave.

Theorem 6 *Let $x \in T$ and $z \in \boldsymbol{R}^S$.*

1. *Let $p \in \Delta_I^{v(x)}$. Under Assumption 1, the function $\varepsilon \mapsto \kappa_p^x(\varepsilon; z)$ is right-differentiable at* 0 *and* $\mathrm{d_H}\kappa_p^x(0; z) = p \cdot z - q \cdot z$ *for any* $q \in \Delta_I^{v(x)}(z)$.
2. *Let $p \in \Lambda_I^{v(x)}(z)$. Under Assumption 2, the function $\varepsilon \mapsto \kappa_p^x(\varepsilon; z)$ is twice Peano right-differentiable at* 0, $\mathrm{d_H}\kappa_p^x(0; z) = 0$, *and* $\mathrm{d_H^2}\kappa_p^x(0; z) = -v'(x)\mathrm{d_H^2}I(v(x)e; z)$.

In a terminology analogous to that of Segal and Spivak (1990), Part 1 pins down the source of the first-order ambiguity aversion, and Part 2 relates, in the absence of the first-order ambiguity aversion, the second-order ambiguity aversion to our extended measure of ambiguity aversion. More specifically, part 1 shows that the first-order ambiguity aversion emerges as the positive first right-derivative of the ambiguity premium only if the benchmark probability p used in its evaluation lies outside $\Delta_I^{v(x)}(z)$. This would be the case if there is a "natural" benchmark probability (in $\Delta_I^{v(x)}$ but) outside $\Delta_I^{v(x)}(z)$, as we explained after Theorem 5. If, instead, we use a benchmark probability in $\Lambda_I^{v(x)}(z)$ to evaluate the ambiguity premium, then the first-order ambiguity aversion disappears and the second-order ambiguity aversion can be related to the ambiguity aversion via

$$\kappa_p^x(\varepsilon, z) \approx -\frac{1}{2}v'(x)\mathrm{d_H^2}I(v(x)e; z)$$

for a small $\varepsilon > 0$.

It is worth noting here that the condition needed for I in Theorem 5 was much weaker than that needed for I in Theorem 6. In Theorem 5, we only required I to be twice Gateaux right-differentiable, while, in Theorem 6, we impose Assumption 2. The difference is due to the ease with which to assess the impact of the deviation in the direction the indicator function e_A of any event A in Theorem 5. Specifically, the vector of changes in realized utility levels caused by εe_A, $v \circ (xe + \varepsilon e_A) - v \circ (xe)$, is equal to $(v(x+\varepsilon) - v(x))e_A$, which is a scalar multiple of e_A regardless of the value of ε. Thus, the directions of these changes are invariant, which makes the difference between twice Hadamard and Gateaux right-differentiability irrelevant as far as the change in utility, $V(xe + \varepsilon e_A) - V(xe)$, is concerned. This is not the case for Theorem 6, because the direction of the vector $v \circ (xe + \varepsilon z) - v \circ (xe)$ in $\boldsymbol{R}^S$ varies as $\varepsilon \to 0+$.

Denote by $\bar{\kappa}_p^x(\varepsilon; z)$ the risk premium of the benchmark utility function $\bar{V}_p$. We then obtain the following characterization of $H^x(z)$ in the non-differentiable case from Theorem 6 and (12).

Corollary 7 *Let $x \in T$, $z \in \boldsymbol{R}^S$, and $p \in \Lambda_I^{v(x)}(z)$. Suppose that* $\mathrm{Var}^p[z] > 0$. *Under Assumption 2,*

$$H^x(z) = \frac{\mathrm{d}_{\mathrm{H}}^2 \kappa_p^x(0; z)}{\dfrac{\partial^2 \bar{\kappa}^x}{\partial \varepsilon^2}(0, z)}.$$

The asymptotic result (13) can be extended to the non-differentiable case, thanks to the definition of the second Peano right-derivative.

Corollary 8 *Let $x \in T$, $z \in \boldsymbol{R}^S$, and $p \in \Lambda_I^{v(x)}(z)$. Suppose that* $\mathrm{Var}^p[z] > 0$. *Under Assumption 2,*

$$\frac{\kappa_p^x(\varepsilon, z)}{\bar{\kappa}_p^x(\varepsilon, z)} \to H^x(z)$$

as $\varepsilon \to 0+$.

9 Conclusion

In this paper, we introduced an Arrow-Pratt-type measure of ambiguity aversion for a class of twice right-differentiable utility functions. The notion of twice right-differentiability we employed is so weak that the class include not only smooth ambiguity models but also maximin and Choquet expected utility functions and other one that are neither differentiable in the standard sense nor biseparable. While we assumed in the main body of the paper that the risk attitude can be represented by an expected utility function, the definitions and results are extended, in an appendix, to the case where it cannot.

This type of measure is particularly useful for a quantitative analysis where the magnitude of ambiguity aversion is estimated, inferred, or calibrated from laboratory findings or market data. Our measure makes it possible to compare ambiguity attitudes of two decision makers even when they apparently have different risk attitudes. Moreover, when the estimated/inferred/calibrated values of our measure are significantly different between two settings, such as laboratory experiments versus asset markets, they indicate that there is room for further improvement in the settings, such as the contents of questionnaires in experiments and the assumption of complete markets on asset markets.

A Lemmas and proof for the differentiable case

Proof of Theorem 1 For each $w \in \boldsymbol{R}^S$, denote by $[w]$ the $S \times S$ matrix of which the s-th diagonal entry is equal to the s-th coordinate of w and the off-diagonal entries are all equal to zero. By the chain rule differentiation,

$$\nabla V(f) = \nabla I(v \circ f)[v' \circ f], \tag{21}$$

for every $f \in T^S$, where the gradients are row vectors. Thus,

$$\nabla^2 V(f) = [v' \circ f]\nabla^2 I(v \circ f)[v' \circ f] + [\nabla I(v \circ f)][v'' \circ f] \tag{22}$$

for every $f \in T^S$. Thus, if $f = xe$ for some $x \in T$, then $v \circ f = v(x)e$, $v' \circ f = v'(x)e$ and $\nabla V(f) = p$. Since $[e]$ coincides with the $I \times I$ identity matrix,

$$\begin{aligned} V(xe) &= v(x), \\ \nabla V(xe) &= v'(x)p^\top, \qquad (23) \\ \nabla^2 V(xe) &= (v'(x))^2 \nabla^2 I(v(x)e) + v''(x)[p]. \qquad (24) \end{aligned}$$

Since $v(V^*(f)) = V(f)$, $v'(V^*(f))\nabla V^*(f) = \nabla V(f)$ and

$$v''(V^*(f))\nabla V^*(f)^\top \nabla V^*(f) + v'(V^*(f))\nabla^2 V^*(f) = \nabla V^2(f)$$

for every $f \in T^S$. Thus, if $f = xe$ for some $x \in T$, then

$$\begin{aligned} V^*(xe) &= x, \qquad (25) \\ \nabla V^*(xe) &= p^\top, \qquad (26) \\ v''(x)pp^\top + v'(x)\nabla^2 V^*(xe) &= \nabla^2 V(xe). \qquad (27) \end{aligned}$$

By (24) and (27),

$$\nabla^2 V^*(xe) = v'(x)\nabla^2 I(v(x)e) + \frac{v''(x)}{v'(x)}([p] - pp^\top) \tag{28}$$

Thus, (2) follows from $z^\top([p] - pp^\top)z = \mathrm{Var}^p[z]$. Since $\nabla^2 \bar{I}(v(x)e) = 0$, (3) can be similarly proved. ///

Proof of Theorem 2 For each $\varepsilon \in \boldsymbol{R}$ with $xe + \varepsilon z \in T^S$, define $\lambda(\varepsilon)$ by $V^*(xe + \lambda(\varepsilon)S^{-1/2}e + \varepsilon\|z\|^{-1}z) - V^*(xe) = 0$, where $\|z\| = \left(\sum_s (z(s))^2\right)^{1/2}$, which is also equal to $\left(\mathrm{Var}^p[z]\right)^{1/2}$ because $E^p[z] = 0$. Then, $\lambda(0) = 0$ and the curve $\left\{f \in T^S \cap L(z) \mid V(f) = V(xe)\right\}$ is parameterized by $\varepsilon \mapsto xe + \lambda(\varepsilon)S^{-1/2}e + \varepsilon\|z\|^{-1}z$. By the implicit function theorem, $\lambda(\varepsilon)$ is a twice continuously differentiable function of ε, $\lambda'(0) = 0$, and

$$\lambda''(0) = -S^{1/2}\left(\frac{z^\top\left(v'(x)\nabla^2 I(v(x)e)\right)z}{\mathrm{Var}^p[z]} + \frac{v''(x)}{v'(x)}\right).$$

Since the curve $\left\{f \in T^S \cap L(z) \mid V(f) = V(xe)\right\}$ is parameterized by $\varepsilon \mapsto xe + \lambda(\varepsilon)S^{-1/2}e + \varepsilon\|z\|^{-1}z$ on the plane $L(z)$, $c^x(z)$ is equal to its curvature when it is regarded as a subset of the plane $L(z)$ with an orthonormal basis $(S^{-1/2}e, \|z\|^{-1}z)$. Thus

$$c^x(z) = \frac{|\lambda''(0)|}{\left((\lambda'(0))^2 + 1\right)^{3/2}} = |\lambda''(0)| = S^{1/2}\left|\frac{z^\top\left(v'(x)\nabla^2 I(v(x)e)\right)z}{\mathrm{Var}^p[z]} + \frac{v''(x)}{v'(x)}\mathrm{Var}^p[z]\right|.$$

Since I is quasi-concave and $E^p[z] = 0$, $z^\top\left(v'(x)\nabla^2 I(v(x)e)\right)z \leq 0$. Since $v'' < 0 < v'$, $v''(x)/v'(x) < 0$. Thus,

$$c^x(z) = -S^{1/2}\left(\frac{z^\top\left(v'(x)\nabla^2 I(v(x)e)\right)z}{\mathrm{Var}^p[z]} + \frac{v''(x)}{v'(x)}\right).$$

We can similarly show that

$$\bar{c}^x(z) = -S^{1/2}\frac{v''(x)}{v'(x)}.$$

///

The proof of Theorem 3 is based on the following lemma. It can be proved based on Taylor's theorem, and we omit the proof.

Lemma 1 *Let D be an open interval in $\boldsymbol{R}$ that contains 0. Let $F : D \to \boldsymbol{R}$ be twice differentiable and satisfy $F(0) = 0$. Define $G : D \to \boldsymbol{R}$ by*

$$G(\delta) = \begin{cases} F'(0) & \text{if } \delta = 0, \\ \dfrac{F(\delta)}{\delta} & \text{otherwise.} \end{cases}$$

Then G is differentiable at 0 and $G'(0) = (1/2)F''(0)$.

Proof of Theorem 3 Let D be the set of all $\delta \in \boldsymbol{R}$ such that $v(x) + \delta \in v(T)$, then D is an open interval that contains 0. For each $\delta \in D$, define $F : D \to \boldsymbol{R}$ by $F(\delta) = I(v(x)e + \delta e_A) - v(x)$. Then $F(0) = 0$, $F'(0) = \nabla I(v(x)e)e_A$, and

$$F''(0) = e_A^\top \nabla^2 I(v(x)e)e_A. \tag{29}$$

Define $G : D \to \boldsymbol{R}$ as in Lemma 1, then

$$\begin{aligned} \rho^x(\varepsilon, A) &= \frac{I\left(v(x)e + (v(x+\varepsilon) - v(x))e_A\right) - v(x)}{v(x+\varepsilon) - v(x)} \\ &= \frac{F(v(x+\varepsilon) - v(x))}{v(x+\varepsilon) - v(x)} = G(v(x+\varepsilon) - v(x)). \end{aligned}$$

By Lemma 1, the right-hand side is a function of ε that is differentiable at $\varepsilon = 0$, and the derivative at $\varepsilon = 0$ is equal to $(1/2)F''(0)v'(x)$. By (29), this completes the proof ///

The proof of Corollary 3 requires the following lemma.

Lemma 2 *Let $A \subset S$ and $x \in T$, then*

$$e_A^\top \left(\underline{v}'(x)\nabla^2 \underline{I}(xe)\right) e_A = e_A^\top \left(v'(x)\nabla^2 I(v(x)e)\right) e_A + \frac{v''(x)}{v'(x)} \mathrm{Var}^p[e_A]. \tag{30}$$

Proof of Lemma 2 Since $v(\underline{I}(f)) = V(f)$ for every $f \in T^S$, $v'(\underline{I}(f))\nabla \underline{I}(f) = \nabla V(f)$ for every $f \in T^S$. Hence $v''(\underline{I}(f))\nabla \underline{I}(f)^\top \nabla \underline{I}(f) + v'(\underline{I}(f))\nabla^2 \underline{I}(f) = \nabla^2 V(f)$ for every $f \in T^S$. When $f = xe$, these equalities can be rewritten as $v'(x)\nabla \underline{I}(\underline{v}(x)e) = \nabla V(xe)$ and $v''(x)\nabla \underline{I}(xe)^\top \nabla \underline{I}(xe) + v'(x)\nabla^2 \underline{I}(xe) = \nabla^2 V(xe)$. By (23) and (24), $\nabla \underline{I}(xe) = \nabla I(v(x)e) = p$ and

$$v''(x)pp^\top + v'(x)\nabla^2 \underline{I}(xe) = (v'(x))^2 \nabla^2 I(v(x)e) + v''(x)[p].$$

Thus,

$$\nabla^2 \underline{I}(xe) = v'(x)\nabla^2 I(v(x)e) + \frac{v''(x)}{v'(x)}\left([p] - pp^\top\right).$$

Hence,

$$e_A^\top \nabla^2 \underline{I}(xe) e_A = e_A^\top \left(v'(x)\nabla^2 I(v(x)e)\right) e_A + \frac{v''(x)}{v'(x)}\mathrm{Var}^p[e_A].$$

Since $\underline{v}'(x) = 1$ by definition, this establishes (30). ///

Proof of Corollary 3 Apply Theorem 3 to (v, I) and $(\underline{v}, \underline{V})$, then, by (30),

$$\frac{\partial \underline{\rho}^x}{\partial \varepsilon}(0, A) = \frac{\partial \rho^x}{\partial \varepsilon}(0, A) + \frac{1}{2}\frac{v''(x)}{v'(x)}\mathrm{Var}^p[e_A]. \tag{31}$$

By rearranging this, we obtain the denominator of $H^x(e_A)$ and complete the proof. ///

Proof of Theorem 4 Define $V^* = v^{-1} \circ V$ and $\bar{V}^* = v^{-1} \circ \bar{V}$. By (25), (10) can be rewritten as $\bar{V}^*\left((x - \kappa^x(\varepsilon, z))\, e + \varepsilon z\right) = V^*(xe + \varepsilon z)$. By differentiating both sides with respect to ε, we obtain

$$\nabla \bar{V}^*\left((x - \kappa^x(\varepsilon, z))\, e + \varepsilon z\right)\left(-\frac{\partial \kappa^x}{\partial \varepsilon}(\varepsilon, z)e + z\right) = \nabla V^*(xe + \varepsilon z)z. \tag{32}$$

By letting $\varepsilon = 0$ and using (26), we obtain

$$\frac{\partial \kappa^x}{\partial \varepsilon}(0, z) = 0 \tag{33}$$

By differentiating both sides of (32) with respect to ε, we obtain

$$\begin{aligned} &\left(-\frac{\partial \kappa^x}{\partial \varepsilon}(\varepsilon, z)e + z\right)^\top \nabla^2 \bar{V}^*\left((x - \kappa^x(\varepsilon, z))\, e + ze\right)\left(-\frac{\partial \kappa^x}{\partial \varepsilon}(\varepsilon, z)e + z\right) \\ &+ \nabla \bar{V}^*\left((x - \kappa^x(\varepsilon, z))\, e + ze\right)\left(-\frac{\partial^2 \kappa^x}{\partial \varepsilon^2}(\varepsilon, z)e\right) \\ = &\, z^\top \nabla^2 V^*(xe + \varepsilon z)z. \end{aligned} \tag{34}$$

By letting $\varepsilon = 0$ and using (26) and (33), we obtain

$$z^\top \nabla^2 \bar{V}^* (xe)\, z - \frac{\partial^2 \kappa^x}{\partial \varepsilon^2}(0, z) = z^\top \nabla^2 V^*(xe) z.$$

Thus,

$$\frac{\partial^2 \kappa^x}{\partial \varepsilon^2}(0, z) = z^\top \left(\nabla^2 \bar{V}^* (xe) - \nabla^2 V^*(xe) \right) z. \tag{35}$$

By using (28) and an analogous result for $\nabla^2 \bar{V}(xe)$ (for which $\nabla^2 \bar{I}(v(x)e) = 0$), we complete the proof. ///

Online Appendix

B Proofs on the smooth ambiguity model

Proof of Proposition 1 The twice continuous differentiability of I follows from that of v and ϕ. If $g = ye$ for some $y \in v(T)$, then $p \cdot g = y = I(g)$ for every $p \in \Delta$. Thus, I is normalized. By differentiating both sides of $\phi(I(g)) = \int_\Delta \phi(p \cdot g)\, \mathrm{d}\mu(p)$ with respect to g, we obtain

$$\phi'(I(g))\nabla I(g) = \int_\Delta \phi'(p \cdot g)\, p^\top \mathrm{d}\mu(p) \tag{36}$$

Since $\phi' > 0$ and $\int_\Delta p\, \mathrm{d}\mu(p) \in \boldsymbol{R}^S_{++}$, $\nabla I(g) \in \boldsymbol{R}^S_{++}$. If $g = ye$ for some $y \in v(T)$, the equality is reduced to $\nabla I(ye) = \int_\Delta p^\top \mathrm{d}\mu(p) = (p^I)^\top$.

By differentiating both sides of (36) with respect to g, we obtain

$$\phi''(I(g))\nabla I(g)^\top \nabla I(g) + \phi'(I(g))\nabla^2 I(g)^\top = \int_\Delta \phi''(p \cdot g)\, pp^\top\, \mathrm{d}\mu(p).$$

Thus,

$$\nabla^2 I(g) = \frac{1}{\phi'(I(g))}\left(\int_\Delta \phi''(p \cdot g)\, pp^\top\, \mathrm{d}\mu(p) - \phi''(I(g))\nabla I(g)^\top \nabla I(z)\right). \tag{37}$$

In particular, for every $y \in v(T)$,

$$\nabla^2 I(ye) = \frac{\phi''(y)}{\phi'(y)}\left(\int_\Delta pp^\top\, \mathrm{d}\mu(p) - p^I(p^I)^\top\right) = \frac{\phi''(y)}{\phi'(y)}\int_\Delta (p - p^I)(p - p^I)^\top\, \mathrm{d}\mu(p). \tag{38}$$

Thus, for every $x \in T$ and every $z \in \boldsymbol{R}^S$,

$$z^\top \nabla^2 I(v(x)e)z = \frac{\phi''(v(x))}{\phi'(v(x))}\int_\Delta \left((p - p^I)\cdot z\right)^2\, \mathrm{d}\mu(p). \tag{39}$$

Since $E^\mu[E^\cdot[z]] = E^{p^I}[z]$ by the law of iterated expectation, the law of total variance implies that

$$\int_\Delta \left((p - p^I)\cdot z\right)^2 \mathrm{d}\mu(p) = \int_\Delta \left(E^p[z] - E^{p^I}[z]\right)^2 \mathrm{d}\mu(p) = \mathrm{Var}^\mu[E^\cdot[z]]. \tag{40}$$

Recall that $w(x) = \phi(v(x))$ for every $x \in T$. By differentiating both sides with respect to x, we obtain

$$w'(x) = \phi'(v(x))v'(x). \tag{41}$$

By differentiating both sides with respect to x, we obtain

$$w''(x) = \phi''(v(x))(v'(x))^2 + \phi'(v(x))v''(x). \tag{42}$$

By dividing each side of (42) by the same side of (41), we obtain

$$\frac{w''(x)}{w'(x)} = \frac{\phi''(v(x))}{\phi'(v(x))}v'(x) + \frac{v''(x)}{v'(x)},$$

that is,

$$\frac{\dfrac{\phi''(v(x))}{\phi'(v(x))}}{\dfrac{v''(x)}{v'(x)}}v'(x) = \frac{-\dfrac{w''(x)}{w'(x)}}{-\dfrac{v''(x)}{v'(x)}} - 1.$$

Thus, by (39) and (40),

$$H^x(z) = \frac{-v'(x)\dfrac{\phi''(v(x))}{\phi'(v(x))}\dfrac{\mathrm{Var}^{\mu}[E^{\cdot}[z]]}{\mathrm{Var}^{p^I}[z]}}{-\dfrac{v''(x)}{v'(x)}} = \left(\frac{-\dfrac{w''(x)}{w'(x)}}{-\dfrac{v''(x)}{v'(x)}} - 1\right)\frac{\mathrm{Var}^{\mu}[E^{\cdot}[z]]}{\mathrm{Var}^{p^I}[z]}.$$

///

Proof of Proposition 2 Since v and w have constant coefficients θ and γ of absolute risk aversion,

$$\frac{-\dfrac{w''(x)}{w'(x)}}{-\dfrac{v''(x)}{v'(x)}} - 1 = \frac{\gamma}{\theta} - 1 = \eta.$$

Second, since the second-order belief and the reduced probability are $\mathscr{N}(\bar{R}\mathbf{1}, \Sigma_M)$

and $\mathscr{N}(\bar{R}\mathbf{1}, \Sigma_R)$, $\mathrm{Var}^{\mu}[E^{\cdot}[R_1]] = 0$ and

$$\frac{\mathrm{Var}^{\mu}[E^{\cdot}[R_2]]}{\mathrm{Var}^{p^I}[R_2]} = \frac{\tau\sigma^2}{\sigma^2} = \tau.$$

By Proposition 1, $H^x(R_1) = 0$ and $H^x(R_2) = \eta\tau$.

Hara and Honda (2022, equation (8)) showed that the solution to (17) is given by

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{\bar{R} - R_{\mathrm{f}}}{(\tau\eta + 3/4)\sigma^2\theta} \begin{pmatrix} 1/2 + \tau\eta \\ 1/2 \end{pmatrix}$$

and $x_0 = W - (x_1 + x_2)$. Thus, the share of the wealth invested in the foreign stock in the wealth invested in the two stocks is

$$\frac{1/2}{(1/2 + \tau\eta) + 1/2} = \frac{1}{2(1 + \tau\eta)}.$$

///

C Extension to the non-expected utility case

We assumed throughout this paper that the decision maker has an expected utility function on the set $\Pi(T)$ of lotteries. In this appendix, we extend the definition of the measure of ambiguity aversion to the case where he does not, in three steps. First, we present a general approach to a measure of ambiguity aversion once an ambiguity-neutral preference against which we evaluate the measure is given. When the decision maker has an expected utility function over lottery acts, this benchmark ambiguity-neutral preference is represented by a subjective expected utility function. When he does not, however, the choice of a benchmark ambiguity-neutral preference requires careful consideration. In the second step, we construct such a benchmark ambiguity-neutral preference, which would respect his preference over lottery acts (not representable by expected utility functions) if there were no ambiguity on the state space. Third, we take this ambiguity-neutral preference as the benchmark preference in the general approach to arrive at our extended measure of ambiguity aversion. We then touch on the nature of the twice continuous differentiability that is needed in this argument. Finally, we discuss how the results on the relation to matching probabilities and ambiguity premiums can be extended to the non-expected utility case.

Suppose that there are a utility function $u : \Pi(T) \to \boldsymbol{R}$ that is continuous with respect to the weak topology and strictly increasing with respect to the first-order stochastic dominance, and an aggregator $I : u(\Pi(T))^S \to \boldsymbol{R}$ that is increasing and normalized. For each $f = (f(s))_s \in \Pi(T)^S$, with $f(s) \in \Pi(T)$ for every s, write $u \circ f = (u(f(s)))_s \in u(\Pi(T))^S$. Define a utility function $V : \Pi(T)^S \to \boldsymbol{R}$ by letting $V(f) = I(u \circ f)$ for every $f \in \Pi(T)^S$. This definition of V is the same as the definition of V that we have used so far, except that there need not be an expected utility function over lottery acts, that is, there need not be a Bernoulli utility function $v : T \to \boldsymbol{R}$ such that $u(P) = \int_T v(x)\,\mathrm{d}P(x)$ for every $P \in \Pi(T)$ (or any monotone transformation of the right-hand side).

Let $\bar{u}$, $\bar{I}$ and $\bar{V}$ be just as u, I, and V. We assume that V is at least as ambiguity-averse as $\bar{V}$ in the sense of Epstein (1999). Then, u and $\bar{u}$ represents the same preference over lottery acts and, thus, we can assume that $u = \bar{u}$. Assume that for each $p \in \Delta$, the function $f \mapsto u(p^{-1} \circ f)$ defined on the set T^S of monetary acts is twice continuously differentiable and its derivative in the direction e is strictly positive. Define $v : T \to \boldsymbol{R}$ by equating $v(x)$ equal to the value of u at the degenerate probability on x. Then, v is twice continuously differentiable and $v' > 0$. By the inverse function theorem, it has an inverse v^{-1}. If u were an expected utility function, then v would the corresponding Bernoulli utility function, but it is not so in the case of non-expected utility functions.

Let $x \in T$ and $p = \nabla I(v(x)e)$. Since $V(xe) = v(x) = \bar{V}(xe)$ and V is at least as ambiguity-averse as $\bar{V}$, $p = \nabla \bar{I}(v(x)e)$. Let $z \in \boldsymbol{R}^S$ and assume that $\mathrm{Var}^p[z] > 0$. Then, the measure of ambiguity aversion for V against $\bar{V}$ is defined as

$$H^x(z) = \frac{-z^\top \left(v'(x)\nabla^2 I(v(x)e)\right) z + z^\top \left(v'(x)\nabla^2 \bar{I}(v(x)e)\right) z}{-\dfrac{v''(x)}{v'(x)}\mathrm{Var}^p[z] - z^\top \left(v'(x)\nabla^2 \bar{I}(v(x)e)\right) z}. \tag{43}$$

This differs from the original definition (1) in that the new term $z^\top \left(v'(x)\nabla^2 \bar{I}(v(x)e)\right) z$ is added and subtracted in the numerator and the denominator on the right-hand side. The original definition corresponds to the case where (v is a Bernoulli utility function and) $\bar{I}$ is the expectation operator under p. Hence, $\nabla^2 \bar{I}(v(x)e) = 0$ and the original definition is a special case of the above definition.

We saw that the original measure (1) of ambiguity aversion is invariant to any increasing and affine transformation of v when it is coupled with an increasing transformation of I in such a way that the resulting aggregator is normalized and the resulting utility function represents the same preference as V. We skip the proof, but this extended measure is invariant to any increasing transformation of u when it is coupled with an increasing transformation of I in such a way that the resulting aggregator is normalized and the resulting utility function represents the same preference as V. Note here that as u is a utility function over the set $\Pi(T)$ of lotteries, we should not limit ourselves to affine transformations but accommodate all increasing transformations to check the invariance of the extended measure.

Having defined the measure of ambiguity aversion of V against an arbitrary $\bar{V}$, we now define an ambiguity-neutral utility function $\bar{V}$ via letting $\bar{I}(g) = u(p \circ (v^{-1} \circ g)^{-1})$ for every $g \in v(T)^S$, where $v^{-1} \circ g$ is the function defined on S that takes values $v^{-1}(g(s))$ at each s and $p \circ (v^{-1} \circ g)^{-1}$ is the probability on T induced from the probability p on the state space S by $v^{-1} \circ g$. We then define $\bar{V}(f) = \bar{I}(u \circ f)$. Our choice of $\bar{I}$ and $\bar{V}$ can be understood as follows. For each monetary act $f = (f(s))_s \in T^S$, with $f(s) \in T$ for every s, write $v \circ f = (v(f(s)))_s \in v(T)^S$. For each Anscombe-Aumann act $f \in \Pi(T)^S$, let $\bar{f} \in T^S$ be the monetary act of certainty equivalents of f, that is, $\bar{f}(s) = v^{-1}(u(f(s)))$ for every s. Then, $v \circ \bar{f} = u \circ f$ and, thus, $\bar{V}(f) = \bar{I}(v \circ \bar{f}) = u(p \circ \bar{f}^{-1})$ for every $f \in \Pi(T)^S$. That is, $\bar{V}$ represents a fictitious preference on the set of two-stage lotteries that shares the same preference as V on the set of simple lotteries, regardless of whether they are given at the first stage (in which case they are monetary acts under p) or the second stage (in which case they are lottery acts). In the terminology of Segal (1990), $\bar{V}$ satisfies the time-neutrality. It also satisfies the compound independence axiom, because $\bar{V}(f)$ depend on f only through $\bar{f}$, just as $V(f) = I(v \circ \bar{f})$.[16] For these reasons, it is appropriate to define the measure of ambiguity aversion for V as a deviation from $\bar{V}$.

For this choice of $\bar{I}$, the Hessian $\nabla^2 \bar{I}(v(x)e)$ can be written as

$$\nabla^2 \bar{I}(v(x)e) = \frac{1}{(v'(x))^2} \left(\nabla_f^2 u(p \circ (xe)^{-1}) - v''(x)[p] \right),$$

where $\nabla_f^2 u(p \circ (xe)^{-1})$ is the Hessian of the function $f \mapsto u(p \circ f)$ defined on the set T^S of monetary acts at the constant act xe. By plugging this into (43), we can

[16]There are other utility functions defined on the set $\Pi(T)^S$ of Anscombe-Aumann acts that represents the same preference over lotteries as V. Segal (1990) gave such an example that satisfies the reduction of compound lotteries axiom but violates the compound independence axiom.

also write

$$H^x(z) = \frac{-z^\top \left(v'(x)\nabla^2 I(v(x)e)\right) z + z^\top \left(\dfrac{1}{v'(x)}\nabla_f^2 u(p\circ(xe)^{-1})\right) z - \dfrac{v''(x)}{v'(x)}E^p[z^2]}{-\dfrac{v''(x)}{v'(x)}\mathrm{Var}^p[z] - z^\top \left(\dfrac{1}{v'(x)}\nabla_f^2 u(p\circ(xe)^{-1})\right) z + \dfrac{v''(x)}{v'(x)}E^p[z^2]}. \tag{44}$$

This is the extended measure of ambiguity aversion to the case where the preference over lottery acts may not be represented by any expected utility function.

We simply assumed in the preceding argument that the function $f \mapsto u(p \circ f)$ defined on the set T^S of monetary acts is twice continuously differentiable. Machina (1982) gave a notion of Fréchet differentiability of utility functions over cumulative distribution functions based on the L^1 norm. Allen (1987), instead, used L^2 norm and Wang (1993) used L^p norm with $p \geq 1$. The utility function v over consumption levels (constant acts) that we defined is different from the local expected utility function (Riesz representation of the Fréchet derivative) of Machina (1982), but, at every consumption level x, our function v and the local expected utility function of Machina at the cumulative distribution function degenerate on x share the same derivative. Segal and Spivak (1997) showed that the risk aversion exhibited by a local expected utility function of Machina is of the first order (that is, the derivative of the risk premium at 0 is strictly positive) if and only if the local utility function at the cumulative distribution function degenerate at x is non-differentiable at x. No sufficient condition for twice differentiability of the function $f \mapsto u(p \circ f)$, however, seems to have been given in the literature.

We now explain how the results on the measure of ambiguity aversion in the expected utility case can be extended to the non-expected utility case. As in the first step of defining the measure of ambiguity aversion in the non-expected utility case, take a non-expected utility function u and two aggregators I and $\bar{I}$. Define v, V, and $\bar{V}$ as before. Define, also, $V^* = v^{-1}\circ V$ and $\bar{V}^* = v^{-1}\circ\bar{V}$. Then, (28) is valid for (I, V^*) and also for $(\bar{I}, \bar{V}^*)$. This implies that Corollary 1 can be extended to the non-expected utility case. As for the ambiguity premium, define $\kappa^x(\varepsilon, z)$ and $\bar{\kappa}^x(\varepsilon, z)$ (the risk premium) as in Section 6. Then, (35) holds. Moreover, since $x - \bar{\kappa}^x(\varepsilon, z) = \bar{V}^*(xe + \varepsilon z)$, by differentiating both sides twice and evaluating at

$\varepsilon = 0$, we obtain

$$\frac{\partial^2 \bar{\kappa}^x}{\partial \varepsilon^2}(0, z) = -z^\top \nabla^2 \bar{V}^*(xe) z.$$

Combining these two results and applying the extension of Corollary 1, we can extend Corollary 4 to the non-expected utility case.

As for the matching probabilities, note that Theorem 3 shows that the derivative of the matching probability is equal to the numerator, multiplied by $-1/2$, of the measure (1) of ambiguity aversion in the expected utility case. In the non-expected utility case, we aim at proving that

$$\frac{\partial \rho^x}{\partial \varepsilon}(0, A) = \frac{1}{2}\left(e_A^\top \left(v'(x) \nabla^2 I(v(x)e) \right) e_A - e_A^\top \left(\frac{1}{v'(x)} \nabla_f^2 u(p \circ (xe)^{-1}) \right) e_A + \frac{v''(x)}{v'(x)} p(A) \right), \tag{45}$$

because $E^p[e_A^2] = p(A)$. To do so, define $\rho^x(\varepsilon, A)$ as in (5). Then, (6) is replaced by $u(f^x(\varepsilon, \rho^x(\varepsilon, A))) = V(xe + \varepsilon e_A)$. For each (ε, r), write $w^x(\varepsilon, r) = u(f^x(\varepsilon, r))$. Then, w^x is twice continuously differentiable, and satisfies $w^x(0, r) = v(x)$ for every $r \in [0, 1]$ and

$$w^x(\varepsilon, \rho^x(\varepsilon, A)) = V(xe + \varepsilon e_A) \tag{46}$$

for every ε. Assume that $\partial^2 w^x(0, p(A))/\partial\varepsilon\partial r = v'(x)$. This assumption is met by expected utility functions and also by utility functions that are quadratic in probabilities in the sense of Machina (1982). By differentiating both sides of (46) with respect to ε twice and evaluating at $\varepsilon = 0$, we obtain

$$\begin{aligned} &\frac{\partial^2 w^x}{\partial \varepsilon^2}(0, \rho^x(0, A)) + 2\frac{\partial^2 w^x}{\partial r \partial \varepsilon}(0, \rho^x(0, A)) \frac{\partial \rho^x}{\partial \varepsilon}(0, \rho^x(0, A)) \\ &\quad + \frac{\partial^2 w^x}{\partial r^2}(0, \rho^x(0, A)) \left(\frac{\partial \rho^x}{\partial \varepsilon}(0, \rho^x(0, A)) \right)^2 + \frac{\partial w^x}{\partial r}(0, \rho^x(0, A)) \frac{\partial^2 \rho^x}{\partial \varepsilon^2}(0, \rho^x(0, A)) \\ &= e_A^\top \nabla^2 V(xe) e_A. \end{aligned} \tag{47}$$

Note that $\rho^x(0, A) = p(A)$, $\partial w^x(0, p(A))/\partial r = 0$, $\partial^2 w^x(0, p(A))/\partial r^2 = 0$, and $\partial^2 w^x(0, p(A))/\partial\varepsilon\partial r = v'(x)$. Also, by the definition of $w^x(\varepsilon, r)$,

$$\frac{\partial^2 w^x}{\partial \varepsilon^2}(0, p(A)) = e_A^\top \nabla_f u(p \circ (xe)^{-1}) e_A.$$

Thus, by applying (24) to the right-hand side of (47) and noting that $E^p[e_A^2] =$

$p(A)$, we obtain

$$e_A^\top \nabla_f u(p \circ (xe)^{-1}) e_A + 2v'(x) \frac{\partial \rho^x}{\partial \varepsilon}(0, \rho^x(0, A)) = (v'(x))^2 e_A^\top \nabla^2 I(v(x)e) e_A + v''(x) p(A).$$

By rearranging the above equality, we obtain (45).

The presence of the last two terms on the right-hand side of (45) distinguishes (45) from Theorem 3. It is easy to see that some sort of additional terms is necessary because the matching probability may be different from the benchmark probability even for lottery acts when the preference over lottery acts cannot be represented by any expected utility function. What is more intriguing is that the last two terms are equal to the change in matching probabilities of lottery acts in the following sense: Define $\bar{\rho}^x(\varepsilon, r)$ by

$$\bar{\rho}^x(\varepsilon, r) v(x + \varepsilon) + (1 - \bar{\rho}^x(\varepsilon, r)) v(x) = w^x(\varepsilon, r),$$

then

$$\frac{\partial \bar{\rho}^x}{\partial \varepsilon}(0, p(A)) = \frac{1}{2} \left(e_A^\top \left(\frac{1}{v'(x)} \nabla_f^2 u(p \circ (xe)^{-1}) \right) e_A - \frac{v''(x)}{v'(x)} p(A) \right).$$

Thus,

$$\frac{\partial \rho^x}{\partial \varepsilon}(0, A) = \frac{1}{2} e_A^\top \left(v'(x) \nabla^2 I(v(x)e) \right) e_A - \frac{\partial \bar{\rho}^x}{\partial \varepsilon}(0, p(A)).$$

Thus, if $\partial \bar{\rho}^x(0, p(A))/\partial \varepsilon < 0$, then the first term of the right-hand side, which would be equal to $\partial \rho^x(0, A)/\partial \varepsilon$ in the expected utility case, overestimates the impact of ambiguity aversion on the reduction of matching probabilities as the prize ε goes up, and the second term $\partial \bar{\rho}^x(0, A)/\partial \varepsilon$ needs to be subtracted to correct it.

D A theorem and proofs for the non-differentiable case

Proof of Theorem 5 By the definition (5),

$$\rho^x(\varepsilon, A) = \frac{I\left(v(x)e + (v(x + \varepsilon) - v(x)) e_A\right) - I(v(x)e)}{v(x + \varepsilon) - v(x)}$$

Since I is Gateaux right-differentiable at x in the direction e_A, $\rho^x(\varepsilon, A) \to \mathrm{d_G}I(v(x)e; e_A)$ as $\varepsilon \to 0+$. Since $\mathrm{d_G}I(v(x)e; e_A) = p(A) = \rho^x(0, A)$ for every $p \in \Delta_I^{v(x)}(e_A)$, this shows that $\rho^x(\,\cdot, A)$ is right-continuous at 0. Moreover,

$$\begin{aligned}
&\frac{\rho^x(\varepsilon, A) - \rho^x(0, A)}{\varepsilon} \\
&= \frac{\dfrac{I\left(v(x)e + (v(x+\varepsilon) - v(x))e_A\right) - I(v(x)e)}{v(x+\varepsilon) - v(x)} - \mathrm{d_G}I(v(x)e; e_A)}{\varepsilon} \\
&= \frac{I\left(v(x)e + (v(x+\varepsilon) - v(x))e_A\right) - I(v(x)e) - (v(x+\varepsilon) - v(x))\mathrm{d_G}I(v(x)e; e_A)}{(v(x+\varepsilon) - v(x))^2} \\
&\quad \times \frac{v(x+\varepsilon) - v(x)}{\varepsilon}.
\end{aligned}$$

As $\varepsilon \to 0+$, the far right-hand side converges to $(1/2)v'(x)\mathrm{d}_{\mathrm{G}}^2 I(v(x)e; e_A)$. This completes the proof. ///

We now establish a chain rule for the second derivative of the composite of two twice Peano-differentiable functions. To do so, we need to introduce some notation. For a function F that is Hadamard right-differentiable at a point a in all directions near a point b, if the Hadamard right-derivative $\mathrm{d_H}F(a; \cdot)$ is Hadamard right-differentiable at b in a direction c, then we write the Hadamard right-derivative $\mathrm{d_H}(\mathrm{d_H}F(a; \cdot))(b; c)$ of $\mathrm{d_H}F(a; \cdot)$ at b in the direction c as $\mathrm{d_H}F(a; b; c)$. If $\mathrm{d_H}F(a, \cdot)$ is linear (as in the case where F is differentiable), then

$$\frac{\mathrm{d_H}F(a; b + \varepsilon z) - \mathrm{d_H}F(a; b)}{\varepsilon} = \frac{\varepsilon \mathrm{d_H}F(a; z)}{\varepsilon} = \mathrm{d_H}F(a; z) \to \mathrm{d_H}F(a; c)$$

as ($\varepsilon \to 0$ and) $z \to c$. Thus, $\mathrm{d_H}F(a; b; c) = \mathrm{d_H}F(a; c)$. The need for the definition of $\mathrm{d_H}F(a; b; c)$, therefore, arises from the nonlinearity of $\mathrm{d_H}F(a, \cdot)$, which is often the case for Hadamard right-derivatives of non-differentiable functions. For a mapping G taking values in $\boldsymbol{R}^M$ with $M \geq 1$, if, for each m, the m-th coordinate function G_m is twice Peano-Gateaux right-differentiable at a point a in a direction b, then we denote by $\mathrm{d_G}G(a; b)$ the vector in $\boldsymbol{R}^M$ of which the m-th row is equal to $\mathrm{d_G}G_m(a; b)$ and by $\mathrm{d}_{\mathrm{G}}^2 G(a; b)$ the vector in $\boldsymbol{R}^M$ of which the m-th coordinate is equal to $\mathrm{d}_{\mathrm{G}}^2 G_m(a; b)$.

Theorem 7 *Let N and M be positive integers, A be an open subset of $\boldsymbol{R}^N$, C be an open subset of $\boldsymbol{R}^M$, $F : C \to \boldsymbol{R}$, and $G : A \to C$. Let $a \in A$ and $b \in \boldsymbol{R}^N$. Suppose that the m-th coordinate function G_m of G is twice Peano-Gateaux right-*

differentiable at a in the direction b for every m. Suppose also that F is twice Peano-Hadamard right-differentiable at $G(a)$ in all directions near $\mathrm{d_G}G(a;b)$, and $\mathrm{d_H}F(a;\cdot)$ is Hadamard right-differentiable at $\mathrm{d_G}G(a;b)$ in the direction $\mathrm{d^2_G}G(a;b)$. Define $H = F \circ G$. Then, H is twice Peano-Gateaux right-differentiable at a in the direction b and

$$\mathrm{d^2_G}H(a;b) = \mathrm{d^2_H}F(G(a);\mathrm{d_G}G(a;b)) + \mathrm{d_H}F(G(a);\mathrm{d_G}G(a;b);\mathrm{d^2_G}G(a;b)).$$

While the function F is assumed to be twice Peano-Hadamard right-differentiable, the composite function H is shown to be only twice Peano-Gateaux right-differentiable. The assumption cannot be weakened to twice Peano-Gateaux right-differentiability. Indeed, Delfour (2019, Example 3.9 in Section 3 of Chapter 3) gave an example of a Gateaux differentiable F and an infinitely differentiable function G such that H does not even satisfy one-sided continuity.

The assumption imposed on $\mathrm{d_G}F(a;\cdot)$ in Theorem 7 is satisfied whenever F is differentiable, concave, or convex. Indeed, if F is differentiable, then $\mathrm{d_H}F(a,\cdot) : \boldsymbol{R}^M \to \boldsymbol{R}$ is linear and, hence, differentiable. If F is concave or convex, then, by Theorem 4.6 of Section 4 of Chapter 3 of Delfour (2019), it is Hadamard right-differentiable and the Hadamard right-derivative, $\mathrm{d_H}F(a,\cdot) : \boldsymbol{R}^M \to \boldsymbol{R}$, is concave or convex. Thus, it is Hadamard right-differentiable at every point in all directions.

The following proof is based on the proof of Proposition 3.1 of Ren and Sen (2001) but differs from it in that the limit operation is one-sided ($\varepsilon > 0$) and neither the linearity of the first derivative nor the quadraticity of the second derivative is assumed or implied.

Proof of Theorem 7 By Theorem 3.5 in Section 3 of Chapter 3 of Delfour (2019), H is Gateaux right-differentiable at a in the direction b, and $\mathrm{d_G}H(a;b) = \mathrm{d_G}F(G(a),\mathrm{d_H}G(a;b))$. For each sufficiently small $\varepsilon > 0$, write

$$\Psi(\varepsilon) = \frac{1}{\varepsilon^2}\left(G(a+\varepsilon b) - G(a) - \varepsilon \mathrm{d_G}G(a,b) - \frac{\varepsilon^2}{2}\mathrm{d^2_G}G(a,b)\right) \in \boldsymbol{R}^M.$$

By the definition of the second Peano-Gateaux right-differentiability, $\Psi(\varepsilon) \to 0$ as $\varepsilon \to 0$. Note that

$$\frac{H(a+\varepsilon b) - H(a) - \varepsilon \mathrm{d_G}H(a;b)}{\varepsilon^2} = \frac{K_1(\varepsilon)}{\varepsilon^2} + \frac{K_2(\varepsilon)}{\varepsilon^2}, \tag{48}$$

where

$$K_1(\varepsilon) = H(a+\varepsilon b) - H(a) - \varepsilon \mathrm{d_H} F\left(G(a); \mathrm{d_G} G(a;b) + \frac{\varepsilon}{2}\mathrm{d_G^2} G(a;b) + \varepsilon\Psi(\varepsilon)\right),$$
$$K_2(\varepsilon) = \varepsilon \mathrm{d_H} F\left(G(a); \mathrm{d_G} G(a;b) + \frac{\varepsilon}{2}\mathrm{d_G^2} G(a;b) + \varepsilon\Psi(\varepsilon)\right) - \varepsilon \mathrm{d_G} H(a;b).$$

Since

$$\mathrm{d_G} G(a;b) + \frac{\varepsilon}{2}\mathrm{d_G^2}(a;b) + \varepsilon\Psi(\varepsilon) \to \mathrm{d_G} G(a;b)$$

as $\varepsilon \to 0+$ and since F is twice Peano-Hadamard right-differentiable at $G(a)$ in the direction $\mathrm{d_G} G(a;b)$,

$$\begin{aligned}\frac{K_1(\varepsilon)}{\varepsilon^2} = & \frac{1}{\varepsilon^2}\Bigg(F\left(G(a) + \varepsilon\left(\mathrm{d_G} G(a,b) + \frac{\varepsilon}{2}\mathrm{d_G^2} G(a;b) + \varepsilon\Psi(\varepsilon)\right)\right) - F(G(a)) \\ & - \varepsilon \mathrm{d_H} F\left(G(a); \mathrm{d_G} G(a;b) + \frac{\varepsilon}{2}\mathrm{d_G^2} G(a;b) + \varepsilon\Psi(\varepsilon)\right)\Bigg) \\ & \to \frac{1}{2}\mathrm{d_H^2} F(G(a); \mathrm{d_G} G(a;b))\end{aligned}$$

as $\varepsilon \to 0+$. Since $\mathrm{d_H} F(a;\cdot)$ is Hadamard right-differentiable at $\mathrm{d_H} G(a;b)$ in the direction $\mathrm{d_H^2} G(a;b)$,

$$\begin{aligned}\frac{K_2(\varepsilon)}{\varepsilon^2} = & \frac{\mathrm{d_H} F\left(G(a); \mathrm{d_G} G(a;b) + \varepsilon\left(\frac{1}{2}\mathrm{d_G^2}(a;b) + \Psi(\varepsilon)\right)\right) - \mathrm{d_H} F(G(a); \mathrm{d_G} G(a;b))}{\varepsilon} \\ & \to \frac{1}{2}\mathrm{d_H} F(G(a); \mathrm{d_G} G(a;b); \mathrm{d_G^2} G(a;b))\end{aligned}$$

as $\varepsilon \to 0+$. Hence,

$$\frac{K_1(\varepsilon) + K_2(\varepsilon)}{\varepsilon^2} \to \frac{1}{2}\left(\mathrm{d_H^2} F(G(a); \mathrm{d_G} G(a;b)) + \mathrm{d_H} F(G(a); \mathrm{d_G} G(a;b); \mathrm{d_G^2}(a;b))\right)$$

as $\varepsilon \to 0+$. By (48), this completes the proof. ///

Proof of Corollary 5 Note, first, that

$$\frac{V(xe+\varepsilon z) - V(xe)}{\varepsilon} = \frac{I\left(v(x)e + (v(x+\varepsilon) - v(x))e_A\right) - I(v(x)e)}{v(x+\varepsilon) - v(x)}\,\frac{v(x+\varepsilon) - v(x)}{\varepsilon}.$$

Since I is Gateaux right-differentiable at $v(x)e$ in the direction e_A, as $\varepsilon \to 0+$, the right-hand side converges to $v'(x)\mathrm{d_G} I(v(x)e; e_A)$. Hence, V is Gateaux right-differentiable at xe in the direction e_A and $\mathrm{d_G} V(xe; e_A) = v'(x)\mathrm{d_G} I(v(x)e; e_A)$.

Write

$$\psi(\varepsilon) = \frac{1}{\varepsilon}\left(v(x+\varepsilon) - v(x) - \varepsilon v'(x) - \frac{\varepsilon^2}{2}v''(x)\right).$$

By Taylor's theorem, $\psi(\varepsilon) \to 0$ as $\varepsilon \to 0$. Note, in addition, that

$$\frac{V(xe+\varepsilon e_A) - V(xe) - \varepsilon \mathrm{d_G} V(xe; e_A)}{\varepsilon^2} = \frac{K_1(\varepsilon)}{\varepsilon^2} + \frac{K_2(\varepsilon)}{\varepsilon^2},$$

where

$$\begin{aligned}
K_1(\varepsilon) &= V(xe+\varepsilon e_A) - V(xe) - \varepsilon\left(v'(x) + \frac{\varepsilon}{2}v''(x) + \varepsilon\psi(\varepsilon)\right)\mathrm{d_H} I\left(v(x)e; e_A\right), \\
K_2(\varepsilon) &= \varepsilon\left(v'(x) + \frac{\varepsilon}{2}v''(x) + \varepsilon\psi(\varepsilon)\right)\mathrm{d_H} I\left(v(x)e; e_A\right) - \varepsilon \mathrm{d_G} V(xe; e_A).
\end{aligned}$$

Then, $K_1(\varepsilon)/\varepsilon^2$ is equal to

$$\frac{I\left(v(x)e + \varepsilon\left(v'(x) + \frac{\varepsilon}{2}v''(x) + \varepsilon\psi(\varepsilon)\right)e_A\right) - I(v(x)e) - \varepsilon\left(v'(x) + \frac{\varepsilon}{2}v''(x) + \varepsilon\psi(\varepsilon)\right)\mathrm{d_H} I\left(v(x)e; e_A\right)}{\left(\varepsilon\left(v'(x) + \frac{\varepsilon}{2}v''(x) + \varepsilon\psi(\varepsilon)\right)\right)^2}$$
$$\times\left(\frac{\varepsilon\left(v'(x) + \frac{\varepsilon}{2}v''(x) + \varepsilon\psi(\varepsilon)\right)}{\varepsilon}\right)^2$$

Since I is twice Peano-Gateaux right-differentiable at $v(x)e$ in the direction e_A, this converges to $(1/2)(v'(x))^2\mathrm{d}^2_{\mathrm{G}} I(v(x)e; e_A)$ as $\varepsilon \to 0+$. Note also that

$$\begin{aligned}
\frac{K_2(\varepsilon)}{\varepsilon^2} &= \frac{\varepsilon\left(v'(x) + \frac{\varepsilon}{2}v''(x) + \varepsilon\psi(\varepsilon)\right)\mathrm{d_H} I\left(v(x)e; e_A\right) - \varepsilon v'(x)\mathrm{d_G} I(v(x)e; e_A)}{\varepsilon^2} \\
&= \left(\frac{1}{2}v''(x) + \psi(\varepsilon)\right)\mathrm{d_H} I\left(v(x)e; e_A\right) \to \frac{1}{2}v''(x)\mathrm{d_H} I\left(v(x)e; e_A\right)
\end{aligned}$$

as $\varepsilon \to 0+$. Hence,

$$\frac{V(xe+\varepsilon e_A) - V(xe) - \varepsilon \mathrm{d_G} V(xe; e_A)}{\varepsilon^2} \to \frac{1}{2}(v'(x))^2\mathrm{d}^2_{\mathrm{G}} I(v(x)e; e_A) + \frac{1}{2}v''(x)\mathrm{d_G} I(v(x)e; e_A)$$

as $\varepsilon \to 0+$. Thus, V is twice Peano-Gateaux right-differentiable at xe in the direction e_A and

$$\mathrm{d}^2_{\mathrm{G}} V(xe; e_A) = (v'(x))^2\mathrm{d}^2_{\mathrm{G}} I(v(x)e; e_A) + v''(x)\mathrm{d_G} I(v(x)e; e_A).$$

Since v^{-1} is (twice continuously) differentiable and V is twice Peano-Gateaux right-differentiable at xe in the direction e_A on the set T^S of monetary acts, by applying Theorem 7 to v^{-1} and V in place of F and G, we can show that $\bar{I}$ is twice Peano-Gateaux right-differentiable at xe in the direction e_A, and

$$\begin{aligned}
&\mathrm{d}^2_{\mathrm{G}}\underline{I}(xe; e_A)\\
=&\,\mathrm{d}^2_{\mathrm{H}}v^{-1}(V(xe); \mathrm{d}_{\mathrm{G}}V(xe; e_A)) + \mathrm{d}_{\mathrm{H}}v^{-1}(V(xe), \mathrm{d}_{\mathrm{G}}V(xe; e_A); \mathrm{d}^2_{\mathrm{G}}V(xe; e_A))\\
=&\,(v^{-1})''(v(x))(v'(x)\mathrm{d}_{\mathrm{G}}I(v(x)e; e_A))^2\\
&+(v^{-1})'(v(x))\left((v'(x))^2\mathrm{d}^2_{\mathrm{G}}I(v(x)e; e_A) + v''(x)\mathrm{d}_{\mathrm{G}}I(v(x)e; e_A)\right)\\
=&-\frac{v''(x)}{v'(x)}(\mathrm{d}_{\mathrm{G}}I(v(x)e; e_A))^2 + v'(x)\mathrm{d}^2_{\mathrm{G}}I(v(x)e; e_A) + \frac{v''(x)}{v'(x)}\mathrm{d}_{\mathrm{G}}I(v(x)e; e_A)\\
=&\,v'(x)\mathrm{d}^2_{\mathrm{G}}I(v(x)e; e_A) + \frac{v''(x)}{v'(x)}\mathrm{d}_{\mathrm{G}}I(v(x)e; e_A)(1-\mathrm{d}_{\mathrm{G}}I(v(x)e; e_A))
\end{aligned}$$

Since $\mathrm{d}_{\mathrm{G}}I(v(x)e; e_A) = p(A)$ for any $p \in \Delta_I^{v(x)}(e_A)$ and $\underline{v}'(x) = 1$, this implies that

$$\mathrm{d}^2_{\mathrm{G}}\underline{I}(xe; e_A) = v'(x)\mathrm{d}^2_{\mathrm{G}}I(v(x)e; e_A) + \frac{v''(x)}{v'(x)}p(A)(1-p(A)),$$

that is,

$$v'(x)\mathrm{d}^2_{\mathrm{G}}I(v(x)e; e_A) - \underline{v}'(x)\mathrm{d}^2_{\mathrm{G}}\underline{I}(xe; e_A) = -\frac{v''(x)}{v'(x)}p(A)(1-p(A)).$$

By applying Theorem 5 to ρ and $\underline{\rho}$, we obtain (20). ///

Proof of Proposition 3 Since V is biseparable,

$$\rho(A) = \frac{V(xe + \varepsilon e_A) - V(xe)}{V((x+\varepsilon)e) - V(xe)}$$

for every sufficiently small $\varepsilon > 0$. The right-hand side of this equality is equal to

$$\frac{I(v(x)e + (v(x+\varepsilon) - v(x))e_A) - I(v(x)e)}{v(x+\varepsilon) - v(x)}.$$

Thus,

$$\rho(A) = \frac{I(v(x)e + \delta e_A) - I(v(x)e)}{\delta}. \tag{49}$$

for every sufficiently small $\delta > 0$. Hence, I is Gateaux right-differentiable at $v(x)e$

in the direction e_A, and $\mathrm{d_G}I(v(x)e; e_A) = \rho(A)$. Thus,

$$\begin{aligned}&\frac{I(v(x)e + \delta e_A) - (I(v(x)e) + \mathrm{d_G}I(v(x)e; \delta e_A))}{\delta^2}\\ =&\frac{I(v(x)e + \delta e_A) - (I(v(x)e) + \rho(A)\delta)}{\delta^2} = 0\end{aligned}$$

by (49). Hence, I is twice Peano-Gateaux right-differentiable at $v(x)e$ in the direction e_A, and $\mathrm{d}^2_{\mathrm{G}}I(v(x)e; e_A) = 0$. Since $0 < \rho(A) < 1$ and $\rho(A) = p(A)$ for any $p \in \Lambda_I^{v(x)}(e_A)$, $\mathrm{Var}^p[e_A] > 0$ for any $p \in \Lambda_I^{v(x)}(e_A)$. Thus, $H^x(e_A) = 0$. ///

Proof of Proposition 4 First, we introduce some definitions. For each positive integer n, define $\mathscr{K}_n(y,z)$ as the set of all $J \in \mathscr{J}$ for which there are an $\varepsilon > 0$ and a $w \in \boldsymbol{R}^S$ such that $\varepsilon \leq 1/n$, $\|w - z\| \leq 1/n$, and $J(ye + \varepsilon w) \leq K(ye + \varepsilon w)$ for every $K \in \mathscr{J}(y,z)$. Then, $\mathscr{K}_n(y,z)$ is nonempty and compact for every n, because so is $\mathscr{J}$. Define $\mathscr{K}(y,z) = \bigcap_n \mathscr{K}_n(y,z)$. Then, $\mathscr{K}(y,z)$ is nonempty and compact because so is $\mathscr{K}_n(y,z)$ for every n. Moreover, $\mathscr{K}(y,z) \subseteq \mathscr{J}(y,z)$. Indeed, let $J \in \mathscr{K}(y,z)$ and $K \in \mathscr{J}$. Then, for each n, there are an ε_n and a $z_n \in \boldsymbol{R}^S$ such that $\varepsilon_n \leq 1/n$, $\|z_n - z\| \leq 1/n$, and $J(ye + \varepsilon_n z_n) \leq K(ye + \varepsilon_n z_n)$. Since $J(ye) = y = K(ye)$,

$$\frac{J(ye + \varepsilon_n z_n) - J(ye)}{\varepsilon_n} \leq \frac{K(ye + \varepsilon_n z_n) - K(ye)}{\varepsilon_n}.$$

By taking the limits of both sides as $n \to \infty$, we obtain $\nabla J(ye)z \leq \nabla K(ye)z$. Hence $J \in \mathscr{J}(y,z)$.

Let $(\varepsilon_n)_n$ and $(z_n)_n$ be sequences such that $\varepsilon_n \to 0+$ and $z_n \to z$ as $n \to \infty$. To show the Hadamard right-differentiability, it suffices to show that for any $J \in \mathscr{J}(y,z)$,

$$\frac{J(ye + \varepsilon_n z_n) - J(ye)}{\varepsilon_n} \to \nabla J(ye)z$$

as $n \to \infty$. Then, for each n, there is a $J_n \in \mathscr{J}$ such that $I(ye + \varepsilon_n z_n) = J_n(ye + \varepsilon_n z_n)$. Since $I(ye) = y = J_n(ye)$,

$$\frac{I(ye + \varepsilon_n z_n) - I(ye)}{\varepsilon_n} = \frac{J_n(ye + \varepsilon_n z_n) - J_n(ye)}{\varepsilon_n}.$$

By the mean-value theorem, for each n, there is a $\theta_n \in [0,1]$ such that

$$\frac{J_n(ye+\varepsilon_n z_n) - J_n(ye)}{\varepsilon_n} = \nabla J_n(ye+\theta_n\varepsilon_n z_n)z_n.$$

Since $ye + \theta_n\varepsilon_n z_n \to ye$ and $\mathscr{J}$ is compact with respect to the C^1 compact-open topology, the set $\{\nabla J_n(ye+\theta_n\varepsilon_n z_n)z_n \mid n = 1,2,\dots\}$ is bounded. Thus, it suffices to prove that every convergent subsequence of $(\nabla J_n(ye+\theta_n\varepsilon_n z_n)z_n)_n$ converges to $\nabla J(ye)z$ for any $J \in \mathscr{J}(y,z)$. To ease notation, assume that $(\nabla J_n(ye + \theta_n\varepsilon_n z_n)z_n)_n$ is itself convergent. We can further assume, without loss of generality, that $(J_n)_n$ is convergent, because $\mathscr{J}$ is compact. Denote its limit by J. Then $J \in \mathscr{J}$ and, since $\nabla J_n \to \nabla J$ uniformly on every compact set, $\nabla J_n(ye+\theta_n\varepsilon_n z_n)z_n \to \nabla J(ye)z$.

It now remains to prove that $J \in \mathscr{J}(y,z)$. Since $\varepsilon_n \to 0+$ and $z_n \to z$, for every m, there is an N_m such that $J_n \in \mathscr{K}_m(y,z)$ for every $n > N_m$. Since $\mathscr{K}_m(y,z)$ is compact, $J \in \mathscr{K}_m(y,z)$. Since this is true for every m, $J \in \mathscr{K}(y,z)$. Since $\mathscr{K}(z) \subseteq \mathscr{J}(y,z)$, $J \in \mathscr{J}(y,z)$.

We now move on to the proof of the twice Peano-Hadamard right-differentiability. Let $(\varepsilon_n)_n$ and $(z_n)_n$ be sequences such that $\varepsilon_n \to 0+$ and $z_n \to z$ as $n \to \infty$. It suffices to show that for any $J \in \mathscr{J}(y,z)$,

$$\frac{I(ye+\varepsilon_n z_n) - (I(ye) + \mathrm{d_H}I(ye;z)\varepsilon_n)}{\varepsilon_n^2} \to \frac{1}{2}z^\top \nabla^2 J(ye)z$$

as $n \to \infty$. Then, for each n, there is a $J_n \in \mathscr{G}$ such that $I(ye+\varepsilon_n z_n) = J_n(ye+\varepsilon_n z_n)$. Since $J(ye) = y = J_n(ye)$ and, by the preceding result on the Hadamard right-derivative, $\mathrm{d_H}I(ye;z_n) \le \nabla J_n(ye)z_n$,

$$\begin{aligned}&\frac{I(ye+\varepsilon_n z_n) - (I(ye) + \mathrm{d_H}I(ye;z_n)\varepsilon_n)}{\varepsilon_n^2}\\ \ge\,& \frac{J_n(ye+\varepsilon_n z_n) - (J_n(ye) + \nabla J_n(ye)z_n\varepsilon_n)}{\varepsilon_n^2}.\end{aligned}$$

By Taylor's theorem, for every n, there is a $\theta_n \in [0,1]$ such that

$$\frac{J_n(ye+\varepsilon_n z_n) - (J_n(ye) + \nabla J_n(ye)z_n\varepsilon_n)}{\varepsilon_n^2} = \frac{1}{2}z_n^\top \nabla^2 J_n(ye+\theta_n\varepsilon_n z_n)z_n.$$

Since $\mathscr{J}$ is compact, the set of the values on the right-hand side over all n is bounded. Thus, it has a convergent subsequence. Moreover, every subsequence has a further subsequence of which the corresponding J_n's constitute a convergent sequence. Let J be the limit of such a subsequence. Then, as in the previous paragraph, we can show that $J \in \mathscr{J}(y,z)$. Moreover, the limit of the further subsequence of the values on the right-hand side is equal to $(1/2)z^\top \nabla^2 J(x)z$, which is, by assumption, independent of the choice of $J \in \mathscr{J}(y,z)$. Thus, the values on left-hand side is bounded and the limit of every convergent subsequence of these values is equal to $(1/2)z^\top \nabla^2 J(ye)z$ for any $J \in \mathscr{J}(y,z)$. Thus,

$$\liminf_n \frac{I(ye+\varepsilon_n z_n) - (I(ye) + \mathrm{d_H} I(ye;z_n)\varepsilon_n)}{\varepsilon_n^2} \geq \frac{1}{2} z^\top \nabla^2 J(ye) z$$

for any $J \in \mathscr{J}(y,z)$.

It now remains to prove that

$$\limsup_n \frac{I(ye+\varepsilon_n z_n) - (I(ye) + \mathrm{d_H} I(ye;z_n)\varepsilon_n)}{\varepsilon_n^2} \leq \frac{1}{2} z^\top \nabla^2 J(ye) z \tag{50}$$

for any $J \in \mathscr{J}(y,z)$. To do so, for each n, we now let $J_n \in \mathscr{J}(y,z_n)$, that is, $\mathrm{d_H} I(ye;z_n) = \nabla J_n(ye) z_n$. Since $I(ye+\varepsilon_n z_n) \leq J_n(ye+\varepsilon_n z_n)$ and $I(ye) = y = J_n(ye)$,

$$\begin{aligned}&\frac{I(ye+\varepsilon_n z_n) - (I(ye) + \mathrm{d_H} I(ye;z_n)\varepsilon_n)}{\varepsilon_n^2}\\ &\leq \frac{J_n(ye+\varepsilon_n z_n) - (J_n(x) + \nabla J_n(ye) z_n \varepsilon_n)}{\varepsilon_n^2}.\end{aligned}$$

By Taylor's theorem, for every n, there is a $\theta_n \in [0,1]$ such that

$$\frac{J_n(ye+\varepsilon_n z_n) - (J_n(ye) + \nabla J_n(ye) z_n \varepsilon_n)}{\varepsilon_n^2} = \frac{1}{2} z_n^\top \nabla^2 J_n(ye+\theta_n \varepsilon_n z_n) z_n.$$

Since $\mathscr{J}$ is compact, the set of the values on the right-hand side over all n is bounded. Thus, it has a convergent subsequence. Moreover, every subsequence has a further subsequence of which the corresponding J_n's constitute a convergent sequence. Let J be the limit of such a subsequence. Then, it can be easily shown that $J \in \mathscr{J}(z)$. Moreover, the limit of the further subsequence of the values on the right-hand side is equal to $(1/2)z^\top \nabla^2 J(ye)z$, which is, by assumption, independent of the choice of $J \in \mathscr{J}(y,z)$. Thus, the values on left-hand side is bounded and the limit of every convergent subsequence of these values is equal to $(1/2)z^\top \nabla^2 J(x)z$ for any $J \in \mathscr{J}(y,z)$. Thus, (50) holds. ///

Proof of Example 1 1. Since J is twice continuously differentiable on $\boldsymbol{R}^S \times \Lambda$, if $p_n \to p$, then $J(\,\cdot\, p_n) \to J(\,\cdot\,, p)$ with respect to the C^2 compact-open topology. Since Λ is compact, so is $\mathscr{J}$. If z is a scalar multiple of e, then $z^\top \nabla^2 J_p(ye)z = 0$ for every $p \in \Lambda$ because J_p is normalized. If not, then $\mathscr{J}(y,z)$ is a singleton because Λ is strictly convex. Thus, Assumption 2 is met.

2. Since Π is finite, $\mathscr{J}$ is compact. To show that Assumption 2 is met, let $y \in \boldsymbol{R}$ and $z \in \boldsymbol{R}^S$. Define a partition $\{S_1, S_2, \dots, S_N\}$ of S as follows. There are $z_1, z_2, \dots, z_N$ such that $z(s) = z_n$ for all $s \in S_n$ and n, and $z_1 < z_2 < \cdots < z_N$.

For each p, $\nabla J_p(ye) = p^\top$ and, hence, $\nabla J_p(ye)z = E^p[z]$. Let $p \in \Lambda$. If there are an s and t such that $z(s) < z(t)$ and $p(s) < p(t)$, let τ be the transposition of swapping s and t. Then $\nabla J_{p\circ\tau}(ye)z < \nabla J_p(ye)z$. Hence, for every $p \in \mathscr{J}(y,z)$, if $z(s) < z(t)$, then $p(s) \geq p(t)$. We now show that the converse also holds, that is, if $p(s) \geq p(t)$ whenever $z(s) < z(t)$, then $p \in \mathscr{J}(y,z)$. Indeed, then, $p(s) \geq p(t)$ for all $n = 1, 2, \dots, N$, $s \in S_1 \cup S_2 \cup \cdots \cup S_n$ and $t \in S_{n+1} \cup S_{k+2} \cup \cdots \cup S_N$. Thus, the $|S_1| + |S_2| + \cdots + |S_n|$ largest coordinates of p must coincide with $(p(s))_{s \in S_1 \cup S_2 \cup \cdots \cup S_n}$. Thus,

$$\sum_{s \in S_1 \cup S_2 \cup \cdots \cup S_n} p(s) \geq \sum_{s \in S_1 \cup S_2 \cup \cdots \cup S_n} p(\pi(s))$$

for every $\pi \in \Pi$. Since every $q \in \Lambda$ can be written as $q = p \circ \pi$ for some $\pi \in \Pi$, this implies that the distribution of z under p is first-order stochastically dominated by that under $q \in \Lambda$ for every $q \in \Lambda$. Since $E^p[z] = \nabla J_p(ye)z$, $J_p \in \mathscr{J}(y,z)$.

Note that the above characterization of $\mathscr{J}(y,z)$ implies that $J_p \in \mathscr{J}(y,z)$ if and only if, for each n, $(p(s))_{s \in S_n}$ coincides with the $(|S_1| + \cdots + |S_{n-1}| + 1)$-th to the $(|S_1| + \cdots + |S_{n-1}| + |S_n|)$-th largest coordinates of p. Since

$$E^p[z] = \sum_s p(s)z(s) = \sum_n \left(\sum_{s \in S_n} p(s)\right) z_n,$$
$$E^p[z^2] = \sum_s p(s)(z(s))^2 = \sum_n \left(\sum_{s \in S_n} p(s)\right) z_n^2,$$

they do not depend on the choice of $J_p \in \mathscr{J}(y,z)$. Since

$$z^\top \nabla^2 J_p(ye)z = \frac{\gamma}{2} z^\top \left([p] - pp^\top\right) z = \frac{\gamma}{2}\left(\sum_s p(s)(z(s))^2 - \left(\sum_s p(s)z(s)\right)^2\right),$$

$z^\top \nabla^2 J_p(ye)z$ does not depend on the choice of $J_p \in \mathscr{J}(y,z)$. Hence, Assumption 2 is met. ///

The following lemma characterizes the second Peano-Gateaux derivative of Vunder Assumption 2, and will be used in the proof of Theorem 6. For each $y \in v(T)$, $z \in \boldsymbol{R}^S$, and $w \in \boldsymbol{R}^S$, denote by $\mathscr{J}(y,z,w)$ the set of all $J \in \mathscr{J}(g,z)$ such that $\nabla J(ye)w \leq \nabla K(ye)w$ for all $K \in \mathscr{J}(g,z)$. Note that $\mathscr{J}(y,z,w)$ is different from $\mathscr{J}(y,w)$ in that the latter requires the last inequality to hold for all $K \in \mathscr{J}$, not just for $K \in \mathscr{J}(y,z)$.

Lemma 3 *Under Assumption 2,*

$$\mathrm{d}^2_{\mathrm{G}} V(xe;z) = (v'(x))^2 z^\top \nabla^2 J(v(x)e)z + v''(x)\nabla J(v(x)e)z^2$$

for all $x \in T$, $z \in \boldsymbol{R}^S$, and $J \in \mathscr{J}(v(x), v'(x)z, v''(x)z^2)$, where $z^2 = (z(s)^2)_s \in \boldsymbol{R}^S$.

In the first term of the right-hand side, J can indeed be any element of $\mathscr{J}(v(x),z)$. Hence, the additional constraint that J is an element of $\mathscr{J}(v(x)e, v'(x)z, v''(x)z^2)$ is needed to correctly evaluate the second term.

Proof of Lemma 3 By Proposition 4, I is twice Peano-Hadamard right-differentiable at $v(x)e$ in the direction z. By applying Theorem 7 to I and v in place of F and G and using the fact that the first and second derivatives of the mapping $\varepsilon \mapsto v\circ(xe+\varepsilon z) = (v(x+\varepsilon z(s)))_s$ of $\boldsymbol{R}$ into $\boldsymbol{R}^S$ are equal to $v'(x)z$ and $v''(x)z^2$, we can show that V is twice Peano-Gateaux right-differentiable at xe in the direction z and

$$\mathrm{d}^2_{\mathrm{G}} V(xe;z) = \mathrm{d}^2_{\mathrm{H}} I(v(x)e; v'(x)z) + \mathrm{d}_{\mathrm{H}} I(v(x)e; v'(x)z; v''(x)z^2).$$

By Proposition 4,

$$\mathrm{d}^2_{\mathrm{H}} I(v(x)e; v'(x)z) = (v'(x)z)^\top \nabla^2 J(v(x))(v'(x)z) = (v'(x))^2 z^\top \nabla^2 J(v(x))z$$

for every $J \in \mathscr{J}(v(x)e, v'(x)z)$. It thus remains to show that

$$\mathrm{d_H}I(v(x)e; v'(x)z; v''(x)z^2) = \nabla J(v(x)e)(v''(x)z^2) = v''(x)\nabla J(v(x)e)z^2$$

for every $J \in \mathscr{J}(v(x)e; v'(x)z; v''(x)z^2)$. For this, it suffices to establish the following, more general, fact: $\mathrm{d_H}I(ye; z; w) = \nabla J(ye)w$ for all $y \in v(T)$, $z \in \boldsymbol{R}^S$, $w \in \boldsymbol{R}^S$, and $J \in \mathscr{J}(y, z, w)$.

Denote by $\mathscr{K}(y, z, w)$ the set of all $J \in \mathscr{J}$ for which there are sequences $(\varepsilon_n)_n$ and $(w_n)_n$ such that $\varepsilon_n \to 0+$ and $w_n \to w$ as $n \to \infty$, and $J \in \mathscr{J}(y, z+\varepsilon_n w_n)$ for every n. We claim that $\mathscr{K}(y, z, w) \subseteq \mathscr{J}(y, z, w)$. (The reverse inclusion also holds but it is not necessary in the subsequent proof.) Indeed, let $J \in \mathscr{K}(y, z, w)$. Then, there are sequences $(\varepsilon_n)_n$ and $(w_n)_n$ such that $\varepsilon_n \to 0+$ and $w_n \to w$ as $n \to \infty$, and $J \in \mathscr{J}(x; y+\varepsilon_n z_n)$ for every n. Hence, $\nabla J(ye)(z+\varepsilon_n w_n) \leq \nabla K(ye)(z+\varepsilon_n w_n)$ for all $K \in \mathscr{J}$ and n. As $n \to \infty$, we obtain $\nabla J(ye)z \leq \nabla K(ye)z$ for all $K \in \mathscr{J}$. Thus, $J \in \mathscr{J}(y, z)$. Hence, $\nabla J(ye)z = \nabla K(ye)z$ for all $K \in \mathscr{J}(y, z)$. Since $\nabla J(ye)(z + \varepsilon_n w_n) \leq \nabla K(ye)(z + \varepsilon_n w_n)$, $\nabla J(ye)w_n \leq \nabla K(ye)w_n$ for all $K \in \mathscr{J}(y, z)$ and n. As $n \to \infty$, we obtain $\nabla J(ye)w \leq \nabla K(ye)w$ for all $K \in \mathscr{J}(y, z)$. Thus, $J \in \mathscr{J}(y, z, w)$. Hence, $\mathscr{K}(y, z, w) \subseteq \mathscr{J}(y, z, w)$.

Let two sequences $(\varepsilon_n)_n$ and $(z_n)_n$ be such that $\varepsilon_n \to 0+$ and $z_n \to z$ as $n \to \infty$. By Proposition 4, for every n, there is a $J_n \in \mathscr{J}(y, z + \varepsilon_n w_n)$ such that $\mathrm{d_H}I(ye; z + \varepsilon_n w_n) = \nabla J_n(ye)(z + \varepsilon_n w_n)$. By taking a subsequence if necessary, we can assume that the sequence $(J_n)_n$ is convergent, and denote its limit by J. As in the proof of Proposition 4, we can show that $J \in \mathscr{K}(y, z, w)$. Since $\mathscr{K}(y, z, w) \subseteq \mathscr{J}(y, z, w)$, $J \in \mathscr{J}(y, z, w)$. Moreover, since $\nabla J_n \to \nabla J$ uniformly on every compact set, $\nabla J_n(ye)(z + \varepsilon_n w_n) \to \nabla J(ye)z$. Hence, $\mathrm{d_H}I(ye; z) = \nabla J(ye)z$. On the other hand, by Proposition 4, $\mathrm{d_H}I(ye; z) = \nabla K(ye)z$ for every $K \in \mathscr{J}(y, z)$. In particular, $\mathrm{d_H}I(ye; z) = \nabla J(ye)z$. Thus,

$$\frac{1}{\varepsilon_n}\left(\mathrm{d_H}I(ye; z + \varepsilon_n w_n) - \mathrm{d_H}I(ye; z)\right) = \frac{1}{\varepsilon_n}\left(\nabla J(ye)(z + \varepsilon_n w_n) - \nabla J(ye)z\right) = \nabla J(ye)w_n$$

for every n. As $n \to \infty$, the right-hand side converges to $\nabla J(ye)w$. Thus, so does the left-hand side. Hence, $\mathrm{d_H}I(ye; z; w) = \nabla J(ye)w$. ///

Proof of Theorem 6 1. Let $(\varepsilon_n)_n$ and $(z_n)_n$ be sequences such that $\varepsilon_n \to 0+$ and $z_n \to z$ as $n \to \infty$. Let $q \in \Delta_I^{v(x)}(z)$. To show the right-differentiability of

$\kappa^x(\cdot, z)$ at 0, and its differential is equal to $p \cdot z - q \cdot z$, it is sufficient to prove that

$$\frac{\kappa_p^x(\varepsilon_n, z) - \kappa_p^x(0, z)}{\varepsilon_n} \to p \cdot z - q \cdot z \tag{51}$$

as $n \to \infty$. Suppose that

$$\limsup_n \frac{\kappa_p^{v(x)}(\varepsilon_n, z) - \kappa_p^{v(x)}(0, z)}{\varepsilon_n} > p \cdot z - q \cdot z.$$

Let δ satisfy

$$\limsup_n \frac{\kappa_p^{v(x)}(\varepsilon_n, z) - \kappa_p^{v(x)}(0, z)}{\varepsilon_n} > \delta > p \cdot z - q \cdot z.$$

By taking a subsequence if necessary, we can assume that

$$\frac{\kappa_p^{v(x)}(\varepsilon_n, z) - \kappa_p^{v(x)}(0, z)}{\varepsilon_n} > \delta$$

for every n. Since $\kappa_p^{v(x)}(0, z) = 0$, $\kappa_p^{v(x)}(\varepsilon_n, z) \geq \delta\varepsilon_n$ for every n. Thus,

$$\bar{V}_p((x - \delta\varepsilon_n)e + \varepsilon_n z) > V(xe + \varepsilon_n z),$$

and, hence,

$$\frac{\bar{V}_p((x - \delta\varepsilon_n)e + \varepsilon_n z) - \bar{V}_p(xe)}{\varepsilon_n} > \frac{V(xe + \varepsilon_n z) - V(xe)}{\varepsilon_n}.$$

As $n \to \infty$, the left-hand side converges to $v'(x)(p \cdot z - \delta)$ and the right-hand side converges to $\mathrm{d_H}V(xe; z)$, which is equal to $v'(x)(q \cdot z)$ by Proposition 4. Thus, $p \cdot z - \delta \geq q \cdot z$, but this is a contradiction. Hence,

$$\limsup_n \frac{\kappa_p^{v(x)}(\varepsilon_n, z) - \kappa_p^{v(x)}(0, z)}{\varepsilon_n} \leq p \cdot z - q \cdot z.$$

We can analogously show that

$$\liminf_n \frac{\kappa_p^{v(x)}(\varepsilon_n, z) - \kappa_p^{v(x)}(0, z)}{\varepsilon_n} \geq p \cdot z - q \cdot z.$$

This complete the proof of part 1.

2. The statement on the right-differentiability and the right-derivative follows from part 1 and the assumption that $p \in \Delta_I^{v(x)}(z)$. To show the twice right-differentiability of $\kappa^x(\cdot, z)$ at 0, and its second right-derivative is equal to $-v'(x)\mathrm{d}^2_{\mathrm{H}} I(v(x)e; z)$, let $(\varepsilon_n)_n$ and $(z_n)_n$ be sequences such that $\varepsilon_n \to 0+$ and $z_n \to z$ as $n \to \infty$. Since $\kappa_p^{v(x)}(0, z) = 0$ and $(\partial \kappa_p^{v(x)}/\partial \varepsilon)(0, z) = 0$, it is sufficient to prove that for any $J \in \mathscr{J}(v(x), z)$ and $p \in \Delta_I^{v(x)}(z)$,

$$\frac{\kappa_p^{v(x)}(\varepsilon_n, z)}{\varepsilon_n^2} \to -\frac{1}{2} z^\top \left(v'(x) \nabla^2 J(v(x)e)\right) z \tag{52}$$

as $n \to \infty$. Suppose that

$$\limsup_n \frac{\kappa_p^{v(x)}(\varepsilon_n, z)}{\varepsilon_n^2} > -\frac{1}{2} z^\top \left(v'(x) \nabla^2 J(v(x)e)\right) z.$$

Let δ satisfy

$$\limsup_n \frac{\kappa_p^{v(x)}(\varepsilon_n, z)}{\varepsilon_n^2} > \delta > -\frac{1}{2} z^\top \left(v'(x) \nabla^2 J(v(x)e)\right) z.$$

By taking a subsequence if necessary, we can assume that

$$\frac{\kappa_p^{v(x)}(\varepsilon_n, z)}{\varepsilon_n^2} > \delta$$

for every n. Then, $\kappa_p^{v(x)}(\varepsilon_n, z) > \delta \varepsilon_n^2$ for every n. Thus,

$$\bar{V}_p((x - \delta \varepsilon_n^2)e + \varepsilon_n z) > V(xe + \varepsilon_n z).$$

By the chain rule for derivatives,

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} \bar{V}_p((x - \delta \varepsilon^2)e + \varepsilon z)\Big|_{\varepsilon=0} = v'(x) p \cdot z.$$

By the chain rule for Hadamard right-derivatives,

$$\mathrm{d_H} V(xe; z) = \mathrm{d_H} I(v(x)e; v'(x)z) = v'(x) p \cdot z.$$

Hence,

$$\lim_n \frac{\bar{V}_p((x-\delta\varepsilon_n^2)e+\varepsilon_n z)-\bar{V}_p(xe)}{\varepsilon_n} = v'(x)p\cdot z = \lim_n \frac{V(xe+\varepsilon_n z)-V(xe)}{\varepsilon_n}.$$

Furthermore, for $z^2 = ((z(s))^2)_s \in \boldsymbol{R}^S$,

$$\frac{\mathrm{d}^2}{\mathrm{d}\varepsilon^2}\,\bar{V}_p((x-\delta\varepsilon)e+\varepsilon z)\big|_{\varepsilon=0} = v'(x)p\cdot z^2 - 2\delta v'(x),$$
$$\mathrm{d}_{\mathrm{H}}^2 V(xe;z) = (v'(x))^2\mathrm{d}_{\mathrm{H}}^2 I(v(x)e;z) + v''(x)p\cdot z^2$$

by the chain rule for derivatives and Lemma 3. By Taylor's theorem,

$$\lim_n \frac{\bar{V}_p((x-\delta\varepsilon_n^2)e+\varepsilon_n z)-(\bar{V}_p(xe)+v'(x)p\cdot z)}{\varepsilon_n^2} = \frac{1}{2}\left(v''(x)p\cdot z^2 - 2\delta v'(x)\right).$$

By definition,

$$\lim_n \frac{V(xe+\varepsilon_n z)-(V(xe)+v'(x)p\cdot z)}{\varepsilon_n^2} = \frac{1}{2}\left((v'(x))^2 z^\top \nabla^2 J(v(x)e)z + v''(x)p\cdot z^2\right).$$

Since

$$\begin{aligned}&\lim_n \frac{\bar{V}_p((x-\delta\varepsilon_n^2)e+\varepsilon_n z)-(\bar{V}_p(xe)+v'(x)p\cdot z)}{\varepsilon_n^2}\\ \geq\ &\lim_n \frac{V(xe+\varepsilon_n z)-(V(xe)+v'(x)p\cdot z)}{\varepsilon_n^2},\end{aligned}$$

this implies that

$$v''(x)p\cdot z^2 - 2\delta v'(x) \geq (v'(x))^2 z^\top\nabla^2 J(v(x)e)z) + v''(x)p\cdot z^2,$$

that is,

$$\delta \leq -\frac{1}{2}z^\top\left(v'(x)\nabla^2 J(v(x)e)\right)z.$$

This is contradiction. Hence,

$$\limsup_n \frac{\kappa_p^{v(x)}(\varepsilon_n, z)}{\varepsilon_n^2} \leq -\frac{1}{2}z^\top\left(v'(x)\nabla^2 J(v(x)e)\right)z.$$

We can analogously show that

$$\liminf_n \frac{\kappa_p^{v(x)}(\varepsilon_n, z)}{\varepsilon_n^2} \geq -\frac{1}{2} z^\top \left(v'(x) \nabla^2 J(v(x)e) \right) z.$$

Thus, (52) follows. ///

References

[1] Abdellaoui, M., A. Baillon, L. Placido, and P. P. Wakker, The rich domain of uncertainty: source function and their experimental implementation, *American Economic Review* 101, 695–723.

[2] S. N. Afriat, 1967, The construction of utility functions from expenditure data, *International Economic Review* 8, 67–77.

[3] Allen, B., 1987, Smooth preferences and the approximate expected utility hypothesis, *Journal of Economic Theory* 41, 340–355.

[4] Anscombe, F. J., and R. J. Aumann, 1963, A definition of subjective probability, *Annals of Mathematical Statistics* 34, 199–205.

[5] Baillon, A., H. Bleichrodt, C. Li, and P. P. Wakker, 2021, Belief hedges: Measuring ambiguity for all events and all models, *Journal of Economic Theory* 198, 105353. DOI: `https://doi.org/10.1016/j.jet.2021.105353`

[6] Baillon, A., Z. Huang, A. Selim, and P. P. Wakker, 2018, Measuring ambiguity attitudes for all (natural) events, *Econometrica* 86, 1839–1858. DOI: `https://doi.org/10.3982/ECTA14370`

[7] Bossaerts, P., P. Ghirardato, S. Guarnaschelli, and W.R. Zame, 2010, Ambiguity in asset markets: Theory and experiment, *Review of Financial Studies* 23: 1325–1359.

[8] Cerreia-Vioglio, S., P. Ghirardato, F. Maccheroni, M. Marinacci, and M. Siniscalchi, 2011, Rational preferences under uncertainty, *Economic Theory* 48, 341–375.

[9] Cerreia-Vioglio, S., F. Maccheroni, and M. Marinacci, 2022, Robust mean-variance approximation, manuscript.

[10] Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio, 2011, Uncertainty averse preferences, *Journal of Economic Theory* 146, 1275–1330.

[11] Chen, H., Ju, N., and J. Miao, 2014, Dynamic asset allocation with ambiguous return predictability. *Review of Economic Dynamics* 17, 799–823.

[12] Collard, F., S. Mukerji, K. Sheppard, and J. Tallon, 2018, Ambiguity and the historical equity premium, *Quantitative Economics* 945–993.

[13] Michel C. Delfour, 2019, *Introduction to Optimization and right-derivative Calculus*, Second Edition, Philadelphia, PA, Society for Industrial and Applied Mathematics.

[14] Debreu, G., 1976, Least concave utility functions, *Journal of Mathematical Economics* 3, 121–129.

[15] Dimmock, S. G., R. Kouwenberg, and P. P. Wakker, 2016, Ambiguity attitudes in a large representative sample, *Management Science* 62, 1363–1380.

[16] Dow, J., and S. R. C. Werlang, 1992, Uncertainty aversion, risk aversion, and the optimal choice of portfolio, *Econometrica* 60, 197–204.

[17] Ellsberg, D., 1961, Risk, ambiguity, and the Savage axiom, *Quarterly Journal of Economics* 75, 643–669.

[18] Epstein, L. G., 1999, A definition of uncertainty aversion, *Review of Economic Studies* 66, 579–608.

[19] Epstein, L. G., and T. Wang, 1994, Intertemporal asset pricing under Knightian uncertainty, *Econometrica* 62, 283–322.

[20] Evans, M. J., and C. E. Weil, 1981-82, Peano derivatives: A survey, *Real Analysis Exchange* 7, 5–23.

[21] Gallant, A. R., M. Jahan-Parvar, H. Liu. 2019. Does smooth ambiguity matter for asset pricing? *Review of Financial Studies* 32, 3617–3666.

[22] Ghirardato, P., and M. Marinacci, 2002, Ambiguity made precise: A comparative foundation, *Journal of Economic Theory* 102, 251–289.

[23] Ghirardato, P., F. Maccheroni, and M. Marinacci, 2004, Differentiating ambiguity and ambiguity aversion, *Journal of Economic Theory* 118, 133–173.

[24] Ghirardato, P., and M. Siniscalchi, 2012, Ambiguity in the small and in the large, *Econometrica* 80, 2827–2847.

[25] Gilboa, I., and D. Schmeidler. 1989. Maxmin Expected Utility with Non-Unique Prior, *Journal of Mathematical Economics* 18, 141–153.

[26] Hansen, L. P., and T. J. Sargent, 2001, Acknowledging misspecification in macroeconomic theory, *Review of Economic Dynamics* 4, 519–535.

[27] Hara, C., and T. Honda, 2022, Implied ambiguity: mean-variance efficiency and pricing errors, *Management Science* 68, 4246–4260.

[28] Hardy, G. H., J. E. Littlewood, and G. Polya, 1952, *Inequalities*, 2nd edition, Cambridge University Press.

[29] Hirsch, M. W., 1976, *Differential Topology*, Springer-Verlag.

[30] Jahan-Parvar, M. R., and H. Liu, 2014, Ambiguity aversion and asset prices in production economies, *Review of Financial Studies* 27, 3060–3097.

[31] Jewitt, I., and S. Mukerji, 2017, Ordering ambiguous acts, *Journal of Economic Theory* 171, 213–267. DOI: `https://doi.org/10.1016/j.jet.2017.07.001`

[32] Ju, N., and J. Miao, 2012, Ambiguity, learning, and asset returns. *Econometrica* 80, 559–591.

[33] Klibanoff, P., M. Marinacci, and S. Mukerji. 2005. A smooth model of decision making under ambiguity, *Econometrica* 73, 1849–1892.

[34] Kreps, D. M., 1988, *Notes on the Theory of Choice*, Westview Press (Boulder, CO).

[35] Lewis, K. K. , 1999, Trying to explain home bias in equities and consumption, *Journal of Economic Literature* 37, 571–608.

[36] Maccheroni, F., M. Marinacci, and D. Ruffino. 2013. Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis. *Econometrica* 81:1075–1113.

[37] Maccheroni, F., M. Marinacci, and A. Rustichini, 2006, Ambiguity aversion, robustness, and the variational representation of preferences, *Econometrica* 74, 1447–1498.

[38] Machina Machina, M. J., 1982, "Expected utility" analysis without the independence axiom, *Econometrica* 50, 277-323.

[39] Machina, M. J., and D. Schmeidler, 1992, A more robust definition of subjective probability, *Econometrica* 60, 745–780.

[40] Mas-Colell, A., 1985, *General Equilibrium Theory: A differentiable approach*, Cambridge (UK), Cambridge University Press.

[41] Mas-Colell, A., M. Whinston, and J. R. Green, 1995, *Microeconomic Theory*, Oxford University Press.

[42] Mukhopadhyay, S. N., 2012, *Higher Order Derivatives*, Taylor and Francis (Boca Raton, FL).

[43] Pratt, J. W., 1964, Risk aversion in the small and in the large, *Econometrica* 32, 122–136.

[44] Ren, Jian-Jian, and Pranab Kumar Sen, 2001, Second order Hadamard differentiability in statistical applications, *Journal of Multivariate Analysis* 77, 187–228.

[45] Rigotti, L., C. Shannon, and T. Strzalecki, 2008, Subjective beliefs and ex ante trade, Econometrica 76, 1167–1190.

[46] R. Tyrrell Rockafellar, 1970, *Convex Analysis*, Princeton University Press (Princeton, NJ).

[47] R. Tyrrell Rockafellar, 2000, Second-order convex analysis, *Journal of Nonlinear and Convex Analysis* 1, 1–16.

[48] R. Tyrrell Rockafellar and Roger J-B Wets, 1998, *Variational Analysis*, Berlin, Springer.

[49] Schmeidler, D., 1989, Subjective probability and expected utility without additivity, *Econometrica* 57, 571–587.

[50] Segal, U., 1990, Two-stage lotteries without the reduction axiom, *Econometrica* 58, 349–377.

[51] Segal, U., and A. Spivak, 1990, First order versus second order risk aversion, *Journal of Economic Theory* 51, 111–125.

[52] Segal, U., and A. Spivak, 1997, First-order risk aversion and non-differentiability, *Economic Theory* 9, 179–183.

[53] Wang, T., 1993, L_p-Fréchet differentiable preference and "local utility" analysis, *Journal of Economic Theory* 61, 139–159.

[54] Wang, F., 2019, Comparative ambiguity aversion, manuscript, HEC Paris.

[55] Yaari, M. E., 1969, Some remarks on measures of risk aversion and on their uses, *Journal of Economic Theory* 1, 315–329.